Essays on Non-Stationary Panel Analysis

Laura Surdeanu

PhD in Economics

Thesis title:

Essays on Non-Stationary Panel Analysis

PhD student:

Laura Surdeanu

Advisors:

Josep Lluís Carrion-i-Silvestre

Tutor:

Josep Lluís Carrion-i-Silvestre

Date:

October 2013

Universitat de Barcelona

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Josep Lluís Carrion-i-Silvestre. He believed in me even when I did not believe in myself. I am extremely grateful to for his help, support and encouragement. THANK YOU!

My love for econometrics started in graduate school at UNT when I took the introductory course in econometrics with Dr. Margie Tieslau. I loved her teaching so much that I registered for all econometrics and econometrics-related courses available. So thank you, Dr. Tieslau!

I would like to thank my parents, Neli and Grigore, for support beyond reason. Horatiu, the best brother in the world, and Miriam always stayed positive. Brindusa, Codruta and Lili helped me stayed sane through these years. I would not have been here without my best friend and husband so thank you, Mihai!

Contents

List of Tables

List of Figures

Introduction

This thesis consists of three self-contained essays on non-stationary panel data. We propose novel approaches to both cointegration and unit root analysis in panel data models. The main contribution of this thesis is allowing for the presence of cross-section dependence through the specification of an approximate common factor model. Early studies assumed that time series in the panel data were either independent or that cross-section dependence could be controlled by including time effects. In macroeconomic, microeconomic and financial applications, cross-section dependence is more a recurrent than a rare characteristic and it is usually caused by the presence of common shocks (oil price shocks or financial crises) or the existence of local productivity spillover effects. Ignoring these factors can lead to spurious statistical inference. More exactly, in the case of unit root testing, the unaccounted cross-section dependence might lead one to conclude that panel data is actually I(0) stationary when in fact it might be I(1) non-stationary. Similarly, the panel data cointegration test statistics might indicate than there are more cointegrating relations than there exist. Thus, recent studies proposed several alternatives to overcome this limitation. One popular approach is the factor structure applied to the error process, an approach that we employ throughout this thesis.

In the first essay we extend the univariate Carrion-i-Silvestre, Kim and Perron (2009) GLS-based unit root tests with multiple structural breaks to panel data. The proposed statistics are general enough that they allow for cross-section dependence and multiple structural breaks in both the level and the trend of the units of the panel. We evaluate the finite-sample properties of these statistics via Monte Carlo simulations. Our simulation study shows that the panel tests perform well, especially for the cases of known structural breaks. We apply these statistics to a panel of annual data covering the period 1870-2008 for 19 OECD countries. We find strong evidence in favor of I(0) stationarity when we apply the unit root tests to idiosyncratic component. However, the empirical analysis also shows that the I(1) non-stationarity of the real per capita GDP is captured by the common factor.

In the second essay we propose a test statistic to determine the cointegration rank of VAR processes both in a unit-by-unit analysis and in a panel data framework. The cross-section dependence is accounted for through the specification of a common factor model, which covers situations where there is cointegration among the cross-section dimension. We perform a Monte Carlo experiment in order to investigate the small-sample properties of the proposed panel statistic and the simulation results indicate a good performance of the tests in terms of empirical size and power. We show that in some cases not accounting for common factors when they are present can lead to overestimating the cointegrating rank. We apply our proposed tests to two empirical applications using the variables involved in the money demand equation and the monetary exchange model. The money demand model detects two stochastic trends while the monetary exchange model detects three stochastic trends.

In the third essay of this dissertation we investigate the cointegration relation between output, physical capital, human capital, public capital and labor for 17 Spanish regions observed over the period 1964-2000. The novelty of our approach is that we allow for cross-section dependence between the members of the panel using a common factor model. This is interesting because we allow the model specification to capture unobservable variables (technological progress, total factor productivity) to be proxied by the common factors, something that has not been widely addressed in the literature. To see if the variables are cointegrated or not, we employ two different techniques at the panel level. More exactly, we compare the statistics from the single-equation method of Westerlund (2008) and Banerjee and Carrion-i-Silvestre (2011, 2013) with those from the VAR framework of Carrion-i-Silvestre and Surdeanu (2011). Moreover, using the VAR method, we identify at least one common cointegrating relation among output, physical capital, human capital, public capital and labor. Finally, we use several estimators to estimate the long-run relation between these variables.

Chapter 1

Panel GLS Unit Root Tests and Common Factors

1.1 Introduction

Ever since Perron (1989) published his seminal paper on unit root tests and structural breaks, the interest on this field rose considerably. Failure to account for structural breaks leads to size distortions and potentially biased estimates of the parameters in which unit root tests build upon. This line of research began when the author showed that the Dickey-Fuller (DF) statistic is biased towards the null hypothesis of unit root if there is an external shock affecting the slope of the time series. He proposed different tests, consistent under both the null and the alternative hypotheses, with the condition that the structural break date is known *a priori*. This condition was later criticized by Christiano (1992) who argued that the date of the structural break is chosen based on pre-test examination of the data so that the analysis becomes conditional on the decision of the practitioner. As a result, the following studies took into consideration that the break date was determined endogenously and examples of such studies for one unknown structural break include Zivot and Andrews (1992), Perron and Vogelsang (1992), Banerjee, Lumsdaine and Stock (1992) and Perron (1997).

The next line of research for univariate case includes studies for two or more structural breaks. Lumsdaine and Papell (1997) argued that unit root test results are sensitive to the number of assumed structural breaks and they extended the Zivot and Andrews (1992) analysis in order to account for two endogenously determined structural breaks. Note that the framework of Zivot and Andrews (1992) allows the structural break only under the alternative hypothesis. For the case when a series has a structural break under the null hypothesis, the rejection of the null hypothesis might indicate that the time series is I(0) stationary with structural breaks when in fact it is I(1) non-stationary with structural breaks. Clemente, Montañés and Reyes (1998) generalize the proposal in Perron and Vogelsang (1992) considering two structural breaks affecting the level of the time series, both under the null and alternative hypotheses. Lee and Strazicich (2003) went one step further and proposed a minimum Lagrange Multiplier (LM) unit root test that allows for two unknown structural breaks in level and trend under both the null and the alternative hypotheses. The authors compared their results with those of Lumsdaine and Papell (1997) using the same dataset and found that the statistic proposed by Lumsdaine and Papell (1997) tends to reject the null hypothesis of unit root more than the LM test.

Other researchers extended the previous analyses to more than two endogenously determined structural breaks. Ohara (1999) and Kapetanios (2005) generalized the Zivot and Andrews (1992) methodology to allow for m structural breaks but only under the alternative hypothesis of I(0) – see Perron (2006) for a complete overview. Carrion-i-Silvestre, Kim and Perron (2009, CKP hereafter) extended the unit root tests based on the GLS detrending procedure proposed in Ng and Perron (2001). Their framework allows for multiple structural breaks under both the null and alternative hypotheses. Also, they allow for structural breaks in both the level and the slope of the time trend of the series. It has been shown that these statistics have better size and power properties than those that allow for structural breaks only under the I(0) stationary alternative. In a recent paper, Westerlund (2012) extended the Amsler and Lee (1995) unit root test to allow for multiple structural breaks in the level of the data. Like in the CKP paper, the break points are allowed under both the null and alternative hypotheses.

With the increasing development of panel data methods, the researchers extended the univariate analysis to panel data framework. However, these types of extensions are relatively limited. The first generation of panel data unit root and stationarity tests with structural breaks assumed that the units of the panel are cross-section independent. For example, Carrion-i-Silvestre, del Barrio Castro and López-Bazo (2001) extended the panel data Dickey-Fuller (DF) unit root test in Harris and Tzavalis (1999) considering one structural break in the level of the time series. Another example of panel unit root test is the work of Im, Lee and Tieslau (2005) who extended the LM-based test while allowing for up to two level shifts.

Finally, Carrion-i-Silvestre, del Barrio Castro and López-Bazo (2005) developed a panel data stationarity test allowing for multiple structural breaks in both the intercept and/or the slope of the time series.

Note that these studies did not account for the cross-sectional dependence that plagued the earlier panel data studies, the so-called first-generation panel tests – see Breitung and Pesaran (2008) and Banerjee and Wagner (2009) for recent overviews of non-stationary panel data analysis. In the recent years it was shown that the cross-section independence assumption is not realistic especially in country or regional studies. For example, one important problem that we have to deal with nowadays is the increase in oil prices. As a result, many macroeconomic variables from one country are very close related with those from a neighboring country. That is, due to a common shock, the cross-section variables of the panel of countries are dependent on one another. If the cross-section dependence is not accounted for, it can cause biased and inconsistent estimates. So the next necessary step is taking into account the dependence between cross-sections while still allowing for structural breaks.

One example of such studies is that of Bai and Carrion-i-Silvestre (2009). The authors treated the cross-section dependence by using common factors originally proposed by Bai and Ng (2004). They proposed as a panel unit root test the square of the modified MSB test defined by Stock (1999) while allowing for multiple structural breaks. The test is invariant in the limit only to level shifts but not to structural breaks affecting the slope of the time trend. Therefore, the authors also proposed a simplified MSB test statistic that is invariant to both level and slope shifts, although the limiting distribution still depends on the number of structural breaks. Another example of a related work is that of Tam (2006) who proposed panel unit root tests that are an extension of the LM-based test and the combination tests of Maddala and Wu (1999) and Choi (2001). The author handles the impact of cross-section dependence by means of bootstrapping. Another study that extends a LM-based unit root test to panel data is the one by Westerlund (2012), who allows for multiple structural breaks in the level of the data. In order to estimate the number of the structural breaks and their location, Westerlund (2012) suggests a procedure based on outlier detection that is valid under both the null and alternative hypotheses, does not require *a priori* knowledge about the number or the location of the structural breaks, and is robust to cross-sectional dependence captured by common factors. However, the disadvantage of the procedure is that the underlying error terms must be normally distributed. Im, Lee and Tieslau (2012) proposed another LM-based panel unit root test that allows for heterogeneous structural breaks in both the level and the slope of the time trend of the series. Their statistic depends only on the number of structural breaks but not on their size or location, and is invariant to nuisance parameters. The authors apply the cross-sectionally augmented ADF (CADF) regression of Pesaran (2007) to their tests as one possible means of correcting for cross-section dependence. Finally, Lee and Wu (2012) suggested a panel unit root test based on the generalized CADF procedure proposed by Pesaran (2007). They incorporate a single-frequency-component Fourier function that is used to approximate the unknown multiple structural breaks. The cross-sectional dependence is modeled by an unobservable I(0) stationary common factor.

However, the panel unit root literature that use the GLS detrending is relatively limited. Very recently, Westerlund (2013) noted that "the only formal treatment known to us is the simulation study of Lopez (2009)". Lopez (2009) proposed a pooled panel unit root test based on GLS detrending that allows for serial and contemporaneous correlation. While the author accounts for cross-sectional dependence by estimating the residual covariance matrix, she does not allow for structural breaks. To the best of our knowledge, none of the existing panel studies that use GLS detrending in their estimation allows for structural breaks under both the null and the alternative hypotheses and in both the intercept and slope of the series.

In this chapter, we propose several panel data unit root tests that are based on the GLS detrending procedure. The statistics are the extension of univariate CKP statistics to panel data. The new tests allow for multiple structural breaks that affect either the level and/or the slope of the time trend. Like in the CKP study, we allow for structural breaks under both the null and the alternative hypotheses. Moreover, we deal with the cross-section dependence through the use of common factors. We then evaluate the finite-sample properties of our statistics via Monte Carlo simulations. Our simulation study shows that the tests perform well for both cases of known and unknown structural breaks. Finally, we apply the proposed tests to a panel of annual per capita real GDP over the period 1870-2008 for 19 OECD countries.

The structure of this chapter is as follows. In Section 1.2 we describe the model, while Section 1.3 presents the unit root test statistics that are investigated. Section 1.4 presents the case for unknown structural breaks. The proposed panel statistics are shown in Section 1.5. Section 1.6 summarizes the Monte Carlo simulation results and the empirical application is carried out in Section 1.7. Finally, this chapter concludes in Section 1.8.

1.2 The model

Let us consider the data generating process (DGP) given by the following system of equations:

$$\begin{aligned} y_{i,t} &= d_{i,t} + F_t' \delta_i + e_{i,t} && (1.2.1) \\ (I-L) F_t &= C(L) w_t && (1.2.2) \\ (1-\theta_i L) e_{i,t} &= B_i(L) \varepsilon_{i,t}, && (1.2.3) \end{aligned}$$

$i = 1, \ldots, N$, $t = 1, \ldots, T$, where the stochastic process $y_{i,t}$ is decomposed as the sum of a deterministic term $d_{i,t}$, a common factor component $F_t' \delta_i$ and the idiosyncratic stochastic component $e_{i,t}$. In this framework the cross-section dependence among time series in the panel data is driven by an approximate common factor model as in Bai and Ng (2002, 2004). F_t denotes a $(r \times 1)$ vector of unobserved common factors and δ_i is a $(r \times 1)$ vector of factor loadings. Note that F_t can be I(0), I(1) or a combination of I(0) and I(1) common factors depending on the rank of $C(1)$. For example, if $C(1) = 0$ then F_t is I(0). If the rank of $C(1)$ is r_1 then they are r_1 common stochastic trends and $r - r_1$ I(0) common factors. If $C(1)$ has full rank then F_t is $I(1)$.

Let $M < \infty$ be a generic positive number, independent of T and N and let $\|A\| = trace\,(A'A)^{1/2}$. We follow Bai and Ng (2004) and define the following assumptions:

Assumption A: (i) for non-random δ_i, $\|\delta_i\| \leq M$; for random δ_i, $E\,\|\delta_i\|^4 \leq M$. (ii) $\frac{1}{N}\sum_{i=1}^{N} \delta_i \delta_i' \xrightarrow{p} \Sigma_\delta$, a $(r \times r)$ positive definite matrix.

Assumption B: (i) $w_t \sim iid\,(0, \Sigma_w)$, $E\,\|w_t\|^4 \leq M$. (ii) $Var\,(\Delta F_t) = \sum_{j=0}^{\infty} C_j \Sigma_w C_j' > 0$. (iii) $\sum_{j=0}^{\infty} j \left\|C_j\right\| < M$. (iv) $C(1)$ has rank r_1, $0 \leq r_1 \leq r$.

Assumption C: (i) for each i, $\varepsilon_{i,t} \sim iid\,(0, \Sigma_{\varepsilon_i})$, $E\,|\varepsilon_{i,t}|^8 \leq M$. (ii) $Var\,(\Delta\varepsilon_{i,t}) = \sum_{j=0}^{\infty} B_{i,j} \Sigma_{\varepsilon_i} B_{i,j}' > 0$. (iii) $\sum_{j=0}^{\infty} j \left\|B_{i,j}\right\| < M$.

Assumption D: $\varepsilon_{i,t}$, w_t and δ_i are mutually independent.

Assumption E: $E\,\|F_0\| \leq M$, and for every $i = 1, \ldots, N$, $E\left|e_{i,0}\right| \leq M$.

Assumptions A and B imply the existence of r common factors. Assumption B permits a combination of I(0) and I(1) common factors in the model. Assumption C(i) allows some weak correlation in $(1 - \theta_i L)\, e_{i,t}$, while C(ii) and C(iii) allow weak cross-section correlation. Assumption D states that the errors $\varepsilon_{i,t}$, w_t and δ_i are mutually independent across i and t. Assumption E defines the initial conditions.

The definition of the deterministic component in (1.2.1) gives rise to three different models. Model 0, where the multiple structural breaks occur in the intercept, is known as the "level shift" model. Model I is known as the "slope change" model and allows for structural breaks only in the slope of the time trend. Finally, Model II allows for multiple structural breaks in both the intercept and the slope of the time trend. These models can be parameterized as

$$d_{i,t} = \sum_{j=0}^{m} z'_{i,t}\left(T^0_{i,j}\right)\psi_{i,j} \equiv z'_{i,t}\left(\lambda^0_i\right)\psi_i,$$

where $z_{i,t}\left(\lambda^0_i\right) = \left[z'_{i,t}\left(T^0_{i,0}\right),\dots,z'_{i,t}\left(T^0_{i,m}\right)\right]'$, $\psi_i = \left(\psi'_{i,0},\dots,\psi'_{i,m}\right)'$ and

$$z_{i,t}\left(T^0_{i,j}\right) = \begin{cases} DU_{i,t}\left(T^0_{i,j}\right) & \text{for Model 0} \\ DT_{i,t}\left(T^0_{i,j}\right) & \text{for Model I} \\ \left(DU_{i,t}\left(T^0_{i,j}\right), DT_{i,t}\left(T^0_{i,j}\right)\right)' & \text{for Model II} \end{cases},$$

and

$$\psi_{i,j} = \begin{cases} \mu_{i,j} & \text{for Model 0} \\ \beta_{i,j} & \text{for Model I} \\ \left(\mu_{i,j},\beta_{i,j}\right)' & \text{for Model II} \end{cases},$$

$0 \le j \le m$, with $DU_{i,t}\left(T^0_{i,j}\right) = 1$ and $DT_{i,t}\left(T^0_{i,j}\right) = t - T^0_{i,j}$ for $t > T^0_{i,j}$ and 0 otherwise, where $T^0_{i,j} = \left\lfloor T\lambda^0_{i,j}\right\rfloor$ represents the true break date for the i-th individual – $\lfloor\cdot\rfloor$ denotes the integer part of the element between brackets – λ^0_i is a $(m\times 1)$-vector with the true break fractions, with the convention that $T^0_{i,0} = 0\ \forall i$. It is worth noticing that the use of "0" as a superscript indicates that the structural breaks are known a priori – the case of unknown structural breaks is addressed below.[1]

As for the break dates, the model that we specify assumes that they admit certain degree of heterogeneity through the definition of

$$T^0_{i,j} = T^0_j + v_{i,j}, \tag{1.2.4}$$

[1]Although we deal with three different specifications involving structural breaks, our setup can also be particularized to the case of no structural breaks considering $m = 0$.

with $v_{i,j} \sim iid\left(0, \sigma_{i,j}^2\right)$ $\forall i, j$, i.e., the break dates are assumed to depart from a common break dates up to a bounded quantity. Note that in the limit, the fraction parameters are common to all individuals since

$$\begin{aligned}
\lambda_{i,j}^0 &= T_{i,j}^0/T \\
&= T_j^0/T + O_p\left(T^{-1}\right) \xrightarrow{p} \lambda_j^0,
\end{aligned}$$

where $\xrightarrow{p}$ denotes convergence in probability. Consequently, although in finite samples the break dates are allowed to be mildly heterogeneous across individuals, in the limit the break fraction vector is common to all individuals, i.e., $\lambda_i^0 \xrightarrow{p} \lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)'$. It is worth mentioning that it is also possible to impose here the constraint that the break points are common to all individuals in finite samples if we set $v_{i,j} = 0\ \forall i, j$ in (1.2.4), so that $\lambda_{i,j}^0 = \lambda_j^0\ \forall i, j$.

The GLS detrended unit root statistics use the transformed data $y_{i,t}^{\bar{\alpha}}$ and $z_{i,t}^{\bar{\alpha}}\left(\lambda_i^0\right)$, which is defined as $y_{i,1}^{\bar{\alpha}} = y_{i,1}$ and $z_{i,1}^{\bar{\alpha}}\left(\lambda_i^0\right) = z_{i,1}\left(\lambda_i^0\right)$ for $t = 1$, and $y_{i,t}^{\bar{\alpha}} = (1 - \bar{\alpha}L)\, y_{i,t}$ and $z_{i,t}^{\bar{\alpha}}\left(\lambda_i^0\right) = (1 - \bar{\alpha}L)\, z_{i,t}\left(\lambda_i^0\right)$ for $t = 2, \ldots, T$, $i = 1, \ldots, N$, with $\bar{\alpha} = 1 + \bar{c}\left(\lambda^0\right)/T$ and $\bar{c}\left(\lambda^0\right)$ being the non-centrality parameter defined in CKP. Let $\tilde{\psi}_i$ be the estimator that minimizes the sum of squared residuals

$$S(\psi_i, \lambda_i^0) = \left(y_{i,t}^{\bar{\alpha}} - z_{i,t}^{\bar{\alpha}\prime}\left(\lambda_i^0\right)\psi_i\right)'\left(y_{i,t}^{\bar{\alpha}_i} - z_{i,t}^{\bar{\alpha}\prime}\left(\lambda_i^0\right)\psi_i\right).$$

Using these estimated parameters we can construct the GLS detrended variable $\tilde{y}_{i,t} = y_{i,t} - z_{i,t}'\left(\lambda_i^0\right)\tilde{\psi}_i$ and compute its first difference

$$\begin{aligned}
\Delta\tilde{y}_i &= \Delta y_i - \Delta z_i'\left(\lambda_i^0\right)\hat{\psi}_i \\
&= -\Delta z_i\left(\lambda_i^0\right)(\hat{\psi}_i - \psi_i) + \Delta F\delta_i + \Delta e_i \\
&= -\Delta z_i\left(\lambda_i^0\right)(z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right))^{-1} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right)\left(F^{\bar{\alpha}}\delta_i + e_i^{\bar{\alpha}}\right) + \Delta F\delta_i + \Delta e_i \\
&= f\delta_i + \xi_i, \qquad (1.2.5)
\end{aligned}$$

where $\xi_i = \Delta e_i - \Delta z_i\left(\lambda_i^0\right)(z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right))^{-1} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}}$ and $f = \Delta F - \Delta z_i\left(\lambda_i^0\right)$ $(z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right))^{-1} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) F^{\bar{\alpha}_i}$. In this case (1.2.5) is a factor model with the new observable variables $\Delta\tilde{y}_{i,t}$ and the estimation of the common factors and factor loadings can be done as in Bai and Ng (2004) using principal components. Let

$$\Delta\tilde{y} = (\Delta\tilde{y}_1, \Delta\tilde{y}_2, ..., \Delta\tilde{y}_N),$$

be the $(T-1)\times N$ data matrix. The estimated principal components of $\Delta F = (\Delta F_2, \Delta F_3, \ldots, \Delta F_T)$, denoted as $\Delta\tilde{F}$, are $\sqrt{T-1}$ times the r eigenvectors corresponding to the first r largest eigenvalues of the $(T-1)\times(T-1)$ matrix $\Delta\tilde{y}\Delta\tilde{y}'$, under the normalization $\Delta\tilde{F}\Delta\tilde{F}'/(T-1) = I_r$. The estimated loading matrix is $\tilde{\Lambda} = \Delta\tilde{y}'\Delta\tilde{F}/(T-1)$, from which we can define the estimated residuals as

$$\Delta\tilde{e}_{i,t} = \Delta\tilde{y}_{i,t} - \Delta\tilde{F}_t'\tilde{\delta}_i. \tag{1.2.6}$$

We can recover the idiosyncratic disturbance terms through cumulation, i.e.,

$$\tilde{e}_{i,t} = \sum_{s=2}^{t} \Delta\tilde{e}_{i,s},$$

whereas the common factors are estimated in the same fashion:

$$\tilde{F}_t = \sum_{s=2}^{t} \Delta\tilde{f}_s.$$

Using these two estimated components we can assess the source of potential I(1) non-stationarity of the observable variables.

1.3 Unit root test statistics

The order of integration of each component can be analyzed using the modified test statistics in Ng and Perron (2001) – hereafter, M-type test statistics. Thus, if we focus on the idiosyncratic component, the M-type unit root tests statistics are defined as:

$$MSB_i^{GLS} = \left(s_i^{-2}T^{-2}\sum_{t=1}^{T}\tilde{e}_{i,t-1}^2\right)^{1/2} \tag{1.3.1}$$

$$MZ_{i,\alpha}^{GLS} = \left(T^{-1}\tilde{e}_{i,T}^2 - s_i^2\right)\left(2T^{-2}\sum_{t=1}^{T}\tilde{e}_{i,t-1}^2\right)^{-1} \tag{1.3.2}$$

$$MZ_{i,t}^{GLS} = \left(T^{-1}\tilde{e}_{i,T}^2 - s_i^2\right)\left(4s_i^2T^{-2}\sum_{t=1}^{T}\tilde{e}_{i,t-1}^2\right)^{-1/2} \tag{1.3.3}$$

where $s_i^2 = (T-k_i)^{-1}\left(\sum_{t=k_i+1}^{T}\tilde{u}_{i,t,k_i}^2\right)/\left(1-\sum_{j=1}^{k_i}\tilde{b}_{i,j}\right)^2$ and k_i is the order of the autoregression selected using the information criteria proposed by Ng and Perron

(2001) and Perron and Qu (2007). The terms $\tilde{u}_{i,t,k_i}$ and $\tilde{b}_{i,j}$ are the OLS estimated coefficients from the regression

$$\Delta\tilde{e}_{i,t} = b_{i,0}\tilde{e}_{i,t-1} + \sum_{j=1}^{k_i} b_{i,j}\Delta\tilde{e}_{i,t-j} + u_{i,t,k_i}.$$

Using these statistics we can test the hypotheses that $e_{i,t}$ is I(1) against the alternative hypothesis that $e_{i,t}$ is I(0), i.e.,

$$\begin{cases} H_0: & \theta_i = 1 \\ H_1: & \theta_i < 1 \end{cases}$$

where θ_i appears in Equation (1.2.3).

As for the common factors, in the case where there is only one $(r = 1)$ common factor, we can proceed to test its order of integration using the M-type test statistics defined above, once it has been GLS detrended. The GLS detrending is performed using the vector of regressors $z_t\left(\lambda^0\right)$ that define the break dates $T_j^0 = E\left(T_{i,j}^0\right) = N^{-1}\sum_{i=1}^N T_{i,j}^0$, $j = 0, 1, \ldots, m$, which makes use of the mild heterogeneous property for the break dates. The GLS detrended estimated common factor is defined as $\tilde{F}_t^d = \tilde{F}_t - z_t'\left(\lambda^0\right)\tilde{\psi}$, where $\tilde{\psi}$ is the quasi-GLS estimation of the parameters. Then, the unit root tests given in (1.3.1)-(1.3.3) can be computed using $\tilde{F}_t^d$ instead of $\tilde{e}_{i,t}$ – the corresponding test statistics are denoted as MSB_F^{GLS}, $MZ_{F,\alpha}^{GLS}$ and $MZ_{F,t}^{GLS}$.

When there is more than one common factor $(r > 1)$, we can assess how many common factors are I(1) and I(0) using the MQ test statistics in Bai and Ng (2004). Start with $q = r$ and proceed in three stages:

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

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

iii. Define $MQ_c^d(q) = T\left[\tilde{v}_c^d(q) - 1\right]$ for the case of no change in the slope of the trend (Model 0) and $MQ_c^d\left(q, \lambda^0\right) = T\left[\tilde{v}_c^d\left(q, \lambda^0\right) - 1\right]$ for the case of changes in the slope of the trend (Models I and II).

(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

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

iii. Define the statistic $MQ_f^d(q) = T\left[\tilde{v}_f^d(q) - 1\right]$ for the case of no change in the slope of the trend (Model 0) and $MQ_f^d\left(q, \lambda^0\right) = T\left[\tilde{v}_f^d\left(q, \lambda^0\right) - 1\right]$ for the case of changes in the slope of the trend (Models I and II).

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 limiting distribution of these unit root test statistics is presented in the following theorem:

Theorem 1.1. *Let $y_{i,t}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, be a stochastic process with the DGP given by (1.2.1) to (1.2.3) and satisfying Assumptions A to E. Also, define $\bar{\alpha} = 1 + \bar{c}\left(\lambda^0\right)/T$, let s_i^2 be a consistent estimate of σ_i^2 and let k_i be chosen in a way that $k_i \to \infty$ and $k_i^3/\min[N, T] \to 0$. Then as $N, T \to \infty$ with $N/T \to 0$*

(a) the statistics applied to the idiosyncratic component converge to:

Model 0:

$$\begin{aligned} MSB_i^{GLS} &\Rightarrow \left(\int_0^1 V_{i,c,\bar{c}}(s)^2 ds\right)^{1/2} \\ MZ_{\alpha,i}^{GLS} &\Rightarrow 0.5\left(V_{i,c,\bar{c}}(1)^2-1\right)\left(\int_0^1 V_{i,c,\bar{c}}(s)^2 ds\right)^{-1} \\ MZ_{t,i}^{GLS} &\Rightarrow 0.5\left(V_{i,c,\bar{c}}(1)^2-1\right)\left(\int_0^1 V_{i,c,\bar{c}}(s)^2 ds\right)^{-1/2} \end{aligned}$$

Models I and II:

$$\begin{aligned} MSB_i^{GLS} &\Rightarrow \left(\int_0^1 V_{i,c,\bar{c}}\left(s,\lambda^0\right)^2 ds\right)^{1/2} \\ MZ_{\alpha,i}^{GLS} &\Rightarrow 0.5\left(V_{i,c,\bar{c}}\left(1,\lambda^0\right)^2-1\right)\left(\int_0^1 V_{i,c,\bar{c}}\left(s,\lambda^0\right)^2 ds\right)^{-1} \\ MZ_{t,i}^{GLS} &\Rightarrow 0.5\left(V_{i,c,\bar{c}}\left(1,\lambda^0\right)^2-1\right)\left(\int_0^1 V_{i,c,\bar{c}}\left(s,\lambda^0\right)^2 ds\right)^{-1/2}, \end{aligned}$$

where $\Rightarrow$ denotes weak convergence to the associated measure of probability, $V_{i,c,\bar{c}}(s)=W_{i,c}(s)-s\left(bW_{i,c}(1)+3(1-b)\int_0^1 uW_{i,c}(u)\,du\right)$, $b=(1-\bar{c})/\left(1-\bar{c}+\bar{c}^2/3\right)$, $V_{i,c,\bar{c}}\left(s,\lambda^0\right)=W_{i,c}(s)-z_2(s)A\left(\lambda^0\right)^{-1}\bar{V}_i\left(\lambda^0\right)$, $W_{i,c}(s)$ is an Ornstein-Uhlenbeck process and the terms $A\left(\lambda^0\right)$ and $\bar{V}_i\left(\lambda^0\right)$ for each cross-section are defined in the Appendix A.

(b) When $r=1$, the limiting distribution for the MSB_F^{GLS}, $MZ_{F,\alpha}^{GLS}$ and $MZ_{F,t}^{GLS}$ test statistics for the different model specifications is the same as the one given by MSB_i^{GLS}, $MZ_{i,\alpha}^{GLS}$ and $MZ_{i,t}^{GLS}$, respectively.

(c) When $r>1$, let $V_{q,c,\bar{c}}(s)$ and $V_{q,c,\bar{c}}\left(s,\lambda^0\right)$ be q-vectors with elements defined by $V_{j,c,\bar{c}}(s)=W_{j,c}(s)-s\left(bW_{j,c}(1)+3(1-b)\int_0^1 uW_{j,c}(u)\,du\right)$ and $V_{j,c,\bar{c}}\left(s,\lambda^0\right)=W_{j,c}(s)-z_2(s)A\left(\lambda^0\right)^{-1}\bar{V}_j\left(\lambda^0\right)$, $j=1,\ldots,q$, respectively.
For Model 0, let $v_*^d(q)$ be the smallest eigenvalues of

$$\Phi_*^d=\frac{1}{2}\left[V_{q,c,\bar{c}}(1)V_{q,c,\bar{c}}(1)'-I_p\right]\left[\int_0^1 V_{q,c,\bar{c}}(s)V_{q,c,\bar{c}}(s)'ds\right]^{-1}.$$

For Models I and II, let $v_*^d\left(q,\lambda^0\right)$ be the smallest eigenvalues of

$$\Phi_*^d(\lambda)=\frac{1}{2}\left[V_{q,c,\bar{c}}\left(1,\lambda^0\right)V_{q,c,\bar{c}}\left(1,\lambda^0\right)'-I_p\right]\left[\int_0^1 V_{q,c,\bar{c}}\left(s,\lambda^0\right)V_{q,c,\bar{c}}\left(s,\lambda^0\right)'ds\right]^{-1}.$$

(c.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, $MQ_c^d(q) \Rightarrow v_*^d(q)$ and $MQ_c^d\left(q,\lambda^0\right) \Rightarrow v_*^d\left(q,\lambda^0\right)$.

(c.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 $p \geq \bar{p}$, $MQ_f^d(q) \Rightarrow v_*^d(q)$ and $MQ_f^d\left(q,\lambda^0\right) \Rightarrow v_*^d\left(q,\lambda^0\right)$.

The proof of Theorem 1.1 is given in Appendix A. The asymptotic critical values for the MSB_i^{GLS}, $MZ_{i,\alpha}^{GLS}$ and $MZ_{i,t}^{GLS}$ – and the ones for the common factor, MSB_F^{GLS}, $MZ_{F,\alpha}^{GLS}$ and $MZ_{F,t}^{GLS}$ – can be found in Carrion-i-Silvestre et al. (2009). The asymptotic critical values for the MQ test statistics can be found in Table 1.1 for Model 0 and in Table 1.2 for Models I and II. For further developments, we have also computed by simulation the asymptotic mean and variance of the limiting distribution of the different test statistics reported in Theorem 1.1.

Finally, the model that we have specified assumes that the number of common factors is known. In practice we will need to estimate it using, for instance, the information criteria proposed in Bai and Ng (2002, 2004). We will analyze the performance of the use of these information criteria in the Monte Carlo simulation section.

1.4 Unknown structural breaks

So far, we have assumed that the vector of break points is known. This assumption might be feasible in some cases, where panels of variables such as, for instance, real exchange rates panels are analyzed and it is known that there is an important event that have affected the time series – in this case, the euro currency birth. However, there might be some cases where this assumption cannot be made and the date of the structural breaks needs to be estimated. Throughout this section, we assume that the number of structural breaks (m) is known and common to all individuals. In principle, it would be possible to specify a maximum number of structural breaks $(m_{\max})$ and estimate the number of structural breaks $(\tilde{m})$ using a panel Bayesian information criterion such as the one proposed in Bai and Ng (2002, 2004).

We proceed to estimate the date of the break points for each time series, without taking into account the other time series of the panel data set. The estimation of the unknown structural breaks for each time series relies on the procedure proposed in

Carrion-i-Silvestre et al. (2009) and consists of the following steps:

1. For a given value of m, compute an initial educated estimation of the date of the break points $\tilde{T}_i = \left(\tilde{T}_{i,1}, \ldots, \tilde{T}_{i,m}\right)'$ and the vector of parameters $\tilde{\psi}_i = \left(\tilde{\psi}'_{i,0}, \ldots, \tilde{\psi}'_{i,m}\right)'$ specifying the model

$$y_{i,t} = z'_{i,t}\left(\lambda_i^0\right)\psi_i + u_{i,t},$$

where the OLS estimation procedure is used to obtain the estimates. The estimates that are obtained in this step are the ones drawn from the minimization of the sum of squared residuals (SSR)

2. Obtain $\bar{c}(\tilde{\lambda}_i)$ using the estimates obtained in the previous step

3. Compute the quasi-difference of the variables and proceed to compute the GLS estimates of the parameters and break dates minimizing the restricted sum of squared residuals (RSSR) – see Carrion-i-Silvestre et al. (2009) for further details

4. Repeat steps 2 and 3 until convergence is achieved and store the estimated break dates

5. Compute the GLS detrended variable $\tilde{y}_{i,t}$ using the final estimates of the break points and the parameters of the model

6. Estimate the common factors, the factor loadings and the idiosyncratic disturbance terms using the method of principal components described in Section 1.2

7. Test the order of integration of the different components using the unit root test statistics proposed in Section 1.3

1.5 Panel data unit root test statistics

Although the analysis that has been conducted so far allows us to test the order of integration of the different stochastic processes involved in the model at a unit-by-unit level, it is possible, in principle, to improve the performance of the statistical inference combining the individual test statistics. In order to pool the individual test statistics we require to introduce the additional assumption of cross-section independence of the idiosyncratic disturbance terms $e_{i,t}$, which implies that the source of cross-section dependence is well captured by the common factor structure. This assumption makes the individual test statistics to be cross-sectionally independent.

We are interested in testing the null hypothesis that all units are I(1) non-stationary against the alternative hypothesis that at least one unit is I(0) stationary, i.e.,

$$\begin{cases} H_0: & \theta_i = 1 \ \forall i \\ H_1: & \theta_i < 1 \text{ for some } i \end{cases} \tag{1.5.1}$$

where θ_i appears in Equation (1.2.3).

We define M in a generic way to define one of the M-type unit root test statistic that we have considered, i.e., for a given unit i we can compute any of the statistics $M_i = \{MSB_i^{GLS}, MZ_{i,\alpha}^{GLS}, MZ_{i,t}^{GLS}\}$. The first way to define a pool panel data test statistic bases on the standardized mean of the individual statistics

$$Z^M = \frac{\sqrt{N}(\overline{M_i} - \bar{\zeta}^M))}{\bar{v}^M} \to N(0,1),$$

where $\overline{M} = N^{-1}\sum_{i=1}^N M_i$, $\bar{\zeta}^M = N^{-1}\sum_{i=1}^N \zeta_i^M$ and $\bar{v}^{M^2} = N^{-1}\sum_{i=1}^N v_i^{M^2}$, where ζ_i^M and $v_i^{M^2}$ are the mean and the variance of the M_i statistic, $M = \{MSB^{GLS}, MZ_\alpha^{GLS}, MZ_t^{GLS}\}$.

The next three panel data tests are based on the combination of the individual p-values. Bai and Ng (2004) noted that pooling based on the p-values not only can be used on unbalanced panels but it has the advantage of allowing heterogeneity across units. Maddala and Wu (1999) define the panel data Fisher-type statistic that can be applied to panel data sets with a small number of cross-section units:

$$P^M = -2\sum_{i=1}^N \ln \tilde{\varphi}_i^M \sim \chi^2_{2N}, \tag{1.5.2}$$

where φ_i^M denotes the p-value of the M_i statistic, $M = \{MSB^{GLS}, MZ_\alpha^{GLS}, MZ_t^{GLS}\}$.

Choi (2001) goes one step further and proposes the following test that is valid for $N \to \infty$:

$$P_m^M = -\frac{1}{\sqrt{N}}\sum_{i=1}^N \left(\ln \tilde{\varphi}_i^M + 1\right) \to N(0,1),$$

and also Choi (2001) proposes the following pool test statistic:

$$C^M = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}\left(\tilde{\varphi}_i^M\right) \to N(0,1),$$

where $\Phi(\cdot)$ denotes the standard Normal cumulative distribution function. In both Choi (2001) statistics, $M = \{MSB^{GLS}, MZ_\alpha^{GLS}, MZ_t^{GLS}\}$.

We simulate the asymptotic mean and variance for each λ^0 and these values are presented in Table 1.4 for 1 structural break and in Table 1.5 for 2 structural breaks. The simulated p-values for the M-class tests are available upon request.

1.6 Monte Carlo simulations

In this section, we conduct a set of simulation experiments in order to investigate the small-sample properties of the proposed tests. We first cover the cases with a single and then multiple structural breaks when the location of the potential structural break is known. Later, we relax the assumption of known structural breaks and we show the simulation results for both a single and multiple endogenous structural breaks. The nominal size of the statistics is set at the 5% level of significance. We present the results of the most general model specification (Model II) with structural breaks in both the slope and the trend. In addition, we focus on the first estimation procedure that has been described above, which bases the results on the individual estimation of the break points. All simulations are performed in GAUSS and the Monte Carlo results reported below are obtained using $1,000$ replications.

1.6.1 Known structural breaks

We begin the analysis by considering the performance of the panel data tests for the case of known single structural break. We account for cross-section dependence by using the common factor structure. The data generating process consists of the following system of equations:

$$y_{i,t} = d_{i,t} + F_t'\delta_i + e_{i,t} \tag{1.6.1}$$

$$d_{i,t} = \alpha_{i,1} DU_{i,t}\left(T_1^0\right) + \beta_{i,1} DT_{i,t}\left(T_1^0\right) \tag{1.6.2}$$

$$F_t = \rho F_{t-1} + w_t \tag{1.6.3}$$

$$e_{i,t} = \theta_i e_{i,t-1} + \varepsilon_{i,t} \tag{1.6.4}$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$, $\varepsilon_{i,t} \sim iid\ N(0,1)$ and $w_t \sim iid\ N\left(0, \sigma_F^2\right)$. Without loss of generality, the values of $\alpha_{i,1}$ and $\beta_{i,1}$ are set equal to 1 for all individuals. For the simulation of the common factor component we specify $\delta_i \sim N(1,1)$, $\rho = \{0.9, 0.95, 1\}$ and $\sigma_F^2 = \{0.5, 1, 10\}$. The number of common factors is estimated using the panel Bayesian information criterion (BIC) in Bai and Ng (2002). The empirical size of the tests for the idiosyncratic component is simulated setting the autoregressive parameter $\theta_i = 1$ while the empirical power is simulated using $\theta_i = 0.9$. The data is generated with $N = 20$ cross-sectional units and three sample sizes $T = \{50, 100, 200\}$.[2] Due to space constraints we only discuss results for $\lambda^0 = 0.5$ with the note that we obtain similar results for $\lambda^0 = \{0.3, 0.7\}$. The full set of results is available at the end of the chapter.

We first consider the size of the statistics and we summarize it in Table 1.6. We can see that all tests perform well since the empirical size is really close to the nominal size of 5%. It is interesting to note that the results are similar for all panel statistics regardless of the order of integration of the common factors. Also, the empirical size does not appear to be affected by the changes in T or σ_F^2. Based on these results we can infer that the performance of the panel tests for the case of one known structural break is good with almost no size distortions.

We then continue the analysis of the proposed panel tests by investigating their power properties. The results are presented in Table 1.7 and they are in line with what it is expected from the asymptotic theory. It is easily seen that the power increases with the increase in T. Note that even for $T = 100$ the power is almost 1 while the tests have perfect power for $T = 200$. Therefore, the results suggest that the all panel statistics have good power properties especially for $T > 50$.

Panel A of Table 1.8 shows the size and power properties of statistics for the common factor, MSB_F^{GLS}, $MZ_{F,\alpha}^{GLS}$ and $MZ_{F,t}^{GLS}$ for one known break. The values for $\rho = 1$ indicate the size of the tests while those for $\rho = \{0.9, 0.95\}$ indicate the power properties. For $T = 50$, or what is considered a small sample, the size and power of the statistics are close to 0. That is, in small samples, the statistics MSB_F^{GLS}, $MZ_{F,\alpha}^{GLS}$ and $MZ_{F,t}^{GLS}$ cannot distinguish between a stationary process and one with unit roots. This might be due to the use of asymptotic critical values and we expect that finite sample critical values will improve the empirical size and power of these statistics. However, as T increases, the empirical size is very close to the 5% nominal size. Also, power increases when T increases.

[2] The simulations have been also carried out for $N = \{40, 60\}$. These results are also reported in this chapter, but we are going to focus our discussion on the results for $N = 20$.

Next, we extend the analysis of panel data unit root tests for the case of two known structural breaks. The DGP has the following form:

$$\begin{aligned}
y_{i,t} &= d_{i,t} + F_t'\delta_i + e_{i,t} && (1.6.5)\\
d_{i,t} &= \alpha_{i,1} DU_{i,t}\left(T_1^0\right) + \alpha_{i,2} DU_{i,t}\left(T_2^0\right) + \beta_{i,1} DT_{i,t}\left(T_1^0\right) \\
&\quad + \beta_{i,2} DT_{i,t}\left(T_2^0\right) && (1.6.6)\\
F_t &= \rho F_{t-1} + w_t && (1.6.7)\\
e_{i,t} &= \theta_i e_{i,t-1} + \varepsilon_{i,t} && (1.6.8)
\end{aligned}$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$, $\varepsilon_{i,t} \sim iid\, N(0,1)$ and $w_t \sim iid\, N\left(0, \sigma_F^2\right)$. As in the previous case, $\delta_i \sim N(1,1)$, $\rho = \{0.9, 0.95, 1\}$ and $\sigma_F^2 = \{0.5, 1, 10\}$. We set $\theta_i = 1$ for the empirical size simulations and $\theta_i = 0.9$ for the empirical power simulations. Without loss of generality, the values of $\alpha_{i,1}$, $\alpha_{i,2}$, $\beta_{i,1}$ and $\beta_{i,2}$ are set equal to 1 for all individuals. The sample size is $T = \{50, 100, 200\}$ and $N = 20$. Since we analyze the case with two structural breaks, we set $\lambda_1^0 = 0.3$ and $\lambda_2^0 = 0.7$.

First, we present the results on the empirical size of the panel unit root tests, which are reported in Table 1.9. We observe that the empirical size of the test statistics is close to the nominal level of 5%. The P_m^M panel statistic based on MSB^{GLS} and MZ_α^{GLS} tests tends to over-reject for $T = 100$ but its empirical size approaches the nominal size as T increases. Overall, for the case of two known structural breaks, the statistics still perform well with very small size distortions. Second, we analyze the power properties of the proposed panel statistics. Table 1.10 presents these results. As expected, in small samples, the power of the tests for the case of two known breaks is slightly lower than the power of the tests for one break. However, for $T = 200$ the power of all panel statistics is 1. Like in the previous case, the power of the tests increases as T increases. Therefore, for the case of one or two known breaks, the panel tests have good power and really small size distortions.

Panel B of Table 1.8 shows the size and power properties of statistics for the common factor, MSB_F^{GLS}, $MZ_{F,\alpha}^{GLS}$ and $MZ_{F,t}^{GLS}$ for two known breaks. The results are similar to those for one known break (Panel A). That is, except for $T = 50$, the empirical size is very close to the nominal size and the empirical power increases when increasing T.

1.6.2 Unknown structural breaks

In the previous subsection, we assumed that the timing of the structural breaks is known. This is not always the case. Actually, many times, the researcher does not know *a priori* the location of the structural breaks and he/she needs the tools to do the proper analysis. Therefore, taking that into consideration, we simulate first the case when there is only one structural break and later on, we extend the analysis to the case with two structural breaks. The main difference between the known and unknown structural break cases consists in estimating the structural breaks. In the case of the unknown structural breaks, we estimate the location of the structural breaks through the global minimization of the RSSR of the GLS detrended model presented in Section 1.4. More explicitly, we estimate the location of unknown structural break by implementing the steps 1 to 7.

The DGP used in this section is the same as for the known break case. The results on the empirical size, contained in Table 1.11, are summarized as follows. Overall, the performance of the statistics for the case of one unknown structural break is similar to that for the case of known structural break with a few exceptions. For $T = \{50, 100\}$ and $\sigma_F^2 = 10$, we can see that the empirical size is slightly above the 5% nominal size. This indicates a tendency of all the panel statistics to over-reject the null hypothesis of unit root. Also, for $T = 200$, the panel statistics P^M and P_m^M show small size distortion. More exactly, the values of these sizes indicate an under-rejection of the null hypothesis. However, the size of the panel statistics Z^M and C^M is really close to the nominal size. As expected, the empirical size approaches the nominal one as T increases. This indicates a good performance of these panel statistics. Next, we continue with the results showing the power properties of the panel tests and we present those results in Table 1.12. Like in the previous cases, the power accuracy of the statistics tends to be low when time-dimension is small. However, the power of the tests seems reasonable in larger samples and increases as sample size increases. Overall, the simulations for the case of one unknown structural break lead us to conclude that the panel statistics have good power in larger samples and perform rather well with small size distortions.

We then look at the size and power properties of statistics for the common factor for one unknown break, which are presented in Panel A of Table 1.13. The size and power of the unit root tests for common factors for $T = 50$ are still close to 0, indicating the statistics are under-sized and have very low power. However, for $T = 100$, the empirical size is very close to the 5% nominal size and the empirical power is even higher than the one in the known break case. For the rest of the time-dimensions, the power of the tests is increasing but it is slightly lower than for the case of one known break. Also, the empirical size of the statistics is close to the nominal size.

In the next step, we consider the size and power of proposed statistics for the case of two unknown structural breaks. The data generating process is the same as for the case with two known structural breaks and consists of Equations (1.6.5) through (1.6.8). Table 1.14 presents the empirical size of the panel statistics. For the majority of cases, the empirical size of the simulated statistics is close to the nominal size. There are a few exceptions when the size is slightly bigger or smaller than the 5% nominal size. For example, we can see a similar trend for all statistics when $T = 50$ and $\sigma_F^2 = 10$. More exactly, these values point out to an over-rejection of the null hypothesis of unit root. However, unlike the results for the case of one unknown structural break for $T = 100$ and $\sigma_F^2 = 10$, the tests have little size distortions for the case of two unknown structural breaks. The Z^M panel statistic based on MSB^{GLS} test performs rather well for $T = 100$ but over-rejects for $T = 200$ and under-rejects for $T = 50$ when $\sigma_F^2 = \{0.5, 1\}$. The P^M and P_m^M statistics show moderate size distortions for $T = 200$. Their values of the size below the nominal size point to an under-rejection of the null hypothesis of unit root. The results for the size of the C^M test indicate that the test has a good performance overall with the exception when $T = 50$ and $\sigma_F^2 = 10$ mentioned earlier. Even though for a few cases the statistics suffer from moderate size distortions, for the majority of cases the tests maintain the nominal level well. Finally, we present in Table 1.15 the results of the power of the tests. As expected, the power accuracy is lower when the time series dimension is small but increases as T increases. For the case of two unknown breaks, the statistics have slightly lower power than for the case of one unknown break. Overall, simulations lead us to conclude that for the case of unknown breaks, the power of the proposed panel tests is lower than in case of known breaks and the statistics tend to under-reject when T is large.

Panel B of Table 1.13 shows the empirical size and power of the unit root statistics for the common factor. For $T = 50$, the results are similar to the ones for one unknown break case. However, as T increases, both empirical size and power of the tests are increasing. The size of the statistics indicate that for the case of two unknown breaks, the unit root statistics for the common factors tend to over-reject when T is large.

1.7 Empirical application

In this section, we present an empirical application of the panel tests described in the previous sections. We use the Maddison dataset used by Dawson and Strazicich (2010) and Kejriwal and Lopez (2010). It consists of annual time series of per capita real GDP over the period 1870-2008 for 19 OECD countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and United States. Note that the data are expressed in 1990 Geary-Khamis dollars. The logarithm of per capita GDP is the output throughout the rest of this chapter. For more details about the data, see Maddison (2009).

We begin the analysis by testing for cross-sectional dependence among the series of the panel. To this end, we apply the CD statistic of Pesaran (2004, 2013) to the panel of OECD countries. One advantage of this test is its robustness to single or multiple structural breaks, making it desirable in the empirical work. The calculated value of the CD test is 20.099, which indicate that we can reject the null hypothesis of no cross-sectional dependence at any acceptable level of significance. As for the common factors component, we estimate it using the method of principal components. In this case, with the maximum number of common factors set at one, the panel BIC information criterion in Bai and Ng (2002) selects one common factor and the estimated common factor is characterized as I(1) stochastic process.

Next, we apply the M-class unit root tests with two structural breaks in trend to each country individually and we show the results in Table 1.16. The second and third columns represent the break dates for each country while the last three columns present the individual M-class statistics. After looking at the break dates, we can see that the World War II period is the most frequent time of structural breaks. This is consistent with the previous studies on the OECD countries like Dawson and Strazicich (2010). We simulate the asymptotic critical values for the MSB^{GLS}, MZ_{α}^{GLS} and MZ_{t}^{GLS} tests and for each pair of λ. Although we do not present the critical values in this chapter, they are available upon request. The MSB^{GLS} statistic suggests the rejection of the null hypothesis at the 5% significance level for Denmark, New Zealand and US and at the 10% level for Germany and UK. The other two statistics, MZ_{α}^{GLS} and MZ_{t}^{GLS} show similar results. Both tests suggest a rejection at the 5% level for Denmark, New Zealand and US and at the 10% level for UK. Overall, for the majority of countries we cannot reject the unit root hypothesis in favor of I(0) stationarity.

Finally, we apply the panel unit root statistics and we show the results in Table 1.17. All panel unit root statistics are able to reject the non-stationary null hypothesis at 10% and two-thirds of them are able to reject the null at the 5% level of significance. Therefore, we can infer that the panel as a whole is I(0) stationary when applied the unit root tests to idiosyncratic component. However, as Table 1.18 shows, the null hypothesis of the unit root tests applied to common factors cannot be rejected. That is, the non-stationarity of the real per capita GDP is driven by the I(1) non-stationary common factor.

1.8 Conclusion

In this chapter, we propose several panel data unit root tests that allow for multiple structural breaks and common factors to control for the presence of cross-section dependence. The test statistics are based on the use of GLS detrending procedure and the structural breaks are allowed under both the null and the alternative hypotheses. The model specification considers both known and unknowns breaks. This chapter derives the limiting distribution of the individual unit root test statistics for the idiosyncratic disturbance term and the common factors. Further, we also show that panel data unit root test statistics can be defined through the combination of the individual test statistics of the idiosyncratic component.

The performance of the statistics that have been proposed is evaluated using a Monte Carlo simulation experiment. The simulations show that the test statistics perform well for the cases of known structural breaks. When the location of the structural breaks is not known *a priori* the panel statistics suffer from under-rejection when the time series dimension is large. Finally, we apply the proposed tests to a panel data set of annual real per capita GDP over the period 1870-2008 for 19 OECD developed countries. All panel statistics rejected the null hypothesis of panel data unit root in favor of I(0) stationarity for the idiosyncratic component of the real per capita GDP. However, all unit root tests for the common factors cannot reject the null hypothesis of unit root. Therefore, we conclude that there are global stochastic trends affecting the real per capita GDP.

Table 1.1: Asymptotic critical values for the $MQ^d(q)$ tests for Model 0

r	1%	5%	10%
1	-13.78	-8.19	-5.82
2	-25.11	-18.16	-14.96
3	-35.27	-27.22	-23.53
4	-45.22	-36.31	-32.16
5	-54.26	-44.87	-40.52
6	-63.65	-53.63	-48.92

The moments of the limiting distribution of the statistics by means of Monte Carlo simulation, using 1,000 steps to approximate the Brownian motion functionals and 100,000 replications.

Table 1.2: Asymptotic critical values for the $MQ^d(q,\lambda^0)$ tests for Models I and II, one structural break case

	$\lambda^0 = 0.1$			$\lambda^0 = 0.2$			$\lambda^0 = 0.3$		
r	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-30.12	-22.28	-18.69	-31.11	-23.44	-19.81	-31.92	-23.83	-20.29
2	-41.08	-32.43	-28.30	-41.76	-33.26	-29.34	-42.09	-33.69	-29.67
3	-50.84	-41.48	-37.12	-51.51	-42.24	-37.93	-51.79	-42.75	-38.35
4	-59.87	-50.08	-45.56	-60.37	-50.99	-46.37	-60.98	-51.24	-46.63
5	-68.92	-58.65	-53.62	-69.45	-59.41	-54.49	-69.82	-59.66	-54.82
6	-77.35	-67.03	-61.90	-78.28	-67.54	-62.61	-78.30	-67.80	-62.77
	$\lambda^0 = 0.4$			$\lambda^0 = 0.5$			$\lambda^0 = 0.6$		
r	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-31.60	-23.97	-20.32	-31.85	-23.86	-20.25	-31.11	-23.23	-19.63
2	-42.07	-33.70	-29.72	-42.20	-33.66	-29.59	-41.90	-33.26	-29.25
3	-51.58	-42.57	-38.24	-51.52	-42.54	-38.16	-51.24	-42.14	-37.85
4	-60.54	-51.17	-46.63	-60.91	-51.11	-46.41	-60.32	-50.69	-46.17
5	-69.72	-59.49	-54.70	-69.09	-59.36	-54.53	-68.90	-59.11	-54.40
6	-78.25	-67.67	-62.68	-77.92	-67.61	-62.62	-77.65	-67.29	-62.33
	$\lambda^0 = 0.7$			$\lambda^0 = 0.8$			$\lambda^0 = 0.9$		
r	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-30.11	-22.29	-18.75	-28.90	-21.14	-17.67	-26.80	-19.38	-16.01
2	-41.07	-32.58	-28.62	-40.02	-31.49	-27.59	-38.01	-29.59	-25.70
3	-50.51	-41.50	-37.15	-50.00	-40.70	-36.50	-48.28	-38.96	-34.64
4	-60.02	-50.40	-45.84	-59.04	-49.52	-44.97	-57.40	-47.80	-43.28
5	-68.98	-58.98	-54.00	-67.99	-58.17	-53.25	-66.78	-56.42	-51.63
6	-77.38	-67.13	-61.96	-76.77	-66.38	-61.45	-75.31	-65.13	-60.10

The moments of the limiting distribution of the statistics by means of Monte Carlo simulation, using 1,000 steps to approximate the Brownian motion functionals and 100,000 replications.

Table 1.3: Mean and variance for the M-class statistics for the cases simulated in this study

	1 known break					
	MSB^{GLS}		MZ_{α}^{GLS}		MZ_{t}^{GLS}	
T	Mean	Variance	Mean	Variance	Mean	Variance
50	0.261456	0.005521	-8.282115	16.489818	-1.9013545	0.294152
100	0.239627	0.004043	-9.935691	24.929351	-2.096539	0.328674
200	0.230758	0.003658	-10.823839	30.655044	-2.205788	0.378106
	2 known breaks					
	MSB^{GLS}		MZ_{α}^{GLS}		MZ_{t}^{GLS}	
T	Mean	Variance	Mean	Variance	Mean	Variance
50	0.216900	0.002273	-11.538986	18.016178	-2.312606	0.220923
100	0.199408	0.001760	-13.776592	28.413185	-2.557161	0.270469
200	0.190975	0.001673	-15.248692	38.824724	-2.684333	0.315443

Simulations are based on 1000 replications. These values are valid for the Monte Carlo simulation presented in Section 1.6. The DGP for the one break case is given by Equations (1.6.1) to (1.6.4). For the case on 1 known break, the values for the mean and the variance are calculated using $\lambda^0 = 0.5$. The DGP for the two breaks case is given by Equations (1.6.5) to (1.6.8). The values for the mean and the variance for the case of 2 known breaks are calculated using $\lambda_1^0 = 0.3$ and $\lambda_2^0 = 0.7$.

Table 1.4: Asymptotic mean and variance for the M-class statistics for 1 known break

	MSB^{GLS}																	
	$\lambda^0=0.1$		$\lambda^0=0.2$		$\lambda^0=0.3$		$\lambda^0=0.4$		$\lambda^0=0.5$		$\lambda^0=0.6$		$\lambda^0=0.7$		$\lambda^0=0.8$		$\lambda^0=0.9$	
T	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
50	0.255	0.004	0.244	0.003	0.244	0.003	0.249	0.004	0.262	0.006	0.278	0.008	0.298	0.011	0.320	0.014	0.349	0.017
100	0.246	0.004	0.234	0.003	0.229	0.003	0.232	0.003	0.240	0.004	0.254	0.006	0.270	0.008	0.289	0.010	0.312	0.012
200	0.241	0.004	0.230	0.004	0.224	0.003	0.224	0.003	0.230	0.004	0.239	0.005	0.255	0.006	0.274	0.009	0.295	0.010
	MZ_α^{GLS}																	
	$\lambda^0=0.1$		$\lambda^0=0.2$		$\lambda^0=0.3$		$\lambda^0=0.4$		$\lambda^0=0.5$		$\lambda^0=0.6$		$\lambda^0=0.7$		$\lambda^0=0.8$		$\lambda^0=0.9$	
T	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
50	-8.250	14.629	-8.702	14.064	-8.785	14.184	-8.617	15.507	-8.320	16.781	-7.854	17.549	-7.180	17.511	-6.469	16.467	-5.725	13.809
100	-9.291	23.069	-10.098	24.238	-10.278	23.797	-10.212	23.795	-9.848	24.086	-9.415	25.031	-8.728	24.994	-7.889	23.194	-6.965	19.601
200	-9.851	29.598	-10.709	31.300	-11.145	31.786	-11.203	31.043	-10.926	31.462	-10.364	30.749	-9.640	29.710	-8.907	29.388	-7.768	23.596
	MZ_t^{GLS}																	
	$\lambda^0=0.1$		$\lambda^0=0.2$		$\lambda^0=0.3$		$\lambda^0=0.4$		$\lambda^0=0.5$		$\lambda^0=0.6$		$\lambda^0=0.7$		$\lambda^0=0.8$		$\lambda^0=0.9$	
T	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
50	-1.887	0.264	-1.962	0.250	-1.971	0.245	-1.943	0.264	-1.897	0.304	-1.842	0.318	-1.757	0.342	-1.670	0.338	-1.576	0.304
100	-2.005	0.360	-2.102	0.343	-2.120	0.330	-2.125	0.324	-2.095	0.331	-2.039	0.356	-1.954	0.368	-1.866	0.364	-1.774	0.337
200	-2.069	0.414	-2.168	0.411	-2.213	0.373	-2.225	0.374	-2.195	0.363	-2.161	0.374	-2.082	0.392	-1.985	0.381	-1.865	0.354

The mean and the variance for each break are simulated using on 10,000 replications.

Table 1.5: Asymptotic mean and variance for the M-class statistics for 2 known breaks

	MSB^{GLS}								
	mean (top triangle), variance (bottom triangle)								
λ^0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1		0.217	0.203	0.194	0.190	0.191	0.197	0.207	0.221
0.2	0.004		0.201	0.191	0.186	0.185	0.189	0.198	0.211
0.3	0.003	0.003		0.194	0.187	0.185	0.187	0.195	0.207
0.4	0.002	0.002	0.002		0.193	0.190	0.192	0.196	0.207
0.5	0.002	0.002	0.002	0.002		0.199	0.200	0.205	0.213
0.6	0.002	0.002	0.002	0.002	0.003		0.215	0.217	0.224
0.7	0.002	0.002	0.002	0.002	0.003	0.004		0.234	0.238
0.8	0.003	0.002	0.002	0.002	0.003	0.004	0.006		0.256
0.9	0.004	0.003	0.003	0.003	0.003	0.004	0.006	0.007	
	MZ_α^{GLS}								
	mean (top triangle), variance (bottom triangle)								
λ^0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1		-12.471	-13.829	-14.893	-15.486	-15.569	-14.995	-13.946	-12.628
0.2	46.632		-13.977	-15.438	-16.225	-16.183	-15.947	-14.935	-13.487
0.3	50.801	51.215		-14.721	-16.013	-16.421	-16.147	-15.357	-13.934
0.4	52.241	53.388	47.737		-14.943	-15.665	-15.729	-15.069	-13.706
0.5	50.722	53.282	52.060	48.311		-14.450	-14.716	-14.387	-13.415
0.6	51.965	51.185	51.734	51.655	48.970		-13.378	-13.185	-12.533
0.7	51.061	52.727	50.559	52.826	52.517	48.374		-11.909	-11.366
0.8	49.087	49.630	49.700	50.724	50.989	49.417	46.587		-10.144
0.9	43.668	46.501	45.750	44.800	48.285	46.458	43.866	40.023	
	MZ_t^{GLS}								
	mean (top triangle), variance (bottom triangle)								
λ^0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1		-2.337	-2.484	-2.588	-2.650	-2.667	-2.637	-2.555	-2.414
0.2	0.511		-2.491	-2.635	-2.722	-2.758	-2.722	-2.656	-2.512
0.3	0.508	0.480		-2.584	-2.709	-2.750	-2.734	-2.681	-2.559
0.4	0.468	0.470	0.437		-2.623	-2.693	-2.707	-2.668	-2.557
0.5	0.445	0.422	0.404	0.414		-2.578	-2.615	-2.576	-2.495
0.6	0.417	0.406	0.396	0.399	0.406		-2.467	-2.486	-2.407
0.7	0.415	0.399	0.376	0.402	0.417	0.430		-2.332	-2.291
0.8	0.408	0.399	0.399	0.397	0.424	0.459	0.454		-2.149
0.9	0.398	0.390	0.375	0.385	0.408	0.436	0.448	0.445	

The mean and the variance for each break are simulated using on $10,000$ replications and $T = 1000$. These values are valid for the Monte Carlo simulation presented in Section 1.6. The DGP for the one break case is given by Equations (1.6.1) to (1.6.4).

Table 1.6: Empirical size of panel unit root statistics for 1 known break for $N = 20$ and $\lambda^0 = 0.5$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.044	0.054	0.048	0.045	0.047	0.044	0.053	0.055	0.058	0.050	0.053	0.053
		0.95	0.048	0.061	0.065	0.055	0.064	0.059	0.069	0.070	0.066	0.057	0.060	0.061
		1	0.052	0.064	0.059	0.051	0.058	0.060	0.064	0.077	0.076	0.062	0.062	0.062
	1	0.9	0.032	0.038	0.034	0.045	0.045	0.044	0.054	0.060	0.053	0.032	0.036	0.038
		0.95	0.052	0.069	0.065	0.055	0.056	0.052	0.071	0.069	0.065	0.066	0.072	0.065
		1	0.046	0.054	0.050	0.037	0.045	0.048	0.051	0.057	0.060	0.045	0.048	0.049
	10	0.9	0.049	0.061	0.053	0.057	0.051	0.046	0.066	0.062	0.057	0.059	0.060	0.055
		0.95	0.043	0.078	0.064	0.060	0.067	0.061	0.074	0.079	0.076	0.063	0.069	0.067
		1	0.041	0.060	0.055	0.065	0.066	0.059	0.075	0.074	0.073	0.054	0.054	0.057
100	0.5	0.9	0.040	0.050	0.050	0.048	0.041	0.048	0.053	0.050	0.065	0.046	0.042	0.049
		0.95	0.047	0.050	0.058	0.045	0.042	0.050	0.059	0.048	0.060	0.052	0.049	0.056
		1	0.051	0.048	0.056	0.045	0.042	0.044	0.056	0.049	0.055	0.055	0.053	0.055
	1	0.9	0.031	0.049	0.052	0.055	0.039	0.054	0.064	0.059	0.066	0.048	0.048	0.050
		0.95	0.043	0.049	0.057	0.050	0.044	0.054	0.055	0.056	0.063	0.051	0.048	0.056
		1	0.038	0.053	0.043	0.054	0.052	0.065	0.067	0.065	0.073	0.054	0.043	0.041
	10	0.9	0.046	0.051	0.055	0.056	0.050	0.057	0.063	0.059	0.067	0.052	0.052	0.057
		0.95	0.051	0.055	0.059	0.056	0.041	0.051	0.061	0.055	0.062	0.056	0.049	0.057
		1	0.045	0.049	0.050	0.039	0.039	0.053	0.055	0.049	0.064	0.049	0.047	0.049
200	0.5	0.9	0.046	0.058	0.053	0.040	0.044	0.038	0.051	0.054	0.043	0.054	0.059	0.051
		0.95	0.066	0.080	0.069	0.063	0.064	0.055	0.070	0.078	0.061	0.075	0.076	0.069
		1	0.043	0.054	0.039	0.045	0.051	0.039	0.060	0.066	0.049	0.051	0.051	0.039
	1	0.9	0.049	0.064	0.050	0.051	0.054	0.042	0.065	0.063	0.048	0.056	0.056	0.043
		0.95	0.059	0.067	0.054	0.046	0.053	0.042	0.063	0.065	0.051	0.063	0.062	0.053
		1	0.055	0.074	0.058	0.067	0.069	0.056	0.080	0.087	0.070	0.065	0.066	0.057
	10	0.9	0.056	0.064	0.051	0.055	0.056	0.047	0.067	0.068	0.056	0.064	0.057	0.050
		0.95	0.049	0.063	0.048	0.050	0.053	0.045	0.061	0.063	0.054	0.056	0.055	0.048
		1	0.055	0.064	0.051	0.052	0.053	0.043	0.061	0.069	0.054	0.056	0.058	0.050

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.7: Empirical power of panel unit root statistics for 1 known break for $N = 20$ and $\lambda^0 = 0.5$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.305	0.354	0.369	0.246	0.262	0.262	0.277	0.306	0.303	0.326	0.358	0.363
		0.95	0.304	0.352	0.357	0.254	0.273	0.277	0.283	0.308	0.302	0.314	0.348	0.351
		1	0.300	0.368	0.392	0.247	0.277	0.274	0.286	0.306	0.304	0.326	0.375	0.381
	1	0.9	0.304	0.341	0.362	0.236	0.259	0.252	0.256	0.283	0.284	0.314	0.350	0.357
		0.95	0.311	0.364	0.384	0.269	0.277	0.275	0.296	0.303	0.303	0.333	0.365	0.375
		1	0.311	0.367	0.384	0.252	0.277	0.274	0.280	0.303	0.302	0.332	0.370	0.378
	10	0.9	0.299	0.366	0.372	0.263	0.283	0.270	0.290	0.311	0.307	0.333	0.361	0.363
		0.95	0.332	0.381	0.389	0.261	0.291	0.291	0.289	0.319	0.317	0.353	0.383	0.391
		1	0.335	0.398	0.410	0.267	0.304	0.303	0.316	0.347	0.335	0.359	0.402	0.406
100	0.5	0.9	0.997	0.994	0.997	0.963	0.972	0.966	0.973	0.977	0.972	0.997	0.998	0.997
		0.95	0.996	0.995	0.998	0.970	0.980	0.976	0.978	0.983	0.981	0.997	0.999	0.998
		1	0.993	0.990	0.996	0.963	0.972	0.965	0.971	0.979	0.973	0.994	0.995	0.994
	1	0.9	0.994	0.986	0.994	0.956	0.964	0.958	0.962	0.972	0.964	0.993	0.996	0.994
		0.95	0.999	0.993	0.996	0.976	0.983	0.978	0.983	0.985	0.982	0.997	0.998	0.995
		1	0.991	0.988	0.993	0.950	0.965	0.959	0.966	0.975	0.966	0.991	0.994	0.993
	10	0.9	0.991	0.989	0.995	0.967	0.977	0.971	0.974	0.979	0.975	0.990	0.997	0.995
		0.95	0.996	0.988	0.997	0.957	0.965	0.964	0.963	0.975	0.972	0.994	0.998	0.997
		1	0.992	0.989	0.994	0.971	0.975	0.970	0.977	0.981	0.975	0.993	0.995	0.994
200	0.5	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.8: Empirical size and power of the M-type unit root statistics for known breaks for $N = 20$ for 1 common factor

			Panel A: 1 known break			Panel B: 2 known breaks		
T	σ_F^2	ρ	MSB_F^{GLS}	$MZ_{F,\alpha}^{GLS}$	$MZ_{F,t}^{GLS}$	MSB_F^{GLS}	$MZ_{F,\alpha}^{GLS}$	$MZ_{F,t}^{GLS}$
50	0.5	0.9	0.000	0.000	0.000	0.000	0.000	0.000
		0.95	0.000	0.000	0.000	0.000	0.000	0.000
		1	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.9	0.000	0.000	0.000	0.000	0.000	0.000
		0.95	0.000	0.000	0.000	0.000	0.000	0.000
		1	0.000	0.000	0.000	0.000	0.000	0.000
	10	0.9	0.000	0.000	0.000	0.000	0.000	0.000
		0.95	0.000	0.000	0.000	0.000	0.000	0.000
		1	0.000	0.000	0.000	0.000	0.000	0.000
100	0.5	0.9	0.093	0.096	0.095	0.055	0.056	0.055
		0.95	0.043	0.046	0.045	0.035	0.031	0.031
		1	0.027	0.028	0.027	0.026	0.021	0.023
	1	0.9	0.091	0.099	0.103	0.055	0.052	0.051
		0.95	0.042	0.045	0.046	0.033	0.032	0.033
		1	0.026	0.026	0.027	0.026	0.022	0.022
	10	0.9	0.102	0.105	0.101	0.051	0.050	0.052
		0.95	0.047	0.054	0.052	0.031	0.030	0.028
		1	0.026	0.029	0.027	0.022	0.023	0.024
200	0.5	0.9	0.465	0.468	0.460	0.320	0.313	0.318
		0.95	0.156	0.158	0.156	0.100	0.097	0.096
		1	0.049	0.047	0.046	0.045	0.043	0.045
	1	0.9	0.488	0.502	0.496	0.347	0.339	0.338
		0.95	0.160	0.165	0.162	0.103	0.100	0.100
		1	0.046	0.047	0.045	0.044	0.042	0.043
	10	0.9	0.529	0.534	0.525	0.351	0.346	0.344
		0.95	0.152	0.162	0.163	0.104	0.100	0.099
		1	0.043	0.048	0.048	0.040	0.038	0.040

The DGP for the one break case is given by Equations (1.6.1) to (1.6.4). For the case on 1 known break, the values for the mean and the variance are calculated using $\lambda^0 = 0.5$. The DGP for the two breaks case is given by Equations (1.6.5) to (1.6.8). The values for the mean and the variance for the case of 2 known breaks are calculated using $\lambda_1^0 = 0.3$ and $\lambda_2^0 = 0.7$.

Table 1.9: Empirical size of panel unit root statistics for 2 known breaks for $N = 20$, $\lambda_1^0 = 0.3$ and $\lambda_2^0 = 0.7$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.043	0.046	0.045	0.047	0.045	0.049	0.059	0.057	0.058	0.050	0.042	0.048
		0.95	0.054	0.058	0.067	0.062	0.056	0.055	0.070	0.064	0.066	0.063	0.062	0.070
		1	0.042	0.055	0.056	0.054	0.047	0.042	0.067	0.057	0.059	0.059	0.053	0.057
	1	0.9	0.038	0.057	0.053	0.048	0.048	0.050	0.059	0.055	0.062	0.051	0.054	0.056
		0.95	0.044	0.051	0.052	0.058	0.057	0.057	0.075	0.069	0.068	0.057	0.049	0.054
		1	0.055	0.061	0.062	0.056	0.055	0.054	0.072	0.067	0.065	0.065	0.060	0.062
	10	0.9	0.039	0.053	0.059	0.052	0.050	0.056	0.062	0.060	0.071	0.051	0.052	0.061
		0.95	0.046	0.059	0.065	0.047	0.047	0.053	0.056	0.058	0.061	0.056	0.058	0.062
		1	0.039	0.048	0.060	0.049	0.047	0.045	0.063	0.060	0.060	0.050	0.050	0.057
100	0.5	0.9	0.053	0.066	0.048	0.054	0.062	0.045	0.066	0.072	0.058	0.056	0.058	0.047
		0.95	0.075	0.082	0.066	0.057	0.059	0.050	0.067	0.071	0.055	0.075	0.079	0.065
		1	0.050	0.067	0.050	0.059	0.065	0.045	0.071	0.073	0.060	0.056	0.059	0.050
	1	0.9	0.055	0.078	0.050	0.059	0.072	0.051	0.076	0.081	0.066	0.067	0.066	0.048
		0.95	0.056	0.062	0.045	0.049	0.052	0.038	0.056	0.061	0.050	0.057	0.061	0.044
		1	0.044	0.062	0.043	0.063	0.062	0.044	0.072	0.072	0.053	0.055	0.056	0.043
	10	0.9	0.060	0.067	0.053	0.063	0.067	0.050	0.075	0.075	0.061	0.060	0.062	0.052
		0.95	0.054	0.073	0.052	0.059	0.065	0.051	0.072	0.077	0.059	0.059	0.063	0.051
		1	0.055	0.051	0.043	0.045	0.051	0.040	0.055	0.054	0.049	0.050	0.050	0.042
200	0.5	0.9	0.049	0.057	0.050	0.046	0.054	0.051	0.064	0.072	0.060	0.053	0.051	0.048
		0.95	0.047	0.061	0.044	0.049	0.048	0.046	0.056	0.059	0.056	0.056	0.053	0.044
		1	0.046	0.056	0.048	0.044	0.046	0.041	0.054	0.059	0.052	0.057	0.053	0.046
	1	0.9	0.046	0.060	0.054	0.049	0.054	0.049	0.059	0.063	0.059	0.059	0.057	0.051
		0.95	0.050	0.053	0.047	0.049	0.056	0.050	0.056	0.064	0.056	0.053	0.050	0.049
		1	0.071	0.086	0.069	0.062	0.071	0.063	0.073	0.080	0.075	0.073	0.075	0.066
	10	0.9	0.055	0.063	0.055	0.049	0.054	0.053	0.063	0.071	0.062	0.061	0.058	0.051
		0.95	0.050	0.060	0.048	0.053	0.058	0.050	0.065	0.065	0.062	0.052	0.050	0.045
		1	0.050	0.054	0.046	0.052	0.057	0.052	0.061	0.063	0.058	0.052	0.049	0.043

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.10: Empirical power of panel unit root statistics for 2 known breaks for $N=20$, $\lambda_1^0=0.3$ and $\lambda_2^0=0.7$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.139	0.183	0.177	0.150	0.146	0.149	0.177	0.173	0.180	0.179	0.174	0.184
		0.95	0.156	0.186	0.197	0.149	0.155	0.156	0.172	0.170	0.183	0.185	0.187	0.195
		1	0.155	0.182	0.182	0.150	0.147	0.152	0.173	0.167	0.171	0.189	0.180	0.187
	1	0.9	0.133	0.162	0.170	0.139	0.131	0.130	0.161	0.148	0.156	0.167	0.160	0.175
		0.95	0.137	0.166	0.168	0.150	0.143	0.152	0.175	0.171	0.177	0.170	0.170	0.174
		1	0.114	0.158	0.163	0.136	0.129	0.135	0.155	0.156	0.163	0.154	0.156	0.166
	10	0.9	0.134	0.168	0.175	0.142	0.140	0.143	0.168	0.157	0.164	0.160	0.164	0.178
		0.95	0.155	0.173	0.186	0.148	0.145	0.151	0.167	0.168	0.170	0.170	0.165	0.181
		1	0.156	0.191	0.193	0.151	0.152	0.166	0.169	0.175	0.190	0.183	0.185	0.199
100	0.5	0.9	0.853	0.856	0.841	0.652	0.661	0.723	0.691	0.709	0.758	0.788	0.813	0.857
		0.95	0.888	0.888	0.866	0.691	0.705	0.772	0.732	0.754	0.798	0.830	0.849	0.877
		1	0.868	0.879	0.854	0.675	0.693	0.746	0.709	0.726	0.783	0.804	0.828	0.873
	1	0.9	0.873	0.873	0.856	0.700	0.709	0.767	0.728	0.751	0.800	0.822	0.843	0.871
		0.95	0.874	0.882	0.864	0.704	0.726	0.771	0.738	0.751	0.797	0.817	0.848	0.878
		1	0.879	0.880	0.856	0.677	0.696	0.751	0.717	0.731	0.778	0.825	0.846	0.870
	10	0.9	0.882	0.877	0.865	0.659	0.687	0.746	0.701	0.718	0.782	0.821	0.844	0.878
		0.95	0.878	0.866	0.852	0.689	0.701	0.746	0.719	0.733	0.779	0.821	0.830	0.855
		1	0.868	0.869	0.854	0.690	0.704	0.750	0.727	0.731	0.782	0.820	0.842	0.866
200	0.5	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.11: Empirical size of panel unit root statistics for 1 unknown break for $N = 20$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.040	0.069	0.059	0.044	0.056	0.053	0.053	0.064	0.058	0.049	0.065	0.058
		0.95	0.053	0.083	0.074	0.062	0.075	0.064	0.066	0.084	0.074	0.065	0.082	0.070
		1	0.041	0.077	0.066	0.049	0.069	0.056	0.066	0.077	0.067	0.052	0.075	0.064
	1	0.9	0.043	0.065	0.052	0.045	0.062	0.055	0.052	0.067	0.062	0.054	0.063	0.054
		0.95	0.048	0.079	0.068	0.063	0.075	0.068	0.068	0.082	0.076	0.065	0.076	0.070
		1	0.051	0.092	0.074	0.066	0.082	0.071	0.073	0.093	0.085	0.066	0.082	0.075
	10	0.9	0.137	0.196	0.175	0.180	0.198	0.192	0.194	0.206	0.198	0.166	0.186	0.176
		0.95	0.138	0.196	0.181	0.180	0.205	0.198	0.189	0.217	0.207	0.165	0.186	0.185
		1	0.134	0.195	0.184	0.173	0.195	0.191	0.181	0.200	0.199	0.158	0.185	0.188
100	0.5	0.9	0.044	0.049	0.064	0.033	0.036	0.047	0.044	0.044	0.055	0.043	0.046	0.061
		0.95	0.052	0.057	0.071	0.050	0.050	0.059	0.054	0.056	0.062	0.058	0.064	0.070
		1	0.044	0.049	0.058	0.034	0.033	0.040	0.039	0.039	0.047	0.040	0.042	0.058
	1	0.9	0.048	0.040	0.058	0.034	0.035	0.040	0.041	0.040	0.046	0.038	0.042	0.057
		0.95	0.048	0.048	0.064	0.038	0.040	0.052	0.044	0.050	0.057	0.045	0.049	0.063
		1	0.038	0.044	0.055	0.033	0.034	0.040	0.045	0.040	0.049	0.040	0.043	0.055
	10	0.9	0.057	0.078	0.095	0.064	0.068	0.075	0.072	0.074	0.085	0.065	0.081	0.093
		0.95	0.093	0.105	0.122	0.092	0.093	0.100	0.096	0.097	0.111	0.100	0.106	0.122
		1	0.101	0.122	0.137	0.108	0.115	0.125	0.113	0.123	0.132	0.115	0.123	0.136
200	0.5	0.9	0.044	0.038	0.049	0.024	0.024	0.025	0.025	0.027	0.026	0.043	0.045	0.047
		0.95	0.041	0.052	0.060	0.034	0.033	0.030	0.041	0.038	0.037	0.054	0.054	0.053
		1	0.037	0.040	0.045	0.030	0.030	0.029	0.033	0.033	0.035	0.041	0.041	0.042
	1	0.9	0.041	0.040	0.045	0.023	0.026	0.026	0.030	0.029	0.029	0.044	0.045	0.043
		0.95	0.049	0.037	0.054	0.025	0.027	0.026	0.031	0.030	0.030	0.051	0.052	0.051
		1	0.037	0.040	0.042	0.024	0.023	0.022	0.028	0.025	0.026	0.042	0.042	0.039
	10	0.9	0.052	0.050	0.066	0.030	0.029	0.032	0.035	0.036	0.037	0.053	0.062	0.062
		0.95	0.047	0.050	0.055	0.024	0.028	0.030	0.034	0.034	0.035	0.051	0.054	0.054
		1	0.078	0.068	0.075	0.052	0.053	0.053	0.059	0.058	0.059	0.073	0.072	0.069

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.12: Empirical power of panel unit root statistics for 1 unknown break for $N = 20$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.164	0.239	0.212	0.151	0.204	0.180	0.179	0.228	0.206	0.193	0.231	0.215
		0.95	0.158	0.241	0.206	0.154	0.191	0.172	0.181	0.219	0.194	0.197	0.239	0.215
		1	0.154	0.247	0.238	0.166	0.208	0.189	0.202	0.245	0.222	0.200	0.236	0.234
	1	0.9	0.124	0.214	0.193	0.124	0.165	0.149	0.149	0.185	0.178	0.149	0.204	0.194
		0.95	0.129	0.220	0.200	0.148	0.186	0.166	0.172	0.211	0.195	0.166	0.211	0.203
		1	0.107	0.196	0.173	0.123	0.150	0.136	0.140	0.174	0.154	0.137	0.195	0.174
	10	0.9	0.082	0.135	0.132	0.076	0.111	0.104	0.089	0.132	0.118	0.090	0.134	0.127
		0.95	0.087	0.129	0.123	0.082	0.105	0.102	0.093	0.123	0.115	0.097	0.134	0.127
		1	0.067	0.122	0.117	0.061	0.088	0.086	0.069	0.111	0.100	0.079	0.120	0.113
100	0.5	0.9	0.742	0.749	0.781	0.667	0.680	0.697	0.692	0.701	0.723	0.754	0.757	0.777
		0.95	0.641	0.651	0.670	0.568	0.574	0.604	0.587	0.598	0.629	0.647	0.660	0.668
		1	0.690	0.684	0.719	0.607	0.611	0.639	0.631	0.641	0.661	0.696	0.696	0.717
	1	0.9	0.598	0.595	0.630	0.532	0.538	0.564	0.552	0.556	0.582	0.598	0.608	0.632
		0.95	0.592	0.598	0.634	0.514	0.527	0.556	0.544	0.552	0.577	0.602	0.610	0.633
		1	0.544	0.568	0.592	0.497	0.508	0.531	0.516	0.528	0.550	0.563	0.570	0.589
	10	0.9	0.613	0.670	0.693	0.588	0.615	0.662	0.610	0.634	0.680	0.643	0.664	0.695
		0.95	0.657	0.702	0.739	0.637	0.658	0.689	0.658	0.681	0.705	0.684	0.702	0.735
		1	0.659	0.729	0.763	0.637	0.671	0.711	0.667	0.687	0.733	0.697	0.734	0.764
200	0.5	0.9	0.788	0.866	0.848	0.857	0.854	0.854	0.860	0.858	0.858	0.860	0.859	0.856
		0.95	0.733	0.803	0.797	0.797	0.795	0.793	0.801	0.801	0.800	0.802	0.801	0.798
		1	0.693	0.753	0.748	0.743	0.741	0.741	0.748	0.748	0.746	0.753	0.753	0.749
	1	0.9	0.677	0.770	0.756	0.755	0.758	0.754	0.769	0.768	0.766	0.766	0.768	0.761
		0.95	0.651	0.723	0.714	0.710	0.708	0.706	0.719	0.719	0.716	0.722	0.722	0.718
		1	0.611	0.685	0.681	0.670	0.671	0.672	0.676	0.677	0.675	0.688	0.689	0.683
	10	0.9	0.635	0.743	0.725	0.739	0.739	0.738	0.747	0.747	0.747	0.727	0.732	0.732
		0.95	0.699	0.789	0.774	0.786	0.789	0.789	0.792	0.792	0.796	0.779	0.779	0.779
		1	0.693	0.787	0.770	0.779	0.779	0.782	0.786	0.789	0.789	0.774	0.775	0.775

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.13: Empirical size and power of the M-type unit root statistics for unknown breaks for $N = 20$ for 1 common factor

			Panel A: 1 unknown break			Panel B: 2 unknown breaks		
T	σ_F^2	ρ	MSB_F^{GLS}	$MZ_{F,\alpha}^{GLS}$	$MZ_{F,t}^{GLS}$	MSB_F^{GLS}	$MZ_{F,\alpha}^{GLS}$	$MZ_{F,t}^{GLS}$
50	0.5	0.9	0.000	0.000	0.000	0.000	0.000	0.000
		0.95	0.000	0.000	0.000	0.000	0.000	0.000
		1	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.9	0.000	0.000	0.000	0.000	0.000	0.000
		0.95	0.000	0.000	0.000	0.000	0.000	0.000
		1	0.000	0.000	0.000	0.000	0.000	0.000
	10	0.9	0.000	0.000	0.000	0.000	0.000	0.000
		0.95	0.000	0.000	0.000	0.000	0.000	0.000
		1	0.000	0.000	0.000	0.000	0.000	0.000
100	0.5	0.9	0.083	0.085	0.082	0.120	0.116	0.117
		0.95	0.049	0.051	0.049	0.076	0.071	0.071
		1	0.041	0.044	0.044	0.064	0.062	0.067
	1	0.9	0.123	0.130	0.127	0.133	0.120	0.120
		0.95	0.069	0.067	0.068	0.088	0.080	0.084
		1	0.057	0.057	0.056	0.071	0.065	0.066
	10	0.9	0.171	0.179	0.179	0.137	0.129	0.132
		0.95	0.097	0.097	0.094	0.086	0.078	0.079
		1	0.074	0.067	0.066	0.081	0.078	0.075
200	0.5	0.9	0.192	0.200	0.202	0.379	0.367	0.365
		0.95	0.062	0.067	0.069	0.157	0.148	0.145
		1	0.023	0.022	0.022	0.096	0.090	0.091
	1	0.9	0.258	0.265	0.265	0.431	0.423	0.423
		0.95	0.084	0.089	0.087	0.187	0.182	0.183
		1	0.032	0.033	0.032	0.115	0.109	0.110
	10	0.9	0.510	0.513	0.511	0.506	0.497	0.494
		0.95	0.192	0.193	0.186	0.229	0.215	0.216
		1	0.071	0.075	0.072	0.113	0.106	0.105

The DGP for the one break case is given by Equations (1.6.1) to (1.6.4). The DGP for the two breaks case is given by Equations (1.6.5) to (1.6.8).

Table 1.14: Empirical size of panel unit root statistics for 2 unknown breaks for $N = 20$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.022	0.038	0.033	0.035	0.034	0.039	0.043	0.041	0.044	0.037	0.037	0.037
		0.95	0.027	0.040	0.033	0.029	0.027	0.030	0.039	0.032	0.041	0.038	0.039	0.035
		1	0.020	0.035	0.033	0.025	0.028	0.036	0.033	0.035	0.043	0.032	0.033	0.039
	1	0.9	0.025	0.039	0.033	0.031	0.030	0.033	0.036	0.036	0.040	0.039	0.037	0.030
		0.95	0.022	0.043	0.043	0.046	0.045	0.045	0.051	0.049	0.051	0.042	0.043	0.044
		1	0.020	0.035	0.031	0.033	0.033	0.037	0.037	0.041	0.042	0.037	0.036	0.037
	10	0.9	0.087	0.112	0.114	0.114	0.119	0.131	0.123	0.129	0.136	0.108	0.114	0.120
		0.95	0.086	0.108	0.105	0.123	0.128	0.133	0.129	0.131	0.137	0.110	0.113	0.116
		1	0.079	0.117	0.109	0.122	0.131	0.137	0.130	0.137	0.140	0.114	0.116	0.123
100	0.5	0.9	0.047	0.043	0.062	0.043	0.029	0.039	0.046	0.040	0.049	0.051	0.044	0.060
		0.95	0.043	0.046	0.052	0.040	0.034	0.040	0.049	0.039	0.047	0.053	0.048	0.051
		1	0.039	0.046	0.059	0.045	0.035	0.040	0.048	0.043	0.050	0.055	0.047	0.061
	1	0.9	0.041	0.037	0.048	0.032	0.027	0.033	0.041	0.031	0.038	0.043	0.039	0.045
		0.95	0.036	0.034	0.044	0.032	0.023	0.025	0.036	0.027	0.034	0.038	0.035	0.043
		1	0.033	0.032	0.040	0.028	0.021	0.029	0.031	0.025	0.035	0.032	0.032	0.040
	10	0.9	0.061	0.064	0.072	0.060	0.053	0.062	0.066	0.060	0.070	0.072	0.068	0.072
		0.95	0.071	0.076	0.079	0.075	0.073	0.076	0.081	0.076	0.079	0.079	0.075	0.080
		1	0.062	0.075	0.079	0.078	0.072	0.076	0.086	0.077	0.082	0.079	0.074	0.080
200	0.5	0.9	0.099	0.042	0.069	0.024	0.022	0.023	0.031	0.025	0.027	0.057	0.047	0.046
		0.95	0.083	0.029	0.054	0.017	0.014	0.014	0.022	0.017	0.020	0.049	0.038	0.034
		1	0.092	0.048	0.057	0.029	0.023	0.024	0.033	0.029	0.029	0.058	0.051	0.052
	1	0.9	0.078	0.032	0.055	0.024	0.019	0.019	0.026	0.021	0.022	0.045	0.034	0.034
		0.95	0.079	0.032	0.052	0.017	0.012	0.010	0.019	0.017	0.018	0.044	0.036	0.035
		1	0.107	0.040	0.070	0.025	0.024	0.023	0.029	0.027	0.027	0.053	0.044	0.046
	10	0.9	0.082	0.048	0.060	0.031	0.026	0.024	0.034	0.030	0.029	0.058	0.050	0.051
		0.95	0.061	0.031	0.044	0.020	0.017	0.017	0.027	0.023	0.022	0.044	0.036	0.032
		1	0.059	0.034	0.044	0.022	0.019	0.019	0.023	0.022	0.021	0.042	0.037	0.037

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.15: Empirical power of panel unit root statistics for 2 unknown breaks for $N = 20$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.055	0.092	0.082	0.087	0.087	0.100	0.098	0.100	0.112	0.091	0.092	0.097
		0.95	0.049	0.084	0.080	0.090	0.083	0.090	0.108	0.098	0.101	0.088	0.085	0.089
		1	0.056	0.100	0.090	0.075	0.078	0.086	0.097	0.094	0.106	0.092	0.098	0.102
	1	0.9	0.055	0.086	0.085	0.076	0.074	0.086	0.091	0.086	0.097	0.081	0.087	0.085
		0.95	0.059	0.098	0.087	0.075	0.070	0.075	0.086	0.082	0.087	0.085	0.098	0.097
		1	0.058	0.086	0.085	0.070	0.076	0.074	0.081	0.081	0.092	0.077	0.085	0.088
	10	0.9	0.149	0.176	0.180	0.169	0.179	0.191	0.177	0.186	0.198	0.166	0.174	0.190
		0.95	0.128	0.178	0.169	0.178	0.182	0.188	0.183	0.187	0.194	0.170	0.180	0.185
		1	0.117	0.170	0.161	0.162	0.175	0.184	0.173	0.181	0.193	0.158	0.170	0.177
100	0.5	0.9	0.433	0.466	0.501	0.409	0.384	0.418	0.441	0.418	0.452	0.472	0.471	0.501
		0.95	0.395	0.414	0.445	0.354	0.334	0.359	0.382	0.360	0.389	0.419	0.415	0.443
		1	0.409	0.408	0.447	0.370	0.333	0.366	0.391	0.365	0.392	0.429	0.419	0.443
	1	0.9	0.340	0.347	0.379	0.306	0.281	0.311	0.330	0.308	0.330	0.362	0.357	0.379
		0.95	0.341	0.360	0.397	0.323	0.291	0.310	0.340	0.314	0.337	0.368	0.364	0.395
		1	0.343	0.351	0.384	0.298	0.277	0.307	0.322	0.305	0.332	0.369	0.359	0.383
	10	0.9	0.269	0.287	0.304	0.246	0.233	0.249	0.260	0.252	0.267	0.294	0.293	0.304
		0.95	0.283	0.296	0.312	0.277	0.259	0.274	0.288	0.279	0.286	0.297	0.295	0.312
		1	0.227	0.258	0.265	0.249	0.238	0.254	0.262	0.252	0.263	0.253	0.255	0.266
200	0.5	0.9	0.593	0.627	0.614	0.610	0.590	0.589	0.625	0.601	0.600	0.650	0.636	0.627
		0.95	0.557	0.582	0.581	0.567	0.548	0.545	0.580	0.562	0.559	0.604	0.590	0.588
		1	0.541	0.562	0.564	0.532	0.517	0.514	0.545	0.525	0.524	0.581	0.567	0.568
	1	0.9	0.532	0.553	0.556	0.519	0.497	0.494	0.542	0.518	0.514	0.571	0.559	0.555
		0.95	0.527	0.517	0.522	0.491	0.471	0.470	0.505	0.488	0.487	0.539	0.522	0.517
		1	0.534	0.544	0.558	0.516	0.503	0.500	0.529	0.514	0.513	0.568	0.547	0.546
	10	0.9	0.489	0.462	0.477	0.414	0.403	0.405	0.425	0.414	0.416	0.481	0.465	0.466
		0.95	0.481	0.413	0.440	0.377	0.368	0.358	0.390	0.374	0.373	0.443	0.423	0.420
		1	0.407	0.375	0.392	0.356	0.353	0.352	0.362	0.356	0.359	0.396	0.381	0.379

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.16: Individual unit root tests and the two break dates

Country	$\tilde{T}_1$	$\tilde{T}_2$	λ used for cv	MSB^{GLS}	MZ_{α}^{GLS}	MZ_t^{GLS}
Australia	1889	1931	(0.2, 0.5)	0.170	-17.130	-2.910
Austria	1913	1947	(0.3, 0.6)	0.165	-17.607	-2.913
Belgium	1918	1943	(0.4, 0.6)	0.200	-12.370	-2.472
Canada	1917	1933	(0.4, 0.5)	0.192	-13.105	-2.509
Denmark	1939	1973	(0.5, 0.8)	0.147**	-23.219**	-3.402**
Finland	1892	1918	(0.2, 0.4)	0.194	-13.342	-2.583
France	1929	1945	(0.4, 0.6)	0.165	-17.122	-2.830
Germany	1913	1944	(0.3, 0.6)	0.152*	-20.914	-3.183
Italy	1918	1945	(0.4, 0.6)	0.295	-4.447	-1.312
Japan	1950	1973	(0.6, 0.8)	0.258	-6.836	-1.763
Netherlands	1918	1945	(0.4, 0.6)	0.168	-17.679	-2.972
New Zealand	1920	1934	(0.4, 0.5)	0.145**	-23.637**	-3.436**
Norway	1916	1939	(0.4, 0.5)	0.161	-17.394	-2.802
Portugal	1936	1961	(0.5, 0.7)	0.203	-10.688	-2.167
Spain	1935	1960	(0.5, 0.7)	0.265	-7.142	-1.890
Sweden	1916	1939	(0.4, 0.5)	0.229	-9.097	-2.083
Switzerland	1907	1944	(0.3, 0.6)	0.184	-14.497	-2.668
UK	1918	1943	(0.4, 0.6)	0.150*	-21.756*	-3.254*
USA	1929	1944	(0.4, 0.6)	0.143**	-24.572**	-3.505**

$\tilde{T}_1$ and $\tilde{T}_2$ represent the break dates and cv denotes the critical value; * and ** denote rejection of the null hypothesis at the 10% and 5% level of significance, respectively.

Table 1.17: Panel unit root tests

Panel test	MSB^{GLS}	MZ_{α}^{GLS}	MZ_t^{GLS}
Z^M	-1.5616*	-1.7372**	-1.6879**
P^M	53.4213**	51.9885*	52.7378*
P_m^M	1.7689**	1.6046*	1.6905**
C^M	-1.7764**	-1.6246*	-1.6712**

Note: * and ** denote rejection of the null hypothesis at the 10% and 5% level of significance, respectively.

Table 1.18: Unit root tests for common factors

MSB_F^{GLS}	$MZ_{F,\alpha}^{GLS}$	$MZ_{F,t}^{GLS}$
0.1818	-14.2714	-2.5951

Table 1.19: Empirical size of panel unit root statistics for 1 known break for $N = 20$ and $\lambda^0 = 0.3$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.043	0.057	0.043	0.037	0.046	0.047	0.040	0.063	0.065	0.042	0.045	0.048
		0.95	0.058	0.067	0.053	0.039	0.056	0.049	0.063	0.069	0.066	0.059	0.057	0.056
		1	0.061	0.074	0.067	0.066	0.074	0.066	0.077	0.088	0.080	0.064	0.072	0.066
	1	0.9	0.043	0.057	0.042	0.048	0.052	0.055	0.056	0.063	0.063	0.050	0.054	0.044
		0.95	0.060	0.067	0.059	0.057	0.059	0.053	0.064	0.067	0.064	0.060	0.067	0.055
		1	0.059	0.078	0.061	0.067	0.074	0.067	0.078	0.088	0.083	0.066	0.068	0.060
	10	0.9	0.050	0.063	0.054	0.051	0.054	0.050	0.060	0.064	0.060	0.056	0.059	0.057
		0.95	0.063	0.072	0.066	0.048	0.065	0.063	0.062	0.077	0.074	0.070	0.075	0.070
		1	0.058	0.073	0.067	0.060	0.061	0.052	0.072	0.073	0.063	0.061	0.068	0.064
100	0.5	0.9	0.051	0.055	0.048	0.041	0.052	0.052	0.050	0.059	0.059	0.054	0.055	0.048
		0.95	0.049	0.058	0.054	0.044	0.051	0.055	0.055	0.063	0.068	0.050	0.053	0.052
		1	0.049	0.060	0.054	0.039	0.046	0.048	0.044	0.053	0.054	0.053	0.055	0.053
	1	0.9	0.040	0.053	0.045	0.041	0.045	0.046	0.044	0.051	0.060	0.046	0.049	0.045
		0.95	0.043	0.064	0.060	0.045	0.052	0.056	0.056	0.068	0.067	0.047	0.062	0.060
		1	0.045	0.061	0.048	0.040	0.048	0.054	0.050	0.064	0.067	0.046	0.051	0.048
	10	0.9	0.050	0.069	0.052	0.054	0.063	0.063	0.064	0.077	0.075	0.049	0.059	0.053
		0.95	0.035	0.053	0.046	0.038	0.048	0.048	0.049	0.060	0.060	0.043	0.049	0.046
		1	0.055	0.074	0.066	0.041	0.055	0.056	0.051	0.067	0.069	0.054	0.068	0.069
200	0.5	0.9	0.043	0.070	0.054	0.054	0.057	0.048	0.067	0.071	0.054	0.056	0.059	0.053
		0.95	0.047	0.060	0.044	0.049	0.051	0.040	0.063	0.061	0.050	0.052	0.050	0.040
		1	0.055	0.074	0.057	0.056	0.058	0.049	0.066	0.073	0.058	0.064	0.072	0.054
	1	0.9	0.043	0.064	0.047	0.047	0.056	0.049	0.066	0.069	0.056	0.051	0.054	0.043
		0.95	0.044	0.071	0.054	0.064	0.065	0.048	0.073	0.075	0.064	0.065	0.063	0.051
		1	0.057	0.079	0.063	0.061	0.065	0.057	0.069	0.076	0.065	0.066	0.073	0.062
	10	0.9	0.059	0.083	0.061	0.068	0.073	0.054	0.086	0.088	0.072	0.067	0.077	0.062
		0.95	0.044	0.070	0.045	0.050	0.059	0.046	0.068	0.070	0.059	0.051	0.057	0.043
		1	0.053	0.077	0.061	0.067	0.069	0.060	0.073	0.081	0.069	0.056	0.060	0.056

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.20: Empirical size of panel unit root statistics for 1 known break for $N = 20$ and $\lambda^0 = 0.7$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.046	0.062	0.053	0.043	0.052	0.050	0.050	0.068	0.063	0.056	0.061	0.057
		0.95	0.056	0.066	0.059	0.044	0.061	0.046	0.061	0.071	0.060	0.062	0.066	0.054
		1	0.056	0.071	0.058	0.052	0.067	0.053	0.059	0.075	0.068	0.060	0.069	0.061
	1	0.9	0.031	0.056	0.039	0.030	0.052	0.043	0.045	0.061	0.056	0.041	0.054	0.041
		0.95	0.050	0.065	0.056	0.035	0.052	0.044	0.047	0.067	0.056	0.055	0.063	0.053
		1	0.043	0.066	0.061	0.036	0.052	0.047	0.045	0.065	0.056	0.051	0.061	0.055
	10	0.9	0.052	0.076	0.055	0.051	0.065	0.056	0.059	0.079	0.073	0.061	0.065	0.057
		0.95	0.056	0.088	0.062	0.040	0.060	0.047	0.056	0.075	0.065	0.067	0.078	0.065
		1	0.046	0.079	0.061	0.056	0.071	0.067	0.072	0.083	0.077	0.064	0.070	0.059
100	0.5	0.9	0.041	0.056	0.064	0.047	0.047	0.058	0.052	0.057	0.074	0.052	0.055	0.064
		0.95	0.041	0.063	0.062	0.059	0.058	0.071	0.064	0.068	0.079	0.054	0.057	0.058
		1	0.053	0.061	0.070	0.049	0.053	0.065	0.057	0.063	0.071	0.060	0.063	0.067
	1	0.9	0.036	0.052	0.056	0.036	0.038	0.049	0.047	0.048	0.064	0.045	0.048	0.054
		0.95	0.038	0.064	0.058	0.049	0.054	0.065	0.061	0.063	0.078	0.047	0.049	0.055
		1	0.050	0.061	0.058	0.046	0.052	0.064	0.058	0.062	0.081	0.054	0.053	0.059
	10	0.9	0.044	0.061	0.059	0.047	0.051	0.062	0.057	0.064	0.075	0.053	0.054	0.057
		0.95	0.055	0.072	0.068	0.058	0.056	0.068	0.063	0.069	0.077	0.062	0.063	0.069
		1	0.049	0.064	0.069	0.049	0.048	0.066	0.061	0.069	0.083	0.059	0.063	0.070
200	0.5	0.9	0.037	0.066	0.061	0.043	0.059	0.050	0.057	0.074	0.061	0.049	0.064	0.056
		0.95	0.041	0.056	0.045	0.037	0.049	0.040	0.045	0.055	0.047	0.038	0.047	0.042
		1	0.046	0.086	0.068	0.066	0.078	0.067	0.076	0.091	0.079	0.058	0.076	0.064
	1	0.9	0.048	0.074	0.061	0.044	0.063	0.049	0.062	0.072	0.062	0.051	0.070	0.060
		0.95	0.035	0.056	0.042	0.041	0.056	0.045	0.053	0.067	0.054	0.038	0.046	0.039
		1	0.048	0.067	0.057	0.040	0.062	0.046	0.057	0.075	0.065	0.056	0.067	0.057
	10	0.9	0.050	0.078	0.063	0.054	0.064	0.054	0.063	0.076	0.069	0.057	0.072	0.059
		0.95	0.035	0.081	0.065	0.051	0.068	0.061	0.064	0.079	0.067	0.056	0.067	0.055
		1	0.034	0.076	0.060	0.054	0.068	0.058	0.063	0.081	0.068	0.053	0.068	0.054

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.21: Empirical size of panel unit root statistics for 1 known break for $N = 40$ and $\lambda^0 = 0.5$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.056	0.061	0.057	0.050	0.054	0.050	0.061	0.062	0.058	0.058	0.055	0.052
		0.95	0.055	0.067	0.063	0.059	0.067	0.064	0.071	0.075	0.068	0.066	0.067	0.069
		1	0.058	0.070	0.065	0.047	0.055	0.050	0.058	0.064	0.060	0.060	0.064	0.067
	1	0.9	0.031	0.051	0.044	0.049	0.053	0.051	0.058	0.062	0.056	0.043	0.044	0.043
		0.95	0.053	0.057	0.057	0.048	0.050	0.050	0.057	0.058	0.055	0.054	0.056	0.056
		1	0.059	0.070	0.069	0.063	0.065	0.060	0.071	0.075	0.065	0.061	0.068	0.071
	10	0.9	0.048	0.057	0.054	0.052	0.056	0.053	0.058	0.061	0.061	0.051	0.055	0.054
		0.95	0.053	0.056	0.061	0.050	0.051	0.049	0.058	0.059	0.053	0.054	0.060	0.062
		1	0.046	0.058	0.057	0.056	0.051	0.055	0.060	0.068	0.068	0.058	0.058	0.057
100	0.5	0.9	0.036	0.046	0.047	0.061	0.063	0.052	0.069	0.072	0.059	0.047	0.052	0.042
		0.95	0.051	0.054	0.060	0.064	0.066	0.056	0.071	0.072	0.063	0.064	0.065	0.053
		1	0.040	0.055	0.059	0.061	0.069	0.056	0.074	0.077	0.063	0.061	0.066	0.055
	1	0.9	0.043	0.046	0.051	0.057	0.062	0.050	0.066	0.073	0.060	0.052	0.064	0.049
		0.95	0.050	0.053	0.059	0.065	0.069	0.060	0.079	0.080	0.068	0.060	0.069	0.055
		1	0.056	0.060	0.066	0.076	0.077	0.062	0.085	0.087	0.071	0.071	0.074	0.058
	10	0.9	0.040	0.041	0.046	0.058	0.059	0.049	0.063	0.068	0.056	0.046	0.048	0.040
		0.95	0.047	0.070	0.071	0.078	0.079	0.063	0.081	0.085	0.074	0.071	0.083	0.065
		1	0.058	0.058	0.066	0.070	0.076	0.056	0.078	0.091	0.069	0.064	0.074	0.059
200	0.5	0.9	0.045	0.054	0.038	0.041	0.032	0.033	0.053	0.040	0.039	0.049	0.041	0.047
		0.95	0.041	0.047	0.039	0.044	0.037	0.042	0.047	0.041	0.043	0.049	0.040	0.047
		1	0.046	0.047	0.037	0.037	0.027	0.027	0.042	0.034	0.041	0.047	0.041	0.046
	1	0.9	0.055	0.061	0.055	0.046	0.041	0.045	0.055	0.045	0.056	0.053	0.052	0.064
		0.95	0.059	0.072	0.053	0.065	0.052	0.055	0.077	0.064	0.064	0.062	0.049	0.061
		1	0.055	0.074	0.054	0.067	0.055	0.061	0.076	0.062	0.070	0.063	0.052	0.062
	10	0.9	0.037	0.058	0.043	0.053	0.048	0.053	0.063	0.054	0.057	0.046	0.038	0.047
		0.95	0.055	0.053	0.035	0.048	0.043	0.041	0.056	0.048	0.051	0.054	0.041	0.045
		1	0.050	0.065	0.061	0.050	0.046	0.054	0.064	0.056	0.064	0.062	0.053	0.061

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.22: Empirical size of panel unit root statistics for 1 known break for $N = 60$ and $\lambda^0 = 0.5$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.045	0.051	0.050	0.045	0.051	0.049	0.051	0.059	0.057	0.050	0.052	0.051
		0.95	0.050	0.058	0.060	0.049	0.055	0.055	0.055	0.063	0.063	0.048	0.061	0.061
		1	0.054	0.056	0.056	0.048	0.052	0.046	0.056	0.059	0.056	0.060	0.058	0.055
	1	0.9	0.042	0.056	0.047	0.047	0.052	0.045	0.055	0.057	0.050	0.052	0.050	0.046
		0.95	0.051	0.071	0.062	0.061	0.061	0.055	0.064	0.064	0.060	0.063	0.065	0.058
		1	0.044	0.049	0.045	0.043	0.050	0.045	0.051	0.053	0.052	0.055	0.050	0.046
	10	0.9	0.032	0.042	0.038	0.039	0.047	0.041	0.046	0.054	0.047	0.036	0.037	0.040
		0.95	0.056	0.071	0.072	0.049	0.056	0.055	0.053	0.065	0.065	0.059	0.072	0.073
		1	0.052	0.066	0.068	0.046	0.053	0.056	0.054	0.062	0.064	0.054	0.062	0.066
100	0.5	0.9	0.042	0.037	0.043	0.050	0.051	0.039	0.054	0.063	0.045	0.044	0.050	0.040
		0.95	0.049	0.048	0.062	0.056	0.064	0.049	0.065	0.075	0.058	0.062	0.072	0.057
		1	0.054	0.051	0.061	0.052	0.055	0.043	0.057	0.063	0.049	0.061	0.075	0.056
	1	0.9	0.038	0.044	0.047	0.056	0.059	0.044	0.065	0.066	0.047	0.052	0.056	0.042
		0.95	0.048	0.049	0.055	0.055	0.064	0.043	0.068	0.075	0.051	0.061	0.068	0.051
		1	0.053	0.048	0.058	0.064	0.068	0.046	0.072	0.077	0.051	0.063	0.069	0.054
	10	0.9	0.043	0.035	0.050	0.048	0.050	0.038	0.054	0.059	0.044	0.053	0.056	0.046
		0.95	0.035	0.052	0.052	0.063	0.068	0.050	0.070	0.077	0.061	0.053	0.061	0.047
		1	0.055	0.059	0.067	0.067	0.072	0.056	0.074	0.085	0.063	0.072	0.083	0.058
200	0.5	0.9	0.051	0.062	0.046	0.052	0.040	0.046	0.060	0.048	0.054	0.059	0.045	0.055
		0.95	0.057	0.062	0.048	0.056	0.044	0.050	0.060	0.052	0.057	0.057	0.048	0.063
		1	0.066	0.072	0.049	0.062	0.053	0.058	0.071	0.058	0.063	0.066	0.049	0.060
	1	0.9	0.069	0.070	0.054	0.057	0.051	0.050	0.062	0.052	0.057	0.074	0.057	0.065
		0.95	0.047	0.051	0.038	0.045	0.033	0.037	0.051	0.041	0.046	0.049	0.038	0.050
		1	0.058	0.058	0.041	0.050	0.035	0.046	0.057	0.048	0.049	0.064	0.044	0.058
	10	0.9	0.046	0.059	0.037	0.049	0.040	0.045	0.058	0.045	0.052	0.047	0.039	0.051
		0.95	0.059	0.063	0.049	0.054	0.045	0.051	0.063	0.051	0.054	0.064	0.052	0.067
		1	0.061	0.058	0.049	0.053	0.041	0.048	0.063	0.047	0.053	0.063	0.049	0.065

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.23: Empirical size of panel unit root statistics for 2 known breaks for $N = 40$, $\lambda_1^0 = 0.3$ and $\lambda_2^0 = 0.7$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.049	0.061	0.067	0.052	0.049	0.055	0.064	0.059	0.059	0.058	0.057	0.062
		0.95	0.054	0.054	0.059	0.047	0.042	0.051	0.057	0.054	0.058	0.052	0.054	0.060
		1	0.049	0.055	0.060	0.061	0.061	0.059	0.071	0.064	0.067	0.059	0.055	0.061
	1	0.9	0.049	0.053	0.056	0.049	0.046	0.046	0.062	0.053	0.056	0.058	0.050	0.051
		0.95	0.044	0.051	0.048	0.055	0.051	0.052	0.065	0.061	0.061	0.057	0.055	0.055
		1	0.040	0.051	0.053	0.056	0.047	0.050	0.065	0.061	0.060	0.051	0.049	0.053
	10	0.9	0.046	0.052	0.053	0.058	0.045	0.047	0.062	0.057	0.060	0.053	0.051	0.055
		0.95	0.051	0.052	0.059	0.051	0.045	0.049	0.060	0.052	0.057	0.050	0.052	0.057
		1	0.044	0.046	0.045	0.055	0.051	0.055	0.062	0.058	0.063	0.045	0.047	0.049
100	0.5	0.9	0.062	0.069	0.052	0.041	0.040	0.054	0.046	0.046	0.059	0.044	0.043	0.059
		0.95	0.061	0.074	0.047	0.038	0.036	0.055	0.043	0.045	0.061	0.036	0.042	0.053
		1	0.064	0.077	0.047	0.037	0.039	0.063	0.044	0.050	0.077	0.034	0.040	0.056
	1	0.9	0.054	0.071	0.046	0.042	0.041	0.056	0.044	0.045	0.069	0.033	0.039	0.049
		0.95	0.077	0.084	0.052	0.042	0.044	0.058	0.046	0.050	0.065	0.037	0.043	0.061
		1	0.067	0.087	0.059	0.044	0.049	0.066	0.048	0.053	0.071	0.034	0.046	0.065
	10	0.9	0.054	0.072	0.043	0.029	0.033	0.056	0.037	0.043	0.065	0.024	0.032	0.049
		0.95	0.079	0.091	0.054	0.034	0.034	0.056	0.037	0.041	0.062	0.040	0.047	0.061
		1	0.075	0.091	0.068	0.044	0.044	0.070	0.047	0.054	0.078	0.041	0.057	0.074
200	0.5	0.9	0.060	0.065	0.050	0.072	0.045	0.046	0.084	0.057	0.049	0.067	0.063	0.045
		0.95	0.052	0.066	0.051	0.070	0.056	0.046	0.078	0.064	0.055	0.066	0.066	0.043
		1	0.043	0.048	0.038	0.052	0.037	0.033	0.065	0.049	0.039	0.046	0.050	0.034
	1	0.9	0.047	0.058	0.045	0.058	0.045	0.041	0.064	0.048	0.048	0.055	0.054	0.043
		0.95	0.053	0.057	0.046	0.062	0.052	0.044	0.073	0.054	0.056	0.058	0.060	0.044
		1	0.054	0.054	0.044	0.064	0.046	0.048	0.078	0.051	0.055	0.056	0.057	0.038
	10	0.9	0.049	0.060	0.045	0.065	0.051	0.049	0.070	0.056	0.053	0.056	0.057	0.042
		0.95	0.059	0.055	0.042	0.048	0.038	0.035	0.060	0.046	0.041	0.059	0.056	0.040
		1	0.057	0.059	0.047	0.061	0.053	0.050	0.069	0.055	0.058	0.057	0.061	0.044

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.24: Empirical size of panel unit root statistics for 2 known breaks for $N = 60$, $\lambda_1^0 = 0.3$ and $\lambda_2^0 = 0.7$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.061	0.071	0.079	0.061	0.054	0.062	0.071	0.067	0.068	0.079	0.071	0.079
		0.95	0.043	0.050	0.058	0.055	0.053	0.056	0.065	0.061	0.067	0.054	0.049	0.055
		1	0.056	0.065	0.072	0.064	0.055	0.066	0.072	0.064	0.072	0.067	0.061	0.072
	1	0.9	0.043	0.042	0.047	0.054	0.054	0.054	0.067	0.057	0.062	0.046	0.043	0.046
		0.95	0.047	0.051	0.056	0.047	0.038	0.046	0.055	0.046	0.053	0.055	0.047	0.053
		1	0.053	0.064	0.061	0.063	0.052	0.058	0.068	0.060	0.065	0.066	0.065	0.063
	10	0.9	0.050	0.046	0.051	0.053	0.045	0.046	0.060	0.050	0.055	0.048	0.043	0.049
		0.95	0.049	0.062	0.065	0.066	0.056	0.056	0.074	0.063	0.070	0.065	0.061	0.064
		1	0.050	0.056	0.064	0.058	0.057	0.060	0.066	0.062	0.072	0.061	0.058	0.065
100	0.5	0.9	0.064	0.079	0.044	0.026	0.032	0.050	0.030	0.035	0.055	0.028	0.032	0.050
		0.95	0.068	0.077	0.045	0.033	0.031	0.060	0.037	0.041	0.066	0.031	0.039	0.053
		1	0.080	0.104	0.061	0.036	0.035	0.066	0.040	0.040	0.077	0.035	0.047	0.069
	1	0.9	0.072	0.083	0.047	0.034	0.039	0.057	0.039	0.044	0.061	0.026	0.033	0.055
		0.95	0.070	0.081	0.049	0.034	0.036	0.052	0.036	0.039	0.058	0.034	0.043	0.055
		1	0.068	0.075	0.042	0.027	0.032	0.046	0.030	0.038	0.057	0.022	0.027	0.046
	10	0.9	0.060	0.077	0.043	0.025	0.033	0.058	0.031	0.039	0.066	0.024	0.036	0.050
		0.95	0.063	0.080	0.037	0.033	0.036	0.056	0.040	0.040	0.066	0.028	0.034	0.045
		1	0.085	0.096	0.065	0.038	0.046	0.071	0.047	0.055	0.082	0.046	0.055	0.071
200	0.5	0.9	0.059	0.056	0.047	0.062	0.040	0.039	0.068	0.047	0.045	0.060	0.058	0.041
		0.95	0.058	0.055	0.040	0.057	0.042	0.038	0.067	0.049	0.044	0.060	0.061	0.035
		1	0.060	0.065	0.046	0.081	0.054	0.051	0.089	0.063	0.056	0.069	0.064	0.042
	1	0.9	0.051	0.061	0.044	0.078	0.057	0.051	0.088	0.062	0.058	0.061	0.061	0.038
		0.95	0.038	0.043	0.026	0.056	0.042	0.039	0.063	0.050	0.049	0.042	0.039	0.021
		1	0.061	0.068	0.051	0.068	0.051	0.050	0.076	0.061	0.056	0.066	0.067	0.044
	10	0.9	0.047	0.055	0.040	0.066	0.047	0.043	0.069	0.056	0.050	0.051	0.056	0.031
		0.95	0.047	0.054	0.038	0.060	0.043	0.038	0.066	0.048	0.043	0.055	0.051	0.033
		1	0.059	0.062	0.050	0.073	0.056	0.050	0.075	0.062	0.055	0.060	0.061	0.045

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.25: Empirical size of panel unit root statistics for 1 unknown break for $N = 40$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.050	0.078	0.061	0.054	0.073	0.054	0.057	0.080	0.061	0.057	0.071	0.063
		0.95	0.049	0.084	0.068	0.045	0.065	0.052	0.052	0.076	0.062	0.053	0.082	0.068
		1	0.059	0.090	0.079	0.052	0.073	0.057	0.060	0.082	0.067	0.069	0.094	0.082
	1	0.9	0.053	0.101	0.081	0.072	0.099	0.090	0.080	0.114	0.097	0.068	0.094	0.079
		0.95	0.067	0.099	0.088	0.078	0.093	0.085	0.079	0.099	0.091	0.073	0.096	0.089
		1	0.080	0.121	0.103	0.073	0.104	0.089	0.085	0.115	0.101	0.091	0.117	0.101
	10	0.9	0.160	0.224	0.215	0.195	0.226	0.220	0.202	0.232	0.225	0.193	0.221	0.218
		0.95	0.182	0.239	0.231	0.220	0.250	0.245	0.226	0.257	0.255	0.205	0.235	0.234
		1	0.156	0.215	0.201	0.200	0.232	0.219	0.206	0.239	0.231	0.188	0.201	0.200
100	0.5	0.9	0.066	0.055	0.076	0.037	0.037	0.048	0.042	0.040	0.057	0.060	0.060	0.077
		0.95	0.069	0.059	0.083	0.039	0.040	0.048	0.043	0.044	0.057	0.065	0.062	0.082
		1	0.054	0.050	0.087	0.034	0.033	0.047	0.039	0.040	0.053	0.052	0.055	0.086
	1	0.9	0.064	0.057	0.091	0.035	0.036	0.051	0.043	0.048	0.067	0.057	0.064	0.092
		0.95	0.050	0.043	0.064	0.033	0.035	0.043	0.040	0.038	0.048	0.045	0.047	0.065
		1	0.072	0.068	0.099	0.055	0.052	0.067	0.060	0.060	0.072	0.073	0.070	0.096
	10	0.9	0.132	0.135	0.160	0.106	0.113	0.133	0.109	0.124	0.138	0.125	0.139	0.160
		0.95	0.144	0.167	0.191	0.143	0.154	0.169	0.148	0.160	0.173	0.154	0.169	0.192
		1	0.137	0.160	0.173	0.148	0.151	0.159	0.150	0.154	0.163	0.146	0.155	0.172
200	0.5	0.9	0.077	0.040	0.055	0.023	0.020	0.021	0.027	0.025	0.025	0.053	0.050	0.043
		0.95	0.088	0.043	0.055	0.026	0.025	0.024	0.029	0.026	0.027	0.053	0.047	0.043
		1	0.085	0.047	0.069	0.029	0.030	0.029	0.035	0.036	0.033	0.053	0.055	0.059
	1	0.9	0.088	0.049	0.063	0.024	0.025	0.023	0.030	0.032	0.031	0.057	0.061	0.057
		0.95	0.110	0.060	0.078	0.034	0.033	0.035	0.037	0.040	0.039	0.071	0.073	0.070
		1	0.108	0.060	0.079	0.038	0.038	0.036	0.044	0.044	0.045	0.072	0.072	0.073
	10	0.9	0.111	0.072	0.098	0.044	0.046	0.049	0.046	0.049	0.051	0.080	0.087	0.086
		0.95	0.124	0.065	0.101	0.042	0.043	0.045	0.044	0.050	0.050	0.088	0.091	0.087
		1	0.113	0.076	0.098	0.052	0.051	0.057	0.055	0.060	0.061	0.084	0.089	0.086

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.26: Empirical size of panel unit root statistics for 1 unknown break for $N = 60$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.058	0.090	0.067	0.047	0.081	0.061	0.055	0.087	0.065	0.059	0.087	0.065
		0.95	0.063	0.107	0.092	0.060	0.093	0.077	0.066	0.103	0.081	0.069	0.101	0.090
		1	0.070	0.111	0.088	0.065	0.096	0.079	0.074	0.103	0.091	0.067	0.104	0.086
	1	0.9	0.069	0.136	0.119	0.078	0.120	0.102	0.081	0.133	0.112	0.084	0.130	0.118
		0.95	0.063	0.125	0.102	0.081	0.107	0.090	0.087	0.115	0.097	0.075	0.117	0.101
		1	0.080	0.133	0.115	0.088	0.123	0.101	0.095	0.128	0.110	0.093	0.126	0.111
	10	0.9	0.168	0.221	0.210	0.205	0.230	0.227	0.207	0.232	0.229	0.190	0.218	0.217
		0.95	0.193	0.250	0.235	0.219	0.251	0.256	0.222	0.258	0.257	0.215	0.238	0.239
		1	0.198	0.256	0.248	0.236	0.264	0.256	0.239	0.264	0.260	0.214	0.247	0.252
100	0.5	0.9	0.057	0.047	0.076	0.035	0.037	0.047	0.035	0.038	0.049	0.052	0.055	0.074
		0.95	0.072	0.044	0.072	0.035	0.032	0.043	0.040	0.037	0.051	0.057	0.056	0.075
		1	0.072	0.051	0.091	0.038	0.036	0.056	0.041	0.041	0.062	0.062	0.061	0.091
	1	0.9	0.054	0.052	0.080	0.039	0.039	0.054	0.041	0.044	0.059	0.051	0.056	0.080
		0.95	0.066	0.061	0.087	0.040	0.045	0.059	0.045	0.051	0.062	0.059	0.068	0.090
		1	0.079	0.067	0.109	0.050	0.052	0.068	0.053	0.055	0.070	0.073	0.080	0.106
	10	0.9	0.103	0.109	0.148	0.081	0.084	0.111	0.088	0.090	0.116	0.102	0.111	0.149
		0.95	0.174	0.186	0.223	0.160	0.167	0.186	0.164	0.167	0.193	0.178	0.191	0.223
		1	0.153	0.172	0.194	0.148	0.155	0.175	0.154	0.161	0.180	0.166	0.173	0.194
200	0.5	0.9	0.117	0.060	0.086	0.037	0.037	0.036	0.041	0.039	0.041	0.069	0.076	0.072
		0.95	0.135	0.057	0.091	0.035	0.036	0.037	0.037	0.039	0.042	0.075	0.081	0.076
		1	0.138	0.053	0.081	0.026	0.026	0.027	0.031	0.030	0.031	0.067	0.071	0.071
	1	0.9	0.127	0.058	0.091	0.029	0.030	0.030	0.033	0.037	0.034	0.070	0.078	0.077
		0.95	0.143	0.059	0.098	0.035	0.036	0.036	0.036	0.038	0.040	0.082	0.084	0.078
		1	0.150	0.053	0.084	0.027	0.027	0.027	0.033	0.031	0.029	0.072	0.078	0.072
	10	0.9	0.162	0.088	0.123	0.043	0.046	0.046	0.047	0.047	0.053	0.107	0.110	0.110
		0.95	0.161	0.080	0.123	0.051	0.054	0.055	0.053	0.058	0.061	0.099	0.107	0.109
		1	0.178	0.110	0.142	0.075	0.083	0.085	0.083	0.090	0.089	0.123	0.124	0.126

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.1) through (1.6.4).

Table 1.27: Empirical size of panel unit root statistics for 2 unknown breaks for $N = 40$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.033	0.043	0.045	0.046	0.043	0.042	0.051	0.050	0.047	0.049	0.045	0.045
		0.95	0.029	0.039	0.038	0.037	0.034	0.038	0.039	0.036	0.044	0.035	0.036	0.043
		1	0.028	0.039	0.038	0.037	0.034	0.039	0.041	0.040	0.044	0.036	0.039	0.045
	1	0.9	0.035	0.042	0.038	0.040	0.038	0.041	0.043	0.042	0.045	0.043	0.041	0.044
		0.95	0.024	0.031	0.035	0.031	0.031	0.037	0.035	0.037	0.040	0.031	0.031	0.036
		1	0.020	0.034	0.036	0.031	0.031	0.030	0.034	0.034	0.034	0.032	0.034	0.038
	10	0.9	0.096	0.120	0.119	0.127	0.132	0.136	0.131	0.135	0.138	0.117	0.120	0.128
		0.95	0.132	0.164	0.158	0.165	0.167	0.180	0.170	0.174	0.187	0.160	0.163	0.165
		1	0.110	0.132	0.131	0.132	0.134	0.144	0.136	0.140	0.146	0.127	0.133	0.138
100	0.5	0.9	0.073	0.058	0.075	0.041	0.035	0.041	0.047	0.037	0.045	0.073	0.060	0.077
		0.95	0.052	0.044	0.063	0.033	0.027	0.039	0.038	0.032	0.044	0.053	0.050	0.060
		1	0.058	0.051	0.071	0.042	0.030	0.040	0.043	0.034	0.044	0.056	0.055	0.071
	1	0.9	0.060	0.050	0.069	0.043	0.034	0.046	0.052	0.038	0.049	0.057	0.053	0.070
		0.95	0.040	0.037	0.050	0.035	0.024	0.029	0.039	0.030	0.034	0.047	0.041	0.049
		1	0.054	0.046	0.061	0.042	0.036	0.047	0.054	0.041	0.052	0.052	0.051	0.061
	10	0.9	0.071	0.079	0.092	0.072	0.065	0.074	0.076	0.068	0.077	0.076	0.074	0.092
		0.95	0.080	0.089	0.102	0.092	0.084	0.091	0.094	0.088	0.091	0.089	0.093	0.104
		1	0.092	0.089	0.103	0.090	0.084	0.091	0.095	0.087	0.094	0.092	0.091	0.105
200	0.5	0.9	0.159	0.049	0.090	0.026	0.024	0.022	0.027	0.026	0.025	0.068	0.056	0.053
		0.95	0.195	0.056	0.106	0.029	0.023	0.022	0.032	0.026	0.025	0.075	0.064	0.060
		1	0.206	0.055	0.110	0.026	0.019	0.021	0.029	0.022	0.021	0.082	0.065	0.063
	1	0.9	0.181	0.045	0.090	0.019	0.016	0.016	0.023	0.016	0.018	0.065	0.048	0.050
		0.95	0.206	0.057	0.105	0.027	0.019	0.018	0.029	0.023	0.019	0.081	0.068	0.066
		1	0.210	0.056	0.096	0.032	0.027	0.027	0.034	0.030	0.032	0.078	0.068	0.064
	10	0.9	0.187	0.068	0.108	0.042	0.038	0.040	0.042	0.043	0.040	0.085	0.076	0.075
		0.95	0.193	0.066	0.100	0.036	0.028	0.028	0.039	0.030	0.032	0.080	0.069	0.068
		1	0.147	0.052	0.091	0.036	0.031	0.032	0.041	0.035	0.037	0.073	0.062	0.061

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Table 1.28: Empirical size of panel unit root statistics for 2 unknown breaks for $N = 60$

			Z^M			P^M			P_m^M			C^M		
T	σ_F^2	ρ	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}	MSB^{GLS}	MZ_α^{GLS}	MZ_t^{GLS}
50	0.5	0.9	0.036	0.041	0.041	0.033	0.028	0.038	0.040	0.038	0.038	0.039	0.038	0.041
		0.95	0.030	0.041	0.038	0.047	0.042	0.046	0.049	0.047	0.051	0.037	0.039	0.038
		1	0.034	0.039	0.041	0.041	0.040	0.044	0.042	0.041	0.044	0.037	0.039	0.042
	1	0.9	0.032	0.040	0.039	0.036	0.031	0.038	0.036	0.035	0.042	0.036	0.039	0.041
		0.95	0.042	0.054	0.056	0.045	0.046	0.051	0.048	0.053	0.053	0.051	0.056	0.058
		1	0.035	0.048	0.046	0.042	0.043	0.047	0.044	0.046	0.051	0.047	0.047	0.047
	10	0.9	0.117	0.136	0.136	0.145	0.153	0.161	0.146	0.156	0.162	0.133	0.139	0.143
		0.95	0.131	0.156	0.152	0.157	0.163	0.170	0.160	0.167	0.172	0.153	0.159	0.164
		1	0.121	0.149	0.148	0.148	0.155	0.158	0.153	0.156	0.164	0.144	0.151	0.152
100	0.5	0.9	0.079	0.067	0.097	0.060	0.042	0.057	0.075	0.048	0.063	0.085	0.076	0.096
		0.95	0.087	0.072	0.100	0.056	0.044	0.058	0.066	0.049	0.064	0.084	0.080	0.098
		1	0.091	0.075	0.110	0.065	0.052	0.065	0.066	0.058	0.068	0.091	0.086	0.109
	1	0.9	0.057	0.040	0.069	0.036	0.029	0.038	0.042	0.031	0.043	0.055	0.053	0.069
		0.95	0.062	0.054	0.074	0.046	0.034	0.042	0.048	0.040	0.046	0.067	0.058	0.074
		1	0.062	0.046	0.070	0.036	0.031	0.036	0.037	0.032	0.037	0.053	0.048	0.069
	10	0.9	0.096	0.100	0.105	0.096	0.087	0.100	0.102	0.090	0.102	0.102	0.100	0.106
		0.95	0.082	0.092	0.105	0.097	0.091	0.095	0.098	0.093	0.101	0.090	0.092	0.104
		1	0.094	0.096	0.101	0.101	0.095	0.101	0.107	0.097	0.102	0.097	0.091	0.102
200	0.5	0.9	0.230	0.059	0.133	0.031	0.025	0.025	0.032	0.027	0.028	0.096	0.073	0.070
		0.95	0.252	0.057	0.121	0.028	0.023	0.022	0.031	0.025	0.024	0.084	0.068	0.063
		1	0.301	0.086	0.169	0.042	0.034	0.033	0.045	0.037	0.036	0.114	0.098	0.090
	1	0.9	0.251	0.060	0.122	0.032	0.025	0.025	0.035	0.026	0.026	0.089	0.075	0.070
		0.95	0.279	0.070	0.139	0.033	0.029	0.031	0.036	0.032	0.032	0.100	0.083	0.078
		1	0.302	0.075	0.150	0.034	0.029	0.028	0.036	0.030	0.033	0.108	0.090	0.087
	10	0.9	0.268	0.090	0.146	0.042	0.037	0.041	0.047	0.040	0.045	0.111	0.102	0.098
		0.95	0.226	0.075	0.126	0.052	0.044	0.045	0.055	0.049	0.048	0.087	0.083	0.081
		1	0.213	0.064	0.116	0.045	0.039	0.039	0.045	0.041	0.040	0.084	0.071	0.071

Simulations are based on 1000 replications considering the 5% level of significance. The DGP is given by Equations (1.6.5) through (1.6.8).

Chapter 2

Panel Cointegration Rank Testing with Cross-Section Dependence

2.1 Introduction

Since the pioneering work of Engle and Granger (1987) the literature on cointegration grew at a rapid pace. Even though this topic has been covered extensively in time series, the analysis of cointegration in panel data is in early stages of development. Most of the initial developments assumed that time series in the panel data were either independent or that cross-section dependence could be controlled by including time effects, which defined the so-called first generation of panel cointegration tests. Unfortunately, the assumption of cross-section independence is crucial for the derivation of the limiting distributions that are obtained in these proposals. However, the use of these panel data statistics that analyze the presence of cointegration assuming cross-section independence can lead to misleading conclusions when used to study, for instance, sectors of economic activity, cities, regions or countries that are closely related. The challenge to overcome this limitation has given rise to the so-called second generation of panel cointegration tests.[1] For a more detailed literature review on panel cointegration see Breitung and Pesaran (2008).

This chapter aims to design a panel cointegration rank test statistic that accommodates for the presence of cross-section dependence through the specification of

[1] Proposals that consider the presence of cross-section dependence among the time series that define the panel data set include Gengenbach, Palm and Urbain (2006), Gengenbach, Westerlund and Urbain (2008), Bai and Carrion-i-Silvestre (2013) and Banerjee and Carrion-i-Silvestre (2013) for the single equation framework, and Groen and Kleinberger (2003) and Breitung (2005) for the vector error correction (VECM) framework.

an approximate common factor model as suggested in the seminal paper of Bai and Ng (2004). Contrary to the existing proposals, our approach allows to estimate the number of cointegrating relations in a panel data giving information about the source of the stochastic trends that might be present in the variables. We assume that the stochastic component of the time series can be decomposed in terms of an idiosyncratic component and a set of common factors. Once the observed variables have been decomposed in these elements, we can investigate the presence of cointegration among the idiosyncratic components – i.e., the individual system involving, for instance, the variables that belong to one particular country – and whether there are non-stationary common factors. Furthermore, with this setup we can distinguish between idiosyncratic stochastic trends and global stochastic trends.

The proposed framework captures several interesting features that are inherent in the economic analysis such as the presence of cross-cointegration – see Banerjee, Marcellino and Osbat (2004) – the approximation of relevant unobserved variables, the distinction between global and idiosyncratic common trends – see Gonzalo and Granger (1995) – and the modeling of the cross-section dependence driven by common factors that are present by the definition of the model under investigation – see Gengenbach, Palm and Urbain (2010).

We propose a test statistic to determine the cointegrating rank in a panel system of equations allowing for the presence of cross-section dependence across the systems of variables in the panel setup. We deal with cross-section dependence by means of approximate common factor models as proposed in Bai and Ng (2004, 2010), Gengenbach, Palm and Urbain (2006), Gengenbach, Westerlund and Urbain (2008), Bai and Carrion-i-Silvestre (2013) and Banerjee and Carrion-i-Silvestre (2013), among others. The contribution of this chapter is twofold. First, the novelty of our approach compared to the existing proposals is that it takes into account the possibility that there might be more than one cointegrating relation among the variables that define the system for each individual. At the same time, our proposal tackles the presence of cross-section dependence among the different systems in a parsimonious way through the use of common factors. The Monte Carlo simulation that has been conducted to investigate the small-sample properties of the proposed panel cointegration rank tests reveals that, in general, the panel data statistics show a good performance in terms of empirical size and power, and encompass the statistical inference that can be drawn from the univariate statistics. More interestingly, the simulations indicate that the consideration of common factors is recommended even when we conduct the analysis for only one individual system, because in some cases the presence of unattended common factors can bias the analysis towards the overestimation of the cointegrating rank. This chapter illustrates the use of the proposed techniques with two empirical applications using the variables involved in the money demand equation and the monetary exchange model. These exemplify the situation discussed above where the presence of common factors appears in a natural way. Global and idiosyncratic stochastic trends are estimated in both cases.

The remainder of this chapter is organized as follows. Section 2.2 presents the model and the assumptions that we adopt in this chapter. Section 2.3 introduces the test statistic used to determine the cointegrating rank. In Section 2.4, we discuss the way in which the individual statistics can be combined to specify a panel data cointegrating rank statistic. Section 2.5 analyzes the finite sample performance of our approach, both in a unit-by-unit framework and in a panel setup by means of Monte Carlo simulation. Section 2.6 presents two empirical applications of the proposed statistics. Finally, some concluding remarks are presented in Section 2.7. The Appendix B collects all the proofs.

2.2 The model

Let $Y_{i,t}$ be a $(k \times 1)$ vector of stochastic process, where k is assumed to be finite throughout this chapter, with the data generating process (DGP) defined as:

$$\begin{aligned} Y_{i,t} &= D_{i,t} + u_{i,t} && (2.2.1)\\ u_{i,t} &= \lambda_i F_t + e_{i,t} && (2.2.2)\\ \left(I_q - L\right) F_t &= C(L)\, w_t && (2.2.3)\\ (I_k - L)\, e_{i,t} &= G_i(L)\, \varepsilon_{i,t}, && (2.2.4) \end{aligned}$$

where $D_{i,t}$ denotes the deterministic component, which in this chapter can be either $D_{i,t} = \mu_i$ – henceforth, this specification is denoted as the intercept model – or $D_{i,t} = \mu_i + \delta_i t$ – hereafter, the linear time trend case – $t = 1, \ldots, T$ and $i = 1, \ldots, N$. The component F_t denotes a $(q \times 1)$ vector of common factors and λ_i is a $(k \times q)$ matrix of factor loadings. Finally, $e_{i,t}$ is a $(k \times 1)$ vector that collects the idiosyncratic stochastic component. Despite the operator $(I - L)$ in Equations (2.2.3) and (2.2.4), where I denotes the identity matrix of appropriate dimension (q or k, as indicated by the subscript), neither F_t or $e_{i,t}$ have to be I(1). In fact, F_t and $e_{i,t}$ can be I(0), I(1), or a combination of both, depending on the rank of $C(1)$ and $G_i(1)$. For instance, 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). The same applies for $G_i(1)$.

Let $M<\infty$ be a generic positive number, not depending on T and N. Throughout this chapter, we use $\|A\| = trace\,(A'A)^{1/2}$ to denote the Euclidean norm of matrix A. The stochastic processes that participate on the definition of the DGP are assumed to satisfy the following assumptions:

Assumption A: (i) for non-random λ_i, $\|\lambda_i\|\leq M$; for random λ_i, $E\|\lambda_i\|^4\leq M$, (ii) $\frac{1}{N}\sum_{i=1}^{N}\lambda_i'\lambda_i \xrightarrow{p} \Sigma_\Lambda$, a $(q\times q)$ positive definite matrix.

Assumption B: (i) $w_t\sim iid\,(0,\Sigma_w)$, $E\|w_t\|^4\leq M$, and (ii) $Var\,(\Delta F_t)=\sum_{j=0}^{\infty}C_j\,\Sigma_w C_j'>0$, (iii) $\sum_{j=0}^{\infty} j\left\|C_j\right\|<M$; and (iv) $C(1)$ has rank q_1, $0\leq q_1\leq q$.

Assumption C: (i) for each i, $\varepsilon_{i,t}\sim iid\,(0,\Sigma_{\varepsilon_i})$, $E\|\varepsilon_{i,t}\|^4\leq M$, (ii) $Var\,(\Delta\varepsilon_{i,t})=\sum_{j=0}^{\infty}G_{i,j}\Sigma_{\varepsilon_i}G_{i,j}'>0$, (iii) $\sum_{j=0}^{\infty} j\left\|G_{i,j}\right\|<M$; and (iv) $G_i(1)$ has rank r, $0\leq r\leq k$.

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

Assumption E: $E\|F_0\|\leq M$, and for every $i=1,\ldots,N$, $E\left\|e_{i,0}\right\|\leq M$.

Assumption A is made on the factor loadings to warrant that the factor structure is identifiable. This assumption is standard in factor analysis. Assumption B implies that the short-run variance of ΔF_t is positive definite but the long-run covariance of ΔF_t can be of reduced rank to permit linear combinations of I(1) factors to be stationary. As stated above, this permits a combination of stationary and non-stationary factors in the model. Assumption C(i) allows some weak correlation in $(I-L)\,e_{i,t}$, whereas C(ii) and C(iii) allow weak cross-section correlation. C(iv) indicates that $G_i(1)$ can be of reduced rank to permit combinations of the I(1) idiosyncratic stochastic trends to be stationary. Assumption D states that the errors $\varepsilon_{i,t}$, w_t and are λ_i are mutually independent groups across i and t. Finally, Assumption E defines the initial conditions.

The definition of the $(k\times q)$ loading matrix λ_i allows us to control the way in which the common factors affect the variables of each individual system. For example, suppose that we partition the λ_i matrix as follows:

$$\lambda_i=\begin{bmatrix}\lambda_{i,1,1} & \lambda_{i,1,2}\\ \lambda_{i,2,1} & \lambda_{i,2,2}\end{bmatrix},$$

so that we can impose restrictions on how the factors affect the elements of $Y_{i,t}$ in

(2.2.1). Thus, some of the factors can only affect one subset of the variables, say, the series that defines the cointegrating space, but not the other variables, and the other way around. Therefore, situations where $\lambda_{i,1,2} = 0$ and/or $\lambda_{i,2,1} = 0$ are covered in this setup. Note that it is possible that both $\lambda_{i,1,2} \neq 0$ and $\lambda_{i,2,1} \neq 0$, which is the general situation that is assumed henceforth.

The presence of cointegration is investigated by assessing the number of stochastic trends that are present in $u_{i,t}$ in (2.2.1). This states two different potential sources for the common trends: first, $e_{i,t}$ in (2.2.2) can capture the idiosyncratic common trends; second, F_t in (2.2.2) can control the effect of global stochastic trends. This distinction yields different qualitative interpretations that are relevant from an economic point of view. Thus, global stochastic trends imply that shocks are permanent to all variables in the panel data set, while idiosyncratic stochastic trends circumscribe permanent shocks to specific individual systems. The standard approach does not distinguish between these two sources, although the simulation experiments that we have carried out below reveal that the statistical inference can be affected if the importance of the global stochastic trends is high. In practice, the inclusion of the common factors informs us about the number of stochastic trends that affect the variables of the system. Moreover, the undesirable effects on the statistical inference are not restricted to the I(1) non-stationary case. As in Andrews (2005), the presence of I(0) stationary common factors is also relevant in our case since common shocks might cause biased and inconsistent estimation when they are left untreated. Common shocks are present in most macroeconomic data – oil shocks, international financial crises, common monetary and fiscal policies or aggregate productivity are some examples – so we should take them into account when computing the test statistics that are used to study the stochastic properties of the variables.

2.3 Cointegrating rank test in the presence of common factors

We estimate the unobservable common factors using the principal component approach suggested in Bai and Ng (2002, 2004). Let us consider the general deterministic component given by $D_{i,t} = \mu_i + \delta_i t$. Taking the first difference of the model we have:

$$\Delta Y_{i,t} = \delta_i + \lambda_i \Delta F_t + \Delta e_{i,t}. \tag{2.3.1}$$

We can define the idempotent matrix $M = I_{T-1} - \iota\left(\iota'\iota\right)^{-1}\iota'$, with ι a $(T-1)\times 1$ vector of ones. Then,

$$\begin{aligned} M\Delta Y_i &= M\Delta F\lambda_i' + M\Delta e_i. \\ y_i &= f\lambda_i' + z_i. \end{aligned}$$

Note that when the deterministic component is $D_{i,t} = \mu_i$, taking first differences removes the constant term, so that in this case we can define $M = I_{T-1}$ and the rest of our discussion applies without any modification. The common factors are extracted as the q eigenvectors corresponding to the q largest eigenvalues of the $(T-1)\times(T-1)$ matrix yy', where $y = [y_1,\ldots,y_N]$ is a $(T-1)\times kN$ matrix that is defined using the $(T-1)\times k$ matrices y_i, $i = 1,\ldots,N$. The matrix of estimated weights, $\hat{\Lambda} = \left(\hat{\lambda}_1,\ldots,\hat{\lambda}_N\right)$, is given by $\hat{\Lambda} = y'\hat{f}$. We can obtain an estimate of z_i from $\hat{z}_i = y_i - \hat{f}\hat{\lambda}_i'$, as in Bai and Ng (2004). Note that we can recover the common factors as $\hat{F}_t = \sum_{j=2}^{t}\hat{f}_j$ and the idiosyncratic component as $\hat{e}_{i,t} = \sum_{j=2}^{t}\hat{z}_{i,j}$.

Once the effects of the common factors are removed, cointegration analysis is then performed focusing on both the idiosyncratic and common factor components. This gives us further insight on the cointegration analysis, since the inference on the cointegrating rank can be distorted if common factors are not accounted for in the model.

In this chapter we propose to determine the rank with a test statistic that is based on the multivariate version of the square of the modified Sargan-Bhargava (MSB) statistic proposed in Stock (1999). Bai and Ng (2010) and Bai and Carrion-i-Silvestre (2013) are among the authors that use the MSB statistic in their papers. One important feature of the MSB statistic that motivates its use in this chapter was pointed out by Ploberger and Phillips (2004). These authors proposed a point optimal test for panel unit root in the presence of incidental trends that has some resemblance to the Sargan-Bhargava test. They show that this statistic possesses some optimality features for panel data within the class of tests that are invariant to heterogeneous trends. Another desirable property of the MSB statistic was mentioned in Bai and Ng (2010), who indicate that the statistic does not require the estimation of the autoregressive parameters, as it would be the case in other single-equation procedures available in the literature.

Without loss of generality, the exposition below uses the estimated idiosyncratic component $\left(\hat{e}_{i,t}\right)$ to describe the statistical procedure, which can be also applied to the estimated common factors $\left(\hat{F}_t\right)$.

The testing procedure builds upon the different rates of convergence of the elements on the $Q_{\hat{e}_i\hat{e}_i} = T^{-1}\hat{e}_i'\hat{e}_i$ matrix under the null hypothesis. Let us assume that the rank of the cointegrating space is $0 \leq r \leq k$. We can define the orthogonal matrix $A = [A_1 : A_2]$ with A_1 a $(k \times r)$ matrix and A_2 a $(k \times m)$ matrix, $m = k - r$, such that the first r elements of the rotated vector $e_{i,t}^A = A'e_{i,t} = \left((A_1'e_{i,t})', (A_2'e_{i,t})'\right)'$ are I(0) and the other m elements are I(1). It should to be understood that $A = A_1$ if $r = k$ and $A = A_2$ if $r = 0$. Accordingly, we partition the long-run variance matrix as

$$\Omega_{\Delta e_i^A \Delta e_i^A} = \begin{bmatrix} \Omega_{11,i} & \Omega_{12,i} \\ \Omega_{21,i} & \Omega_{22,i} \end{bmatrix}.$$

We have $\eta\left(T^{-1}Q_{\hat{e}_i\hat{e}_i}\hat{\Omega}^{-1}_{\Delta\hat{e}_i\Delta\hat{e}_i}\right) = \eta\left(T^{-1}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}}\Delta\hat{e}_i^{\hat{A}}} Q_{\hat{e}_i^{\hat{A}}\hat{e}_i^{\hat{A}}} \hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}}\Delta\hat{e}_i^{\hat{A}}}\right)$, where $\hat{A} = [\hat{A}_1 : \hat{A}_2]$ being $\hat{A}_1$ and $\hat{A}_2$ the matrices of the eigenvectors associated with the smallest r and largest m eigenvalues, respectively, of $T^{-2}\sum_{t=1}^T \hat{e}_{i,t}\hat{e}_{i,t}'$, and $\eta(\cdot)$ denotes the eigenvalues of the matrix between parentheses. The determination of the number of stochastic trends in the system relies on the sequential testing procedure that follows:

1. First, assume that the cointegrating rank is zero, i.e., set $m = k$.
2. Specify the null hypothesis that there are m stochastic trends against the alternative hypothesis that there are less than m common stochastic trends.
3. Estimate $\hat{A}_2$ as the m eigenvectors that correspond to the m largest eigenvalues of $T^{-1}Q_{\hat{e}_i\hat{e}_i}$.
4. Define the univariate MSB statistic as

$$\begin{aligned} MSB_{j,i}(m) &= \eta^{\min}\left(T^{-1}Q_{\hat{e}_i^{\hat{A}_2}\hat{e}_i^{\hat{A}_2}}\hat{\Omega}^{-1}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}\right) \\ &= \eta^{\min}\left(T^{-1}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}} Q_{\hat{e}_i^{\hat{A}_2}\hat{e}_i^{\hat{A}_2}} \hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}\right) \\ &= \hat{\eta}_{i,1}, \end{aligned} \qquad (2.3.2)$$

   where the subscript $j = \{\mu, \tau\}$ refers to the deterministic component used in the model – μ for the $D_{i,t} = \mu_i$ intercept model and τ for the $D_{i,t} = \mu_i + \delta_i t$ linear time trend case – $\eta_{i,1} < \cdots < \eta_{i,m}$ are the eigenvalues of $T^{-1}Q_{\hat{e}_i^{\hat{A}_2}\hat{e}_i^{\hat{A}_2}}\hat{\Omega}^{-1}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}$ sorted in ascending order, and $\eta^{\min}(\cdot)$ denotes the minimum eigenvalue operator.

5. Compare the value of the $MSB_{j,i}(m)$, $j=\{\mu,\tau\}$, statistic with the corresponding critical values from the left tail of the distribution – i.e., the null hypothesis that there are m stochastic trends is rejected if $MSB_{j,i}(m)$ is smaller than the corresponding critical value.

6. If the null hypothesis of m common stochastic trends is rejected, specify the null hypothesis of $m-1$ stochastic trends and return to step 2. The process continues until either the null hypothesis is not rejected or when $m=0$ is achieved.

The estimation of $\Omega_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}$ can be obtained in a parametric way from the estimation of the VECM model specification. Expressed in matrix notation, we have

$$\begin{aligned}
\Delta\hat{e}_i^{\hat{A}_2} &= \hat{e}_{i,-1}^{\hat{A}_2}\Pi_i + \Delta\hat{e}_i^{\hat{A}_2}\Gamma_{i,p_i}(L) + \varepsilon_i \\
\Delta\hat{e}_i^{\hat{A}_2}\left(I_m - \Gamma_{i,p_i}(L)\right) &= \hat{e}_{i,-1}^{\hat{A}_2}\Pi_i + \varepsilon_i \\
\Delta\hat{e}_i^{\hat{A}_2} &= \hat{e}_{i,-1}^{\hat{A}_2}\Pi_i\left(I_m - \Gamma_{i,p_i}(L)\right)^{-1} + \varepsilon_i\left(I_m - \Gamma_{i,p_i}(L)\right)^{-1},
\end{aligned}$$

where p_i denotes the number of lags of $\Delta\hat{e}_{i,i}^{\hat{A}_2}$ that are considered. Following Ng and Perron (2001), define $\hat{\Omega}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}} = \left(\left(I_m - \hat{\Gamma}_{i,p_i}(1)\right)^{-1}\right)' T^{-1}\hat{\varepsilon}_i'\hat{\varepsilon}_i\left(I_m - \hat{\Gamma}_{i,p_i}(1)\right)^{-1}$, where the lag order of the model p_i is estimated using the modified information criterion in Qu and Perron (2007) assuming that the cointegrating rank is zero – note that under the null hypothesis we assume that there are m stochastic trends in the system defined by the m variables of $\hat{e}_{i,t}^{\hat{A}_2}$.

The same procedure can be applied to the q estimated common factors, using $\hat{F}_t$ instead of $\hat{e}_{i,t}$. In this case, $A=[A_1 : A_2]$ with A_1 a $(q\times q_0)$ matrix and A_2 a $(q\times q_1)$ matrix, $q=q_0+q_1$, where q_0 and q_1 are the number of I(0) and I(1) common factors, respectively – it should be understood that $A=A_1$ if $q_0=q$ and $A=A_2$ if $q_1=q$. As above, $\hat{A}_1$ and $\hat{A}_2$ are the matrices of the eigenvectors associated with the smallest q_0 and largest q_1 eigenvalues, respectively, of $T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'$. The test statistic applied to the estimated common factors is also given by (2.3.2), which is denoted as $MSB_j^F(q_1)$, $j=\{\mu,\tau\}$.

The limiting distribution of the MSB statistics is established in the following theorem.

Theorem 2.1. *Let $Y_{i,t}$, $i=1,\ldots,N$, $t=1,\ldots,T$, be a $(k\times 1)$ vector of stochastic processes with the DGP given by (2.2.1) to (2.2.4) and satisfying Assumptions A to E. Furthermore, it is assumed that $p_i\to\infty$ and $p_i^3/\min[T,N]\to 0$ as $T\to\infty$,*

$N \to \infty$, $N/T \to 0$.

(1) Under the null hypothesis that there are m stochastic trends, the MSB statistic given in (2.3.2) applied to the idiosyncratic component converges to:

$$\begin{array}{ll} \textit{(a) For the intercept model:} & MSB_{\mu,i}(m) \Rightarrow \eta^{\min}\left(\int_0^1 W_i(s) W_i(s)' ds\right) \\ \textit{(b) For the linear trend model:} & MSB_{\tau,i}(m) \Rightarrow \eta^{\min}\left(\int_0^1 V_i(s) V_i(s)' ds\right), \end{array}$$

where $\Rightarrow$ denotes weak convergence, $W_i(s)$ is an $(m \times 1)$ vector of independent standard Brownian motions, and $V_i(s) = W_i(s) - sW_i(1)$ is an $(m \times 1)$ vector of independent Brownian bridges.

(2) Under the null hypothesis that there are q_1 stochastic trends, the MSB statistic given in (2.3.2) applied to the common factors converges to:

$$\begin{array}{ll} \textit{(a) For the intercept model:} & MSB_{\mu}^{F}(q_1) \Rightarrow \eta^{\min}\left(\int_0^1 W(s) W(s)' ds\right) \\ \textit{(b) For the linear trend model:} & MSB_{\tau}^{F}(q_1) \Rightarrow \eta^{\min}\left(\int_0^1 V(s) V(s)' ds\right), \end{array}$$

where $W(s)$ is a $(q_1 \times 1)$ vector of independent standard Brownian motions, and $V(s) = W(s) - sW(1)$ is a $(q_1 \times 1)$ vector of independent Brownian bridges.

The proof of Theorem 2.1 is given in the Appendix B. It has to be stressed that our framework treats as a special case the situation in which the time series are assumed to be cross-section independent – in this case we only need to impose $\lambda_i = 0\ \forall i$ in (2.2.2). The estimation of the parameters of the deterministic component can be done specifying the model in first differences so that $y_i = z_i$, where $y_i = M\Delta Y_i$ and $z_i = M\Delta e_i$ – as before, $M = I_{T-1}$ for the intercept and $M = I_{T-1} - \iota(\iota'\iota)^{-1}\iota'$ for the time trend deterministic specifications. Then, defining $\hat{e}_{i,t} = \sum_{j=2}^{t} y_{i,j}$ we compute the MSB statistic as described previously. Its limiting distribution is given by Theorem 2.1.

The asymptotic critical values for the $MSB_j(m)$ statistic, $j = \{\mu, \tau\}$, are reported in Table 2.1, which have been generated using 1,000 steps to approximate the Brownian motion functionals and 10,000 replications. In order to improve the performance of the statistics in empirical analyses, we have also computed critical values for different finite sample sizes.[2] For subsequent purposes, we have also

[2]These finite sample critical values are computed using the autoregressive spectral density estimator $\hat{\Omega}_{\Delta\hat{e}_i^{\hat{A}2}\Delta\hat{e}_i^{\hat{A}2}}$ defined above with the specification of the upper bound for p_i as $p_{\max} = T^{1/3}$ for $T = \{100, 200, 500\}$ – see Perron and Ng (1996). As pointed out by the referee, small sample critical values are influenced by nuisance parameters, although the specification of the autoregressive correction with $p_{\max}$ lags aims to control for the presence of correlation.

computed the mean and the variance of the statistics.

The MSB statistic presented in this chapter is consistent under the alternative hypothesis that states that there are less common stochastic trends than the ones specified under the null hypothesis. The following theorem shows the rate at which the MSB statistic diverges under the alternative hypothesis.

Theorem 2.2. *Let $Y_{i,t}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, be a $(k \times 1)$ vector of stochastic processes with the DGP given by (2.2.1) to (2.2.4) and satisfying Assumptions A to E. Under the alternative hypothesis the MSB statistics $MSB_{j,i}(m)$ and $MSB_j^F(q_1)$, $j = \{\mu, \tau\}$, are $O_p\left(T^{-1}\right)$.*

The proof is given in the Appendix B. The result in Theorem 2.2 shows that the MSB statistic converges to zero under the alternative hypothesis. Therefore, the statistic is consistent.

2.4 Panel data cointegrating rank tests

The individual MSB statistics can be pooled to define panel data statistics, which are expected to increase the performance of the statistical inference when estimating the cointegrating rank. In this section we define three different panel data statistics depending on the way in which the individual information is combined. It is important to mention that all these statistics are independent if we include the additional assumption that the idiosyncratic component is independent across the cross-section dimension. In all cases, the null hypothesis is that all N individual systems have m stochastic trends, while the alternative hypothesis is that there are $m-1$ stochastic trends:

$$\begin{cases} H_0: & m \text{ stochastic trends} \quad \forall i = 1, \ldots, N \\ H_1: & m-1 \text{ stochastic trends} \quad \forall i = 1, \ldots, N \end{cases} . \tag{2.4.1}$$

Note that for the panel data based procedure to make sense, we need to assume that all individual systems have the same number of stochastic trends. The sequential testing procedure works as described in the previous section. First, we set $m = k$ and test the null hypothesis given in (2.4.1). If this null hypothesis is rejected, then we define $m = k-1$ and proceed to test the null hypothesis in (2.4.1). The procedure continues until the null hypothesis is not rejected.

This chapter focuses on three different panel data statistics. The first panel data

MSB (PMSB) statistic is based on the standardized mean of the individual statistics:

$$PMSB_j^Z(m) = \frac{\sqrt{N}(\overline{MSB}_j(m) - E(MSB_j(m)))}{\sqrt{Var(MSB_j(m))}}, \tag{2.4.2}$$

where $\overline{MSB}_j(m) = N^{-1}\sum_{i=1}^{N} MSB_{j,i}(m)$, and $E(MSB_j(m))$ and $Var(MSB_j(m))$ are the mean and the variance of the $MSB_{j,i}(m)$ statistic computed from (2.3.2), $j = \{\mu, \tau\}$.

It is possible to define two additional panel data statistics based on the combination of the individual p-values. Maddala and Wu (1999) define the panel data Fisher-type statistic $PMSB_j^F(m) = -2\sum_{i=1}^{N} \ln \hat{\varphi}_i \sim \chi^2_{2N}$, where φ_i denotes the p-value of the $MSB_{j,i}(m)$ statistic, $j = \{\mu, \tau\}$. Finally, since the $PMSB_j^F(m)$ statistic is valid for finite N, Choi (2001) suggests to compute the following test when $N \to \infty$:

$$PMSB_j^C(m) = \frac{-2\sum_{i=1}^{N} \ln \hat{\varphi}_i - 2N}{\sqrt{4N}} \qquad j = \{\mu, \tau\}. \tag{2.4.3}$$

The limiting distribution of the panel statistics is given in the following theorem.

Theorem 2.3. *Let $Y_{i,t}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, be a $(k \times 1)$ vector of stochastic processes with the DGP given by (2.2.1) to (2.2.4) and satisfying Assumptions A to E, together with the assumption that the idiosyncratic component is independent across the cross-section dimension. Under the null hypothesis that there are m stochastic trends, with $p_i \to \infty$ and $p_i^3/\min[T,N] \to 0$ as $T \to \infty$, $N \to \infty$, $N/T \to 0$, the panel statistics given in (2.4.2) and (2.4.3) converge to:*

$$\begin{aligned} PMSB_j^Z(m) &\Rightarrow N(0,1) \\ PMSB_j^C(m) &\Rightarrow N(0,1), \ j = \{\mu, \tau\}. \end{aligned}$$

The proof is given in the Appendix B. The simulated mean and variance of the $MSB_{j,i}(m)$ statistic that are involved in the computation of the $PMSB_j^Z(m)$ statistic, $j = \{\mu, \tau\}$, under the null hypothesis are presented in Table 2.1 for different sample sizes and different number of stochastic trends. As for the p-value of the $MSB_{j,i}(m)$ statistic, $j = \{\mu, \tau\}$, we have simulated look-up tables for different values of T, which are available upon request.

2.5 Monte Carlo simulation

We now analyze the finite sample performance of the MSB panel cointegration rank tests for the two deterministic specifications considered in this chapter. The DGP is based on Toda (1995) and Saikkonen and Lütkepohl (2000) and has the following form:

$$
\begin{aligned}
Y_{i,t} &= \lambda_i F_t + e_{i,t} && (2.5.1)\\
\begin{pmatrix} e_{1,i,t} \\ e_{2,i,t} \end{pmatrix} &= \begin{pmatrix} \psi_i & 0 \\ 0 & I_{k-r} \end{pmatrix} \begin{pmatrix} e_{1,i,t-1} \\ e_{2,i,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,i,t} \\ \varepsilon_{2,i,t} \end{pmatrix} && (2.5.2)\\
F_t &= \rho F_{t-1} + \sigma_F w_t, && (2.5.3)
\end{aligned}
$$

where $\varepsilon_{i,t} \sim iid\ N(0, I_k)$, $i = 1, \ldots, N$, $t = 1, \ldots, T$ and $j = 1, \ldots, q$. The system is defined by $k = 3$ variables, $Y_{i,t} = (y_{1,i,t}, y_{2,i,t}, y_{3,i,t})'$, with the deterministic specifications given by $D_{i,t} = \mu_i + \delta_i t$, with $\mu_i \sim U[-1, 1]$ and $\delta_i \sim U[-0.5, 0.5]$, where U denotes the uniform distribution. The idiosyncratic cointegrating rank is investigated using $\psi_i = aI_r$ with $a = \{0.5, 0.8, 1\}$. When there is not cointegration among any of the time series, Equation (2.5.2) reduces to $e_{i,t} = e_{i,t-1} + \varepsilon_{i,t}$. On the other hand, when all time series are I(0), Equation (2.5.2) reduces to $e_{i,t} = \psi_i e_{i,t-1} + \varepsilon_{i,t}$.

We distinguish three different configurations for the common factor component:

- Setup 1: one common factor ($q = 1$) with $\lambda_i = (0, 0, c_{1,i})'$, $c_{1,i} \sim U[1, 2]$. Note that in this case the common factor only affects the variable y_3 of each individual system.

- Setup 2: one common factor ($q = 1$) that affects two variables of the system (y_2 and y_3) through the specification of $\lambda_i = (0, c_{1,i}, c_{2,i})'$, $c_{1,i} \sim U[1, 2]$, $c_{2,i} \sim U[1, 2]$.

- Setup 3: two common factors ($q = 2$), each of them affecting one variable of the system, with

$$
\lambda_i = \begin{bmatrix} 0 & 0 \\ c_{2,1,i} & 0 \\ 0 & c_{3,2,i} \end{bmatrix}; \quad c_{2,1,i} \sim U[1, 2], c_{3,2,i} \sim U[1, 2].
$$

  Note that in this case the first factor only affects y_2 whereas the second factor only affects y_3.

As for the parameters affecting the definition of the common factors, we have conducted simulations for all the combinations of $\rho = \{0.9, 0.95, 1\}$ and $\sigma_F^2 = \{0.5, 1, 10\}$ with $w_t \sim N(0,1)$. The number of common factors is estimated using the panel Bayesian information criterion (BIC) in Bai and Ng (2002) with $q_{\max} = 6$ as the maximum number of common factors. It is worth mentioning that panel BIC always detected the true number of common factors.

The empirical size and power of the statistics are obtained using $1,000$ replications for all different combinations of $N = \{1, 10, 20, 40\}$ and $T = \{100, 200, 500\}$. The nominal size is set at the 5% level of significance. For conciseness, we report only the results for $N = 1$ and for $N = 20$. The results for $N = \{10, 40\}$ are qualitatively similar to those for $N = 20$ so they are not shown to save space. Furthermore, since the performance of the different panel data statistics is similar, we only present the results for the $PMSB^Z$ test, although the complete set of results is available upon request. Finally, the simulations were carried out in GAUSS using the COINT 2.0 library.

2.5.1 Ignoring the cross-section dependence

To motivate the importance of the presence of common factors, we first investigate the MSB statistic when the common factors are not considered. Since the purpose of this exercise is to illustrate the potential pitfalls that can affect the cointegration analysis in this situation, we only use here the first setup, i.e., the one that assumes that there is only one common factor that affects one variable of the individual systems. Finally and due to space constraint, we report here a summarized version of the results of the simulation experiments that we have conducted.

Unit-by-unit analysis

Tables 2.2 and 2.3 present the results for the deterministic specification given by a constant term and a time trend, respectively. As it can be seen, the presence of one I(0) common factor does not affect the inference when we focus on the individual systems ($N = 1$), unless the importance of the common factor is large ($\sigma_F^2 = 10$). Thus, when $\sigma_F^2 = 10$ the statistical procedure tends to detect fewer stochastic trends than there exist when the true $m > 0$. Furthermore, for $\sigma_F^2 > 0.5$ we observe that, as the sample size increases, the probability of selecting the correct number of stochastic trends decreases for $m > 0$. This is an undesirable effect, since we should expect a better performance of the statistical analysis as T increases.

These features are similar for the two deterministic functions, although the model that includes the time trend tends to show a better performance.

When the common factor is I(1) and $m = 0$, the probability of selecting the correct number of stochastic trends falls dramatically: at best, it is equal to 0.242 for $T = 100$ and $\sigma_F^2 = 0.5$ for the intercept model, and equal to 0.368 for $T = 200$ and $\sigma_F^2 = 0.5$ for the time trend model. This is something to be expected, since the analysis is not accounting for the stochastic trend that introduces the common factor. When the true $m = 1$, the probability of selecting the correct number of stochastic trends is nearly 0.95, which is also to be expected since the stochastic trend of the individual system is the one affected by the I(1) common factor. An interesting result is obtained when $m = 3$ and $\sigma_F^2 = 10$, for which the probabilities of correct selection are 0.771 ($T = 100$), 0.793 ($T = 200$) and 0.808 ($T = 500$) for the intercept model. Intuitively, we should expect that the presence of an I(1) common factor should not cause any distortion in this case, where the individual system has three stochastic trends. Notwithstanding, these values are below 0.95, showing that leaving the common factors untreated causes undesirable effects. In contrast, the probabilities of correct selection are 0.952 ($T = 100$), 0.967 ($T = 200$) and 0.953 ($T = 500$) once the common factor has been accounted for in the model.

Tables 2.2 and 2.3 report the performance of the MSB statistic when the common factors are considered. Now, the probability of selecting the correct number of stochastic trends is close to the nominal size in almost all cases, clearly outperforming the inference that is obtained when the common factors are disregarded.

Panel data analysis

The performance of the statistical inference worsens when we use the panel data $PMSB^Z$ statistic – see Tables 2.4 and 2.5. The lack of considering the cross-section dependence introduced by the common factor, regardless of whether it is I(0) or I(1), introduces serious biases towards concluding that $\hat{m} < m$. This feature is clearly evidenced for the cases where $\sigma_F^2 > 0.5$. As expected, when $m = 0$ the presence of an I(0) common factor does not cause any problems.

Another interesting situation is when the true number of stochastic trends is $m = 0$ and the common factor is I(1), since in this case the probability of selecting $m = 0$ is around 0.5 and the probability of selecting $m = 1$ is around 0.5. This situation makes sense given that the statistical procedure is detecting the common factor, but only for half of the times. The same is also found when the true number of stochastic trends is $m = 1$. Thus, we can conclude that either $m = 0$ or $m = 1$ with the same probability irrespective of whether the true m is 0 or 1.

When the common factors are allowed for, the $PMSB^Z$ panel statistic shows probabilities of correct selection that are close to 95%, which reveals the damage caused by ignoring the common factors, even if they are I(0). The presence of unattended common factors can lead to distorted conclusions, since there is a tendency to indicate that there are more cointegrating relations than there exist in the system ($\hat{m} < m$). This evidences the importance of accounting for cross-section dependence both in unit-by-unit and panel data analyses.

2.5.2 Considering the cross-section dependence

Unit-by-unit analysis

The simulations reported in Tables 2.2 and 2.3 for $a = 0.5$ and $N = 1$ reveal some important features. First, we can see that the results do not depend on the stochastic properties of the common factor, provided that the performance of the MSB statistics is similar regardless of the order of integration of the common factor, and for all values of σ_F^2. Second, the performance of the MSB statistic depends on how close to one is the autoregressive parameter a. Thus, for $a = 0.5$ the MSB statistic selects the correct number of stochastic trends in most cases. As expected, the behavior of the procedure worsens for $a = 0.8$, although it tends to select the correct number of stochastic trends as T increases. The results for $a = 0.8$ are available upon request. Note that this picture is in sharp contrast with the previous one when the common factors are not considered.

Panel data analysis

Tables 2.4 to 2.7 collect the simulation results for the $PMSB^Z$ statistic when $a = 0.5$ and $N = 20$. The performance is similar for the two deterministic specifications, regardless of the setup that is used and the stochastic properties of the common factors.[3] As it can be seen, the probability of selecting the correct number of idiosyncratic stochastic trends is close to 0.95, which indicates the good properties that show the $PMSB^Z$ statistic. As expected, results not reported here show that the performance is slightly worse for $a = 0.8$. This is more evident for the $PMSB_\tau^F$ and

[3] Moon, Perron and Phillips (2007) show that panel unit root test suffer from power loss when a deterministic trend is considered. However, in our simulation setup we have not observed such power reduction problem.

$PMSB_{\tau}^{C}$ statistics when $T = 100$. However, as T increases, all panel data statistics select the correct number of stochastic trends.

The behavior of the MSB statistic applied to the estimated common factors ($MSB_j^F(q_1)$, $j = \mu, \tau$) is analyzed in Tables 2.8 and 2.9 for the intercept and time trend models, respectively. At first sight, the pattern of the results behaves as expected, i.e., the performance of the MSB statistic improves as T and σ_F^2 increase. However, for small T and σ_F^2 the key feature that defines the performance of the statistic is whether the common factors are I(0) or I(1).

When the common factors are I(1), the probability of selecting the correct number of common factors is close to 0.95 in all cases. When the common factors are I(0), the MSB statistic tends to overestimate the number of common stochastic trends when $\rho = 0.95$ and $T = 100$ – this is especially evident for setup 3. These results resemble others reported previously in the literature using alternative statistics – see, for instance, Stock and Watson (1988) and Bai and Ng (2004) – and evidence the fact that we are dealing within a time series framework, i.e., $MSB_j^F(q_1)$, $j = \mu, \tau$, is not a panel data statistic. Finally, note that the results are qualitatively similar for the two deterministic specifications.

2.6 Empirical illustrations

In this section we illustrate the use of the panel data cointegrating rank test that has been proposed in the chapter. We focus on two empirical applications that have been widely analyzed in monetary economics, and constitute two examples of the situations discussed in the introduction where the use of common factors is justified. The increasing economic integration experienced by some developed economies and the international capital mobility can be thought as main factors to induce cross-section dependence among variables such as interest rates, exchange rates, GDP and money aggregates. These variables are the main ingredients of the models that we investigate below.

2.6.1 The money demand model

The first illustration focuses on the long-run money demand function, a topic that has been studied previously in Stock and Watson (1993), Hoffman, Rasche and Tieslau (1995), and Mark and Sul (2003), among others. We employ the data set used in Mark and Sul (2003), which consists of annual observations that covers the period that goes from 1957 until 1996 for nineteen countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Switzerland, United Kingdom and the United States. See Mark and Sul (2003) for further details about the data definitions and sources.

The three-dimensional vector $Y_{i,t}$ for the i-th country at time t has the form:

$$Y_{i,t} = (m_{i,t}, gdp_{i,t}, R_{i,t})', \tag{2.6.1}$$

$i = 1, \ldots, 19$, $t = 1957, , \ldots, 1996$, where $m_{i,t}$ is the logarithm of the real money aggregate $M1$, $gdp_{i,t}$ is the logarithm of the real GDP and $R_{i,t}$ is the nominal short-term interest rate.

We start our analysis by checking whether the cross-section dependence exists among the series of the panel. To this end, we apply the panel CD statistic of Pesaran (2004, 2013) to the panels generated by each variable and to the whole panel defined by all variables taken together.[4] Panel A of Table 2.10 presents the values of the test, showing that the null hypothesis of cross-section independence is strongly rejected for all panel data sets. It is therefore natural to conclude that the time series in our panel sets are cross-section dependent.

In order to investigate the stochastic properties of the panel data sets of each variable we compute the panel data unit root test in Bai and Ng (2004), Moon and Perron (2004) and Pesaran (2007). These three proposals capture the presence of cross-section dependence using common factor models, although the underlying assumptions of the model specification are different. Thus and as pointed out in Bai and Ng (2010), the approaches in Moon and Perron (2004) and Pesaran (2007) focus on the idiosyncratic component, with the additional implicit assumption that both the idiosyncratic and common factor components have the same order of integration. However, the proposal in Bai and Ng (2004) tests the unit root hypothesis both on the idiosyncratic and common factor components separately, so that each component might have different orders of integration. Consequently, the results drawn from the Bai-Ng procedure are more informative since they allow us to identify the potential source of non-stationarity and, thus, characterize the stochastic properties of the observable data.

[4]The order of the autoregressive model is selected using the t-sig criterion in Ng and Perron (1995), with the maximum number of lags defined by $p_{\max} = \left[12(T/100)^{1/4}\right]$. For money supply and GDP, the regressions include a time trend, while for interest rates they include an intercept. This specification is also used below when computing the unit root test statistics.

Let us first focus on the results of the Bai and Ng (2004) statistics reported in Panel A of Table 2.10. The panel data ADF statistic of the idiosyncratic component (ADF_e) indicates that the null hypothesis of panel unit root cannot be rejected for the $m_{i,t}$ and $gdp_{i,t}$ variables, while it is strongly rejected for the $R_{i,t}$ variable at the 5% level of significance. The procedure has estimated two common factors for the $gdp_{i,t}$, one common factor for $m_{i,t}$ and six common factors for $R_{i,t}$.[5] In all cases, the common factors are characterized as I(1) stochastic processes, which implies that $Y_{i,t} \sim I(1)$.

Pesaran's (2007) CIPS statistic gives similar conclusions, since the panel unit root hypothesis (for the idiosyncratic component) is not rejected for $gdp_{i,t}$ and $m_{i,t}$, while the conclusion for the $R_{i,t}$ variable depends on the order of the autoregressive correction – the null hypothesis of panel unit root is not rejected for $p > 2$, where p denotes the order of the autoregressive correction. Finally, the t_a and t_b statistics from Moon and Perron (2004) indicate that (the idiosyncratic component of) $m_{i,t}$ and $R_{i,t}$ are I(0), whereas $gdp_{i,t}$ is I(1). As mentioned above, the Bai and Ng (2004) procedure is less restrictive and more informative provided that it investigates the stochastic properties of both the idiosyncratic and common factor components without imposing the constraint that the order of integration of the components has to be the same. If we rely on these results, we conclude that $Y_{i,t} \sim I(1)$.

Given that the panel data set of the different variables has been characterized as I(1), we then proceed to analyze the cointegrating rank. Following the previous analyses in the literature mentioned above, we estimate the time trend model specification. The computation of the panel data BIC selects two common factors. For Denmark and Iceland, the MSB statistic for the individual systems of idiosyncratic terms reported in Panel A of Table 2.11 indicates that there are two idiosyncratic stochastic trends, i.e., the cointegrating rank is one.[6] For the rest of the countries, the MSB statistic detects three idiosyncratic stochastic trends, i.e., the cointegrating rank is zero. One possible explanation for the absence of cointegration at the individual level might the low power of the individual MSB statistic. However, at the panel level the results from the cointegration tests are unanimous in favor of cointegration. The panel data $PMSB^{Z}_{\tau}$, $PMSB^{F}_{\tau}$, and $PMSB^{C}_{\tau}$ statistics in Panel A of Table 2.11 indicate the existence of two idiosyncratic stochastic trends, i.e., the cointegration rank is one. Finally, the MSB statistic applied to the two estimated

[5]In the two empirical applications of this chapter, the number of common factors is selected using the panel BIC information criterion as in Bai and Ng (2002) specifying a maximum of six factors.

[6]In the two empirical applications of this chapter, the lag order of the model p_i is estimated using the modified information criterion in Qu and Perron (2007) with a maximum of six lags.

common factors classifies all of them as I(1) stochastic processes. This conclusion is also found when using the MQ tests in Bai and Ng (2004).

In all, we have detected the existence of two global stochastic trends that affect the variables in our panel data system. These global stochastic trends represent cross-cointegration relations that can exist among the same macroeconomic variable for the different countries. Furthermore, all panel data statistics that have been computed indicate that the idiosyncratic component of these variables, which captures the shocks that affect the individual economies, is affected by two common trends.

2.6.2 The monetary exchange model

The second application examines the monetary exchange rate model during the post-Bretton Woods era. We use the Mark and Sul (2001) dataset, which has also been analyzed in Rapach and Wohar (2004) and Gengenbach, Westerlund and Urbain (2008). The data consists of quarterly observations for the logarithms of the real GDP, money supply and nominal exchange rate from the first quarter of 1973 up to the first quarter of 1997 for nineteen countries: Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Italy, Japan, Korea, the Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom and the United States. See Mark and Sul (2001) for further details on the database.

The variables used in the analysis are the relative money supply ($m^*_{i,t} = m_{i,t} - m_{US,t}$), the relative real output or GDP ($gdp^*_{i,t} = gdp_{i,t} - gdp_{US,t}$) and the nominal exchange rate ($exch_{i,t}$) for country i at time t, where the United States has been used as the benchmark. As it can be seen, the definition of the variables of this model implies the existence of common factors by construction, since money and output are expressed as deviations from the corresponding US variables – i.e., the factors – and the exchange rate indicates the price of the corresponding currency in terms of the US dollar. Therefore, at this stage of the analysis we should expect to detect at least two common factors – note that this exemplifies one of the situations discussed in the introduction. As expected, the CD statistics presented in Panel B of Table 2.10 indicate that the null hypothesis of cross-section independence is strongly rejected in all cases.

The stochastic properties of the three panels are investigated using the panel data unit root test as in Bai and Ng (2004), Moon and Perron (2004) and Pesaran (2007). The first statistic (ADF_e) indicates that the null hypothesis of panel unit root cannot be rejected for the $gdp^*_{i,t}$ and $exch_{i,t}$ variables, but it is strongly rejected for the $m^*_{i,t}$ variable at the 5% level of significance. As for the common factors component, Panel B of Table 2.10 shows that there have been estimated six common factors for $m^*_{i,t}$ and $exch_{i,t}$, and one common factor for $gdp^*_{i,t}$. Pesaran's (2007) CIPS statistic indicates that $exch_{i,t}$ is I(1) regardless of the order of the autoregressive correction (p), but it gives mixed results for $gdp^*_{i,t}$ and $m^*_{i,t}$. More precisely, $gdp^*_{i,t}$ and $m^*_{i,t}$ are I(1) for $p = \{3,4,5\}$ and $p = \{0,3,4,5\}$, respectively. Finally, the t_a and t_b statistics from Moon and Perron (2004) indicate that the idiosyncratic components of all three variables are I(1).

Since there is overwhelming evidence in favor of non-stationarity, we proceed with the cointegration analysis using the specification that considers a time trend. If we look at the individual systems, Panel B of Table 2.11 reveals that the cointegrating rank is one for Norway and Spain. For the rest of the countries, the individual MSB statistic indicates the absence of any cointegrating relation between the variables of the monetary exchange model. The panel data $PMSB^Z_\tau$, $PMSB^F_\tau$ and $PMSB^C_\tau$ tests lead to the same conclusion, indicating that the cointegration rank is zero. Panel B of Table 2.11 also presents the results from the analysis of the common factors. As it can be seen, the computation of the panel data BIC selects three common factors, which are characterized by the MQ tests in Bai and Ng (2004) and the MSB statistic as I(1) non-stationary common factors. Overall, the panel data statistics that indicate that the idiosyncratic component of these variables, is affected by three common trends.

2.7 Conclusion

In this chapter we propose a new test statistic to estimate the cointegrating rank both in a unit-by-unit analysis and in a panel data framework. Our proposal covers the cross-section dependence through the specification of approximate common factor models, which is a relevant situation from both theoretical and empirical point of views. This setup allows us to cover strong cross-section dependence cases, i.e., cases where the time series of one individual system are cointegrated with times series of other individual system (cross-cointegration), as well as cases where the factors appear by construction due to model specification.

The performance of the proposed tests is investigated with Monte Carlo simulations. In general, the panel data based MSB statistic provides better estimation of the number of stochastic trends present in each individual system than the univariate one. More interestingly, the simulations reveal that the existence of common factors can lead to misleading conclusions even if the analysis is carried out at a unit-by-unit basis. This is relevant from an empirical point of view considering that, in most cases, the cointegration analysis is conducted by focusing on one country whose economic system is related to that of other countries or ruled by international organizations such as in the case of the European Union. Therefore, the theoretical proposal presented in this chapter has also a significant empirical contribution. We have illustrated the application of the techniques to two popular empirical models. The analysis detected two stochastic trends for the case of money demand and three stochastic trends for the case of monetary exchange model.

Table 2.1: Critical values, mean and variance of the MSB_μ and MSB_τ statistics

CRITICAL VALUES FOR INDIVIDUAL TESTING

MSB_μ statistic

	$T = 100$		$T = 200$		$T = 500$		$T = 1,000$	
k	5%	10%	5%	10%	5%	10%	5%	10%
1	0.0629	0.0853	0.0610	0.0841	0.0593	0.0810	0.0576	0.0769
2	0.0318	0.0377	0.0298	0.0357	0.0280	0.0336	0.0271	0.0328
3	0.0227	0.0257	0.0203	0.0234	0.0188	0.0218	0.0181	0.0210
4	0.0186	0.0203	0.0160	0.0179	0.0147	0.0165	0.0136	0.0152
5	0.0158	0.0170	0.0133	0.0145	0.0120	0.0132	0.0111	0.0123
6	0.0141	0.0150	0.0116	0.0124	0.0102	0.0111	0.0093	0.0102

MSB_τ statistic

	$T = 100$		$T = 200$		$T = 500$		$T = 1,000$	
k	5%	10%	5%	10%	5%	10%	5%	10%
1	0.0420	0.0520	0.0397	0.0496	0.0389	0.0480	0.0364	0.0456
2	0.0268	0.0307	0.0243	0.0282	0.0228	0.0267	0.0218	0.0255
3	0.0206	0.0226	0.0178	0.0200	0.0166	0.0189	0.0155	0.0178
4	0.0172	0.0186	0.0146	0.0162	0.0131	0.0145	0.0122	0.0136
5	0.0150	0.0160	0.0125	0.0135	0.0109	0.0120	0.0100	0.0111
6	0.0135	0.0143	0.0109	0.0117	0.0095	0.0103	0.0086	0.0093

MEAN AND VARIANCE

MSB_μ statistic

	$T = 100$		$T = 200$		$T = 500$		$T = 1,000$	
k	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
1	0.50363	0.29974	0.50454	0.32442	0.49960	0.31146	0.50445	0.34863
2	0.09293	0.00387	0.09049	0.00386	0.08806	0.00378	0.08580	0.00364
3	0.04653	0.00042	0.04356	0.00042	0.04191	0.00040	0.04048	0.00039
4	0.03107	0.00010	0.02864	0.00011	0.02708	0.00010	0.02578	0.00009
5	0.02379	0.00004	0.02115	0.00003	0.01995	0.00004	0.01884	0.00003
6	0.01970	0.00002	0.01710	0.00002	0.01571	0.00002	0.01475	0.00002

MSB_τ statistic

	$T = 100$		$T = 200$		$T = 500$		$T = 1,000$	
k	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
1	0.17830	0.02079	0.17637	0.02195	0.16825	0.02122	0.16622	0.02267
2	0.06172	0.00108	0.05943	0.00109	0.05705	0.00106	0.05539	0.00095
3	0.03709	0.00019	0.03431	0.00018	0.03281	0.00018	0.03132	0.00017
4	0.02696	0.00006	0.02454	0.00006	0.02270	0.00005	0.02172	0.00005
5	0.02161	0.00002	0.01900	0.00002	0.01744	0.00002	0.01651	0.00002
6	0.01831	0.00001	0.01570	0.00001	0.01427	0.00001	0.01323	0.00001

k denotes the number of stochastic trends under the null hypothesis. Simulations are based on 10,000 replications, with DGP given by (2.2.1) to (2.2.4). $p_{max} = T^{1/3}$ for $T < 1,000$, and 0 otherwise

Table 2.2: Probability of selecting the correct number of stochastic trends. Intercept model, setup 1 with $a = 0.5$, unit-by-unit analysis, MSB_{μ} statistic

			Factors are not considered				Factors are considered			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	0.972	0.910	0.932	0.950	1.000	0.947	0.953	0.972
		0.95	0.805	0.940	0.951	0.961	0.990	0.963	0.962	0.974
		1	0.242	0.952	0.957	0.959	0.750	0.898	0.935	0.960
	1	0.90	0.946	0.881	0.927	0.933	0.999	0.956	0.956	0.971
		0.95	0.663	0.921	0.929	0.947	0.986	0.952	0.945	0.962
		1	0.186	0.953	0.943	0.955	0.786	0.919	0.932	0.958
	10	0.90	0.747	0.638	0.708	0.718	1.000	0.952	0.952	0.957
		0.95	0.314	0.826	0.803	0.760	0.995	0.959	0.958	0.950
		1	0.060	0.959	0.837	0.771	0.753	0.899	0.933	0.952
200	0.5	0.90	0.999	0.902	0.915	0.925	1.000	0.955	0.951	0.947
		0.95	0.957	0.923	0.927	0.928	1.000	0.954	0.941	0.947
		1	0.217	0.954	0.951	0.956	0.755	0.942	0.950	0.950
	1	0.90	0.999	0.856	0.871	0.909	1.000	0.959	0.940	0.948
		0.95	0.950	0.898	0.932	0.940	1.000	0.963	0.953	0.958
		1	0.167	0.950	0.938	0.940	0.783	0.953	0.947	0.951
	10	0.90	0.998	0.421	0.452	0.501	1.000	0.955	0.947	0.942
		0.95	0.801	0.616	0.703	0.712	1.000	0.952	0.945	0.957
		1	0.055	0.945	0.846	0.793	0.799	0.949	0.954	0.967
500	0.5	0.90	1.000	0.884	0.911	0.930	1.000	0.953	0.956	0.960
		0.95	1.000	0.887	0.927	0.950	1.000	0.938	0.956	0.969
		1	0.119	0.950	0.944	0.946	0.818	0.954	0.947	0.955
	1	0.90	1.000	0.802	0.829	0.860	1.000	0.952	0.957	0.956
		0.95	0.999	0.833	0.877	0.885	1.000	0.951	0.949	0.950
		1	0.110	0.952	0.938	0.941	0.830	0.942	0.950	0.964
	10	0.90	1.000	0.321	0.224	0.173	1.000	0.940	0.950	0.952
		0.95	1.000	0.379	0.358	0.398	1.000	0.957	0.961	0.948
		1	0.072	0.948	0.855	0.808	0.816	0.954	0.948	0.953

Table 2.3: Probability of selecting the correct number of stochastic trends. Time trend model, setup 1 with $a = 0.5$, unit-by-unit analysis, MSB_τ statistic

			Factors are not considered				Factors are considered			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	0.746	0.876	0.910	0.958	0.822	0.889	0.902	0.953
		0.95	0.548	0.899	0.904	0.960	0.827	0.879	0.896	0.970
		1	0.352	0.896	0.914	0.944	0.766	0.849	0.897	0.953
	1	0.90	0.615	0.857	0.898	0.929	0.819	0.882	0.909	0.959
		0.95	0.365	0.870	0.899	0.940	0.809	0.865	0.898	0.945
		1	0.200	0.897	0.911	0.951	0.778	0.863	0.879	0.960
	10	0.90	0.327	0.735	0.712	0.699	0.834	0.885	0.902	0.962
		0.95	0.108	0.871	0.776	0.741	0.811	0.878	0.898	0.947
		1	0.057	0.904	0.811	0.761	0.745	0.843	0.869	0.963
200	0.5	0.90	0.941	0.897	0.927	0.940	0.893	0.952	0.941	0.954
		0.95	0.841	0.939	0.932	0.942	0.899	0.956	0.939	0.946
		1	0.368	0.962	0.961	0.951	0.826	0.945	0.947	0.948
	1	0.90	0.947	0.833	0.883	0.914	0.894	0.947	0.948	0.959
		0.95	0.738	0.900	0.938	0.942	0.888	0.955	0.947	0.961
		1	0.254	0.961	0.939	0.945	0.837	0.952	0.952	0.951
	10	0.90	0.824	0.480	0.563	0.589	0.900	0.949	0.953	0.953
		0.95	0.358	0.801	0.762	0.746	0.888	0.951	0.951	0.965
		1	0.067	0.947	0.838	0.777	0.856	0.943	0.954	0.963
500	0.5	0.90	0.979	0.860	0.892	0.911	0.896	0.948	0.963	0.962
		0.95	0.969	0.894	0.915	0.928	0.880	0.948	0.953	0.959
		1	0.297	0.942	0.950	0.946	0.816	0.948	0.954	0.945
	1	0.90	0.998	0.756	0.810	0.846	0.898	0.947	0.950	0.938
		0.95	0.982	0.805	0.866	0.898	0.899	0.945	0.951	0.949
		1	0.241	0.944	0.933	0.952	0.839	0.953	0.953	0.960
	10	0.90	0.995	0.162	0.184	0.193	0.899	0.946	0.946	0.959
		0.95	0.930	0.337	0.428	0.480	0.901	0.937	0.956	0.944
		1	0.063	0.944	0.846	0.777	0.830	0.959	0.948	0.953

Table 2.4: Probability of selecting the correct number of stochastic trends. Intercept model, setup 1 with $a = 0.5$, $N = 20$, panel data analysis, $PMSB_{\mu}^{Z}$ statistic

			Factors are not considered				Factors are considered			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	1.000	0.572	0.635	0.745	1.000	0.971	0.983	0.976
		0.95	0.990	0.580	0.779	0.841	1.000	0.986	0.974	0.966
		1	0.528	0.803	0.905	0.915	0.954	0.981	0.980	0.972
	1	0.90	1.000	0.228	0.317	0.471	1.000	0.983	0.975	0.969
		0.95	0.981	0.288	0.571	0.675	1.000	0.983	0.979	0.969
		1	0.520	0.683	0.817	0.801	0.948	0.983	0.977	0.978
	10	0.90	1.000	0.000	0.023	0.020	1.000	0.980	0.979	0.974
		0.95	0.974	0.034	0.141	0.089	1.000	0.986	0.983	0.969
		1	0.468	0.545	0.309	0.095	0.955	0.978	0.978	0.972
200	0.5	0.90	1.000	0.523	0.411	0.518	1.000	0.974	0.963	0.973
		0.95	1.000	0.498	0.609	0.730	1.000	0.976	0.969	0.972
		1	0.492	0.817	0.888	0.919	0.982	0.981	0.965	0.976
	1	0.90	1.000	0.165	0.074	0.126	1.000	0.976	0.971	0.976
		0.95	1.000	0.103	0.293	0.458	1.000	0.981	0.966	0.974
		1	0.485	0.687	0.800	0.817	0.985	0.982	0.971	0.974
	10	0.90	1.000	0.000	0.000	0.001	1.000	0.974	0.967	0.970
		0.95	0.999	0.001	0.019	0.021	1.000	0.968	0.968	0.977
		1	0.505	0.509	0.295	0.115	0.991	0.979	0.967	0.973
500	0.5	0.90	1.000	0.524	0.214	0.215	1.000	0.975	0.967	0.966
		0.95	1.000	0.363	0.357	0.429	1.000	0.973	0.969	0.975
		1	0.516	0.794	0.880	0.899	1.000	0.979	0.968	0.963
	1	0.90	1.000	0.140	0.025	0.013	1.000	0.973	0.960	0.962
		0.95	1.000	0.069	0.046	0.082	1.000	0.973	0.962	0.964
		1	0.494	0.680	0.818	0.814	1.000	0.962	0.959	0.963
	10	0.90	1.000	0.000	0.000	0.000	1.000	0.971	0.968	0.966
		0.95	1.000	0.000	0.000	0.000	1.000	0.972	0.969	0.958
		1	0.510	0.499	0.303	0.095	1.000	0.977	0.983	0.965

Table 2.5: Probability of selecting the correct number of stochastic trends. Time trend model, setup 1 with $a = 0.5$, $N = 20$, panel data analysis, $PMSB^Z_\tau$ statistic

			Factors are not considered				Factors are considered			
T	σ^2_F	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	0.973	0.492	0.731	0.818	1.000	0.986	0.986	0.986
		0.95	0.885	0.652	0.865	0.886	0.998	0.984	0.980	0.982
		1	0.650	0.820	0.923	0.930	0.973	0.980	0.974	0.969
	1	0.90	0.957	0.250	0.482	0.559	1.000	0.985	0.978	0.977
		0.95	0.829	0.472	0.707	0.756	0.998	0.978	0.980	0.979
		1	0.600	0.678	0.817	0.815	0.972	0.979	0.979	0.977
	10	0.90	0.925	0.092	0.062	0.017	1.000	0.983	0.982	0.981
		0.95	0.772	0.274	0.221	0.068	0.999	0.977	0.985	0.981
		1	0.493	0.534	0.330	0.083	0.965	0.977	0.985	0.971
200	0.5	0.90	0.999	0.246	0.439	0.576	1.000	0.973	0.971	0.974
		0.95	0.968	0.392	0.657	0.772	1.000	0.972	0.975	0.969
		1	0.605	0.773	0.883	0.908	0.991	0.961	0.969	0.969
	1	0.90	0.996	0.045	0.095	0.196	1.000	0.966	0.972	0.975
		0.95	0.978	0.177	0.425	0.578	1.000	0.964	0.973	0.973
		1	0.599	0.673	0.783	0.808	0.992	0.963	0.973	0.976
	10	0.90	0.999	0.003	0.003	0.002	1.000	0.962	0.973	0.971
		0.95	0.961	0.050	0.059	0.032	1.000	0.967	0.959	0.964
		1	0.504	0.530	0.290	0.081	0.996	0.961	0.968	0.977
500	0.5	0.90	1.000	0.160	0.140	0.192	1.000	0.973	0.975	0.965
		0.95	0.999	0.192	0.343	0.455	1.000	0.978	0.979	0.974
		1	0.584	0.784	0.890	0.905	1.000	0.974	0.978	0.966
	1	0.90	1.000	0.014	0.007	0.007	1.000	0.966	0.983	0.965
		0.95	1.000	0.021	0.067	0.105	1.000	0.976	0.974	0.970
		1	0.561	0.674	0.781	0.790	1.000	0.972	0.975	0.965
	10	0.90	1.000	0.000	0.000	0.000	1.000	0.969	0.963	0.974
		0.95	1.000	0.000	0.000	0.000	1.000	0.971	0.968	0.965
		1	0.502	0.535	0.306	0.079	0.999	0.976	0.986	0.967

Table 2.6: Probability of selecting the correct number of stochastic trends. Intercept model, setups 2 and 3 with $a = 0.5$, $N = 20$, panel data analysis, $PMSB_{\mu}^{Z}$ statistic

			Setup 2				Setup 3			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	1.000	0.976	0.978	0.973	1.000	0.980	0.979	0.978
		0.95	1.000	0.988	0.977	0.965	1.000	0.985	0.981	0.968
		1	0.955	0.988	0.981	0.977	0.880	0.984	0.982	0.970
	1	0.90	1.000	0.985	0.977	0.971	1.000	0.989	0.980	0.980
		0.95	1.000	0.985	0.984	0.961	1.000	0.978	0.982	0.982
		1	0.951	0.985	0.979	0.974	0.875	0.990	0.976	0.975
	10	0.90	1.000	0.980	0.981	0.975	1.000	0.974	0.979	0.979
		0.95	1.000	0.982	0.981	0.969	1.000	0.970	0.983	0.982
		1	0.957	0.983	0.980	0.969	0.862	0.980	0.976	0.974
200	0.5	0.90	1.000	0.982	0.972	0.967	1.000	0.983	0.975	0.983
		0.95	1.000	0.977	0.970	0.974	1.000	0.979	0.977	0.961
		1	0.984	0.982	0.958	0.974	0.946	0.982	0.972	0.967
	1	0.90	1.000	0.979	0.974	0.971	1.000	0.975	0.970	0.977
		0.95	1.000	0.978	0.963	0.977	1.000	0.979	0.966	0.978
		1	0.986	0.986	0.972	0.975	0.951	0.983	0.974	0.971
	10	0.90	1.000	0.979	0.965	0.972	1.000	0.974	0.969	0.980
		0.95	1.000	0.973	0.975	0.979	1.000	0.978	0.973	0.971
		1	0.991	0.985	0.965	0.972	0.960	0.981	0.961	0.961
500	0.5	0.90	1.000	0.979	0.972	0.968	1.000	0.977	0.970	0.965
		0.95	1.000	0.973	0.970	0.980	1.000	0.964	0.970	0.964
		1	1.000	0.981	0.971	0.966	0.993	0.974	0.961	0.968
	1	0.90	1.000	0.973	0.959	0.966	1.000	0.966	0.969	0.971
		0.95	1.000	0.976	0.965	0.962	1.000	0.972	0.970	0.976
		1	1.000	0.965	0.959	0.974	0.992	0.978	0.979	0.965
	10	0.90	1.000	0.975	0.970	0.965	1.000	0.960	0.963	0.965
		0.95	1.000	0.970	0.961	0.960	1.000	0.969	0.975	0.964
		1	1.000	0.983	0.978	0.964	0.992	0.975	0.973	0.974

Table 2.7: Probability of selecting the correct number of stochastic trends. Time trend model, setups 2 and 3 with $a = 0.5$, $N = 20$, panel data analysis, $PMSB_{\tau}^{Z}$ statistic

			Setup 2				Setup 3			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	1.000	0.980	0.980	0.984	1.000	0.987	0.985	0.989
		0.95	0.998	0.987	0.980	0.978	0.998	0.984	0.976	0.989
		1	0.968	0.974	0.982	0.974	0.937	0.978	0.978	0.973
	1	0.90	1.000	0.986	0.979	0.971	0.999	0.984	0.982	0.983
		0.95	0.999	0.976	0.981	0.976	0.995	0.980	0.978	0.990
		1	0.975	0.980	0.985	0.978	0.935	0.982	0.981	0.980
	10	0.90	1.000	0.981	0.984	0.976	1.000	0.987	0.980	0.980
		0.95	0.998	0.977	0.989	0.981	0.996	0.979	0.983	0.977
		1	0.969	0.978	0.978	0.976	0.933	0.983	0.978	0.976
200	0.5	0.90	1.000	0.974	0.971	0.975	1.000	0.968	0.977	0.969
		0.95	1.000	0.967	0.975	0.968	1.000	0.974	0.964	0.969
		1	0.993	0.959	0.970	0.971	0.982	0.979	0.967	0.979
	1	0.90	1.000	0.963	0.972	0.971	1.000	0.960	0.976	0.976
		0.95	1.000	0.957	0.972	0.971	1.000	0.962	0.967	0.965
		1	0.994	0.962	0.968	0.968	0.984	0.969	0.977	0.975
	10	0.90	1.000	0.962	0.975	0.973	1.000	0.969	0.965	0.978
		0.95	1.000	0.966	0.960	0.967	1.000	0.952	0.960	0.973
		1	0.996	0.961	0.967	0.980	0.986	0.973	0.958	0.970
500	0.5	0.90	1.000	0.974	0.976	0.956	1.000	0.975	0.979	0.970
		0.95	1.000	0.975	0.974	0.978	1.000	0.972	0.973	0.969
		1	1.000	0.975	0.976	0.965	1.000	0.972	0.961	0.970
	1	0.90	1.000	0.968	0.981	0.962	1.000	0.969	0.976	0.974
		0.95	1.000	0.977	0.971	0.962	1.000	0.965	0.976	0.960
		1	1.000	0.972	0.973	0.979	0.999	0.979	0.976	0.966
	10	0.90	1.000	0.961	0.966	0.970	1.000	0.973	0.978	0.964
		0.95	1.000	0.972	0.968	0.970	1.000	0.974	0.981	0.964
		1	0.999	0.975	0.983	0.959	0.998	0.967	0.978	0.967

Table 2.8: Probability of selecting the correct number of common factors. Intercept term, unit-by-unit analysis, MSB_{μ}^{F} test

			Setup 1, $q=1$				**Setup 2,** $q=1$				**Setup 3,** $q=2$			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	0.766	0.543	0.534	0.521	0.737	0.654	0.605	0.596	0.313	0.181	0.117	0.105
		0.95	0.351	0.268	0.257	0.248	0.316	0.292	0.262	0.256	0.035	0.026	0.016	0.013
		1	0.943	0.943	0.946	0.947	0.949	0.947	0.947	0.949	0.922	0.930	0.944	0.948
	1	0.90	0.730	0.596	0.581	0.575	0.714	0.674	0.640	0.636	0.263	0.203	0.152	0.147
		0.95	0.301	0.243	0.236	0.230	0.283	0.273	0.255	0.246	0.029	0.019	0.016	0.015
		1	0.936	0.942	0.946	0.947	0.941	0.944	0.951	0.953	0.953	0.954	0.956	0.956
	10	0.90	0.686	0.678	0.677	0.676	0.680	0.684	0.679	0.681	0.247	0.234	0.241	0.240
		0.95	0.259	0.258	0.258	0.257	0.260	0.257	0.261	0.261	0.021	0.021	0.021	0.021
		1	0.953	0.951	0.951	0.951	0.952	0.951	0.953	0.953	0.954	0.956	0.957	0.959
200	0.5	0.90	0.998	0.812	0.809	0.803	0.998	0.964	0.906	0.903	0.972	0.747	0.484	0.478
		0.95	0.771	0.528	0.514	0.512	0.746	0.673	0.617	0.613	0.319	0.197	0.111	0.107
		1	0.947	0.958	0.959	0.961	0.951	0.957	0.957	0.957	0.955	0.970	0.964	0.966
	1	0.90	0.999	0.912	0.910	0.912	0.999	0.989	0.972	0.972	0.978	0.829	0.686	0.683
		0.95	0.768	0.638	0.637	0.629	0.756	0.725	0.687	0.686	0.278	0.219	0.168	0.158
		1	0.942	0.948	0.951	0.950	0.948	0.941	0.951	0.952	0.956	0.963	0.963	0.965
	10	0.90	0.998	0.994	0.994	0.994	0.998	0.996	0.996	0.996	0.955	0.938	0.927	0.927
		0.95	0.729	0.698	0.697	0.697	0.724	0.713	0.709	0.709	0.237	0.235	0.240	0.239
		1	0.954	0.952	0.952	0.952	0.954	0.951	0.949	0.949	0.952	0.953	0.955	0.955
500	0.5	0.90	1.000	0.876	0.879	0.873	1.000	0.997	0.980	0.981	1.000	0.881	0.718	0.724
		0.95	1.000	0.836	0.838	0.834	1.000	0.989	0.953	0.953	1.000	0.813	0.609	0.608
		1	0.937	0.951	0.953	0.953	0.941	0.943	0.950	0.950	0.940	0.947	0.958	0.957
	1	0.90	1.000	0.974	0.972	0.972	1.000	1.000	0.996	0.995	1.000	0.970	0.923	0.923
		0.95	1.000	0.931	0.932	0.930	1.000	0.999	0.989	0.988	0.999	0.930	0.815	0.815
		1	0.960	0.964	0.963	0.964	0.962	0.960	0.960	0.960	0.936	0.939	0.948	0.948
	10	0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.998	0.994	0.995
		1	0.941	0.943	0.943	0.943	0.941	0.940	0.943	0.943	0.948	0.947	0.947	0.947

Table 2.9: Probability of selecting the correct number of common factors. Time trend, unit-by-unit analysis, MSB_{τ}^{F} test

			Setup 1, $q=1$				**Setup 2,** $q=1$				**Setup 3,** $q=2$			
T	σ_F^2	ρ	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$	$m=0$	$m=1$	$m=2$	$m=3$
100	0.5	0.90	0.355	0.266	0.251	0.237	0.326	0.293	0.266	0.263	0.050	0.041	0.028	0.024
		0.95	0.137	0.108	0.104	0.101	0.120	0.113	0.106	0.102	0.004	0.004	0.004	0.004
		1	0.936	0.945	0.946	0.950	0.948	0.947	0.950	0.951	0.927	0.933	0.942	0.949
	1	0.90	0.317	0.279	0.273	0.264	0.304	0.289	0.283	0.282	0.035	0.028	0.022	0.021
		0.95	0.124	0.103	0.099	0.094	0.110	0.112	0.101	0.100	0.007	0.006	0.007	0.006
		1	0.945	0.943	0.946	0.947	0.945	0.942	0.947	0.948	0.944	0.938	0.943	0.946
	10	0.90	0.284	0.279	0.278	0.278	0.285	0.280	0.276	0.276	0.028	0.026	0.030	0.030
		0.95	0.094	0.092	0.092	0.092	0.091	0.096	0.089	0.089	0.002	0.005	0.004	0.004
		1	0.948	0.955	0.956	0.956	0.950	0.953	0.953	0.953	0.944	0.944	0.939	0.939
200	0.5	0.90	0.778	0.661	0.652	0.645	0.757	0.737	0.692	0.689	0.480	0.385	0.285	0.273
		0.95	0.370	0.274	0.272	0.264	0.334	0.304	0.292	0.288	0.056	0.039	0.036	0.030
		1	0.943	0.962	0.962	0.960	0.949	0.954	0.959	0.959	0.945	0.954	0.951	0.953
	1	0.90	0.803	0.735	0.731	0.725	0.794	0.779	0.745	0.742	0.451	0.402	0.357	0.354
		0.95	0.335	0.303	0.293	0.292	0.324	0.300	0.301	0.298	0.033	0.025	0.025	0.024
		1	0.950	0.953	0.951	0.951	0.952	0.953	0.948	0.949	0.938	0.943	0.948	0.950
	10	0.90	0.771	0.755	0.755	0.755	0.772	0.766	0.767	0.767	0.399	0.388	0.380	0.380
		0.95	0.291	0.284	0.284	0.284	0.289	0.291	0.284	0.283	0.040	0.040	0.030	0.030
		1	0.945	0.942	0.942	0.942	0.945	0.945	0.945	0.945	0.956	0.953	0.954	0.954
500	0.5	0.90	0.995	0.944	0.945	0.941	0.995	0.990	0.977	0.976	0.983	0.918	0.854	0.854
		0.95	0.913	0.776	0.773	0.770	0.902	0.864	0.824	0.826	0.741	0.605	0.484	0.473
		1	0.934	0.946	0.949	0.949	0.940	0.943	0.947	0.947	0.935	0.938	0.948	0.950
	1	0.90	0.999	0.986	0.986	0.986	0.999	0.998	0.993	0.993	0.988	0.958	0.928	0.927
		0.95	0.900	0.847	0.846	0.844	0.894	0.873	0.867	0.866	0.700	0.638	0.572	0.571
		1	0.938	0.935	0.934	0.935	0.941	0.935	0.942	0.942	0.944	0.944	0.943	0.944
	10	0.90	0.994	0.993	0.993	0.993	0.994	0.993	0.993	0.993	0.983	0.980	0.975	0.975
		0.95	0.900	0.894	0.893	0.892	0.899	0.897	0.893	0.893	0.707	0.702	0.697	0.695
		1	0.952	0.954	0.954	0.954	0.953	0.954	0.953	0.953	0.959	0.958	0.957	0.957

Table 2.10: Panel data unit root tests

PANEL A: MONEY DEMAND MODEL

	Pesaran (2004, 2013) CD test		Moon and Perron (2004)			
	CD	p-val	t_a	p-val	t_b	p-val
$gdp_{i,t}$	21.610	0.000	-1.151	0.125	-0.998	0.159
$m_{i,t}$	19.017	0.000	-2.748	0.003	-3.026	0.001
$R_{i,t}$	22.535	0.000	-14.502	0.000	-6.540	0.000
Whole	27.824	0.000				

	Bai and Ng (2004) statistics					
	ADF_e	p-val	MQ_f	$(\hat{q},\hat{q}_1)$	MQ_c	$(\hat{q},\hat{q}_1)$
$gdp_{i,t}$	-0.711	0.238	-10.243	(2,2)	-14.346	(2,2)
$m_{i,t}$	1.191	0.883	-4.155	(1,1)	-7.462	(1,1)
$R_{i,t}$	-4.064	0.000	-37.939	(6,6)	-36.753	(6,6)

	Pesaran (2007) CIPS statistic					
	CADF(0)	CADF(1)	CADF(2)	CADF(3)	CADF(4)	CADF(5)
$gdp_{i,t}$	-2.275	-2.532	-2.475	-2.405	-1.997	-1.912
$m_{i,t}$	-2.380	-2.311	-2.285	-2.099	-2.127	-1.881
$R_{i,t}$	-3.184	-2.676	-2.314	-2.195	-1.918	-2.079

PANEL B: MONETARY EXCHANGE MODEL

	Pesaran (2004, 2013) CD test		Moon and Perron (2004)			
	CD	p-val	t_a	p-val	t_b	p-val
$gdp_{i,t}$	26.969	0.000	-0.143	0.443	-0.091	0.464
$m_{i,t}$	29.277	0.000	-1.147	0.126	-0.711	0.239
$exch_{i,t}$	66.118	0.000	0.305	0.620	3.498	1.000
Whole	28.846	0.000				

	Bai and Ng (2004) statistics					
	ADF_e	p-val	MQ_f	$(\hat{q},\hat{q}_1)$	MQ_c	$(\hat{q},\hat{q}_1)$
$gdp_{i,t}$	-0.921	0.178	-7.408	(1,1)	-11.562	(1,1)
$m_{i,t}$	-3.169	0.001	-17.426	(6,6)	-30.490	(6,6)
$exch_{i,t}$	1.832	0.967	-36.238	(6,6)	-37.207	(6,6)

	Pesaran (2007) CIPS statistic					
	CADF(0)	CADF(1)	CADF(2)	CADF(3)	CADF(4)	CADF(5)
$gdp_{i,t}$	-3.662	-2.939	-2.869	-2.193	-2.461	-2.388
$m_{i,t}$	-1.189	-3.126	-2.874	-2.518	-1.961	-2.678
$exch_{i,t}$	-2.024	-2.043	-1.969	-1.884	-1.802	-1.790

$\hat{q}$ denotes the total (stationary and non-stationary) number of estimated common factors and $\hat{q}_1$ the number of non-stationary ones. The number between parentheses in CADF(p), $p = 0, 1, ..., 5$, indicates the lag augmentation used in the cross-sectionally augmented ADF test computation

Table 2.11: Individual and panel data cointegration rank statistics

Panel A: Money demand model				**Panel B: Monetary exchange model**			
Individual idiosyncratic systems				**Individual idiosyncratic systems**			
	$m=3$	$m=2$	$m=1$		$m=3$	$m=2$	$m=1$
AUS	0.042	0.077	0.177	AUS	0.027	0.145	0.246
AUT	0.030	0.048	0.453	AUT	0.065	0.053	0.141
BEL	0.036	0.038	0.440	BEL	0.037	0.108	0.346
CAN	0.055	0.060	0.364	CAN	0.050	0.124	1.084
DEN	0.027**	0.055	0.127	DEN	0.044	0.059	0.396
FIN	0.032	0.173	0.543	FIN	0.059	0.082	0.280
FRA	0.036	0.248	0.559	FRA	0.021	0.162	0.231
GER	0.030	0.145	0.457	GER	0.051	0.062	0.131
ICE	0.023**	0.115	0.530	GRE	0.078	0.229	0.237
IRE	0.063	0.103	0.274	ITA	0.042	0.121	0.212
JAP	0.035	0.097	0.239	JAP	0.030	0.094	0.300
NET	0.040	0.049	0.443	KOR	0.032	0.107	0.108
NEZ	0.044	0.092	0.319	NET	0.033	0.175	0.182
NOR	0.037	0.059	0.384	NOR	0.004**	0.058	0.542
POR	0.028	0.137	0.630	SPA	0.018**	0.052	0.760
SPA	0.029	0.057	0.296	SWE	0.167	0.140	0.736
SWI	0.084	0.148	0.585	SWI	0.034	0.047	0.098
UK	0.040	0.040	0.495	UK	0.067	0.090	0.502
US	0.038	0.078	0.323				
Panel data cointegrating rank tests				**Panel data cointegrating rank tests**			
	$m=3$	$m=2$	$m=1$		$m=3$	$m=2$	$m=1$
$PMSB_{\tau}^{Z}$	-1.884**	3.905	7.716	$PMSB_{\tau}^{Z}$	3.248	5.708	5.431
$PMSB_{\tau}^{F}$	78.660**	23.093	5.005	$PMSB_{\tau}^{F}$	40.658	9.713	10.967
$PMSB_{\tau}^{C}$	4.664**	-1.710	-3.785	$PMSB_{\tau}^{C}$	0.549	-3.098	-2.950
Common factors				**Common factors**			
	Test	$\hat{q}_1$			Test	$\hat{q}_1$	
MSB_{τ}^{F}	0.414	2		MSB_{τ}^{F}	0.126	3	
MQ_f	-6.756	2		MQ_f	-24.535	3	
MQ_c	-6.958	2		MQ_c	-29.088	3	

m denotes the number of stochastic trends that is specified under the null hypothesis. The maximum number of common factors is six. The country specific numbers are the MSB statistics. In panel A, the 5% critical values for the individual MSB tests are 0.027 (for $m=3$), 0.033 (for $m=2$) and 0.049 (for $m=1$). In panel B, the 5% critical values for the individual MSB tests are 0.021 (for $m=3$), 0.027 (for $m=2$) and 0.042 (for $m=1$).** Denotes rejection of the null hypothesis at the 5% level of significance. $\hat{q}_1$ is the estimated number of non-stationary common factors

Chapter 3

Estimation of Production Functions: The Spanish Regional Case

3.1 Introduction

The estimation of production function that relates the output of a firm, region or country to different combinations of factors of production – usually physical capital and labor – have devoted lot of interest in empirical economics – see Aschauer (1989), Munnell (1990), García-Milá and McGuire (1992), Holtz-Eakin (1994), Baltagi and Pinnoi (1995), and García-Milá, McGuire and Porter (1996) for the US, Merriman (1990) for Japan, Berndt and Hansson (1992) for Sweden, Evans and Karras (1994) for a sample of industrialized countries, Dalamagas (1995) for Greece, Otto and Voss (1996) for Australia, and Wylie (1996) for Canada. These studies estimate production functions including not only physical capital and labor as inputs, but also human and public capitals as productive factors.

Early studies of production function employ time series data, focusing on an individual region or country. For example, for the case of the aggregated Spanish economy, Serrano (1997) uses annual data observed over the period 1964-1991 and finds no evidence of cointegration between gross value added, human capital, private physical capital and labor. Sosvilla-Rivero and Alonso (2005) obtain a different result and find that in Spain, gross domestic product (GDP), physical capital, human capital and labor define a cointegration relation. They employ annual data that is observed over the period 1910-1995. The contradictory results indicate that the empirical evidence from the time series studies is mixed. One plausible explanation is the low power of the univariate unit root and cointegration tests that were used in these studies.

More recent studies show that the power of unit root and cointegration test statistics can be improved when both the time series and cross-section dimensions are combined. Thus, another category of studies that estimate long-run production functions using panel data tools emerges. For example, Serrano (1996) employs a panel of regional data observed over the years 1980-1991 and simply avoids the risk of a spurious regression working with a model that relates the gross value added, human capital, physical capital and labor in first difference.[1] Bajo-Rubio and Díaz-Roldán (2005) go one step further and add public capital to the production function. The authors investigate the relation between GDP, private capital, public capital, human capital and labor using data for the 17 Spanish regions over 1965-1995. They find that these variables are cointegrated. Recently, Márquez, Ramajo and Hewings (2011) investigate the relation between gross value added, public capital, private capital and labor for the 17 Spanish regions observed over the period 1972-2000. The authors find that indeed, there is cointegration among these variables.

One critical problem with the panel data studies for the Spanish regions mentioned above is the assumption of cross-section independence made. This is an unrealistic and far too restrictive assumption from an empirical point of view, especially since regions are so closely related to each other – Spanish regions share a common institutional framework with a common fiscal system that is used to finance the regional governments. If the independence assumption is violated then we might expect to have, on the one hand, biased and inconsistent estimates of the parameters and, on the other hand, spurious statistical inference – see Andrews (2005). More specifically, in the case of non-stationary panel data, the unaccounted cross-section dependence might lead to conclude that panel data is actually I(0) stationary when in fact it might be I(1) non-stationary – see Banerjee, Marcellino and Osbat (2005). Similarly, the panel data cointegration test statistics might indicate than there are more cointegrating relations than there exist – see Carrion-i-Silvestre and Surdeanu (2011). Consequently, accounting for the presence of cross-section dependence is crucial to draw meaningful conclusions from the analysis.

Cross-section dependence is more a recurrent than a rare characteristic that is present in macroeconomic time series of different units – i.e., countries, regions or sectors. There are different sources of cross-section dependence that can be expected to affect the units of a panel data set. For instance, cross-section dependence

[1] Note that this leads to the estimation of a short-run relation among the variables since the long-run one would require to use the variables in levels, not in first difference.

is usually caused by the presence of common shocks (oil price shocks or financial crises) or the existence of local productivity spillover effects. Further, the economic literature on output stochastic convergence implies the existence of a long-run relation (cointegration relation) among the different economies, so that the use of macroeconomic variables such as the output or production should account for the presence of this long-run relation across the cross-section – the so-called cross-cointegration concept, as defined in Banerjee, Marcellino and Osbat (2005). This implies that cross-section dependence is more the rule than the exception. Therefore, in country or regional level studies is practically impossible to ignore the effect of cross-section dependence in the analysis of the models that are to be estimated. Bai and Ng (2002, 2004) recognize early on this problem and lay down the foundation of the theoretical panel framework with common factors. The use of common factor models is particularly useful to capture the presence of cross-section that is pervasive or strong, i.e., the sort of cross-section dependence that affects all units of the panel data.

However, as Banerjee, Eberhardt and Reade (2010) mention, the empirical work on the estimation of production function in panel data using the common factor technique is relatively limited. Two examples related to our study is the work by Costantini and Destefanis (2009) and Banerjee and Carrion-i-Silvestre (2011). Costantini and Destefanis (2009) analyze the production function for the Italian regions over the 1970-2003 period and find that the regional value added, physical capital and human capital augmented labor are cointegrated. They also find that ignoring the cross-section dependence biases upward the estimates for the returns to scale. Similarly, Banerjee and Carrion-i-Silvestre (2011) estimate a production function with GDP, labor and capital stock using a panel of 19 developed countries covering the period 1951-2007, concluding that there is a long-run relation among these variables. Note that these last studies use the single-equation framework while our proposal focuses on both single-equation and vector autoregressive (VAR) frameworks. The advantage of a VAR model is knowing exactly how many cointegration relations or, conversely, how many stochastic trends exist among the units of the panel. To the best of our knowledge, none of the existing studies for the Spanish economy take into consideration the (strong) cross-section dependence among the members of the panel either in a single or a system-based approaches.

In this chapter, we reexamine the cointegration relation among the output, physical capital, human capital, public capital and labor for the 17 Spanish regions observed over the period 1964-2000. In order to analyze the order of integration of the variables in our model we apply the panel data unit root test statistics in Bai and Ng (2004), Moon and Perron (2004) and Pesaran (2007) and the panel data stationarity test statistics in Hadri (2000). All these test statistics account for the existence of cross-section dependence in different ways. In general, the application of these statistics leads to the same qualitative conclusion, i.e., that all panel data sets are characterized as I(1) non-stationary panels. We then use the panel cointegration statistics recently proposed in Carrion-i-Silvestre and Surdeanu (2011) using a VAR framework and in Westerlund (2008) and Banerjee and Carrion-i-Silvestre (2011, 2013) for the single-equation framework. All the cointegration statistics allow for cross-section dependence through the use of common factors. By using the vector autoregressive model we are able to determine the exact number of cointegrating vectors that exist in the model. Then, we compute the panel data estimators proposed in Bai, Kao and Ng (2009) and Kapetanios, Pesaran and Yamagata (2011) to estimate the long-run production function for the Spanish regions.

Our analysis bases on the Cobb and Douglas (1928) production function specification. This specification form is widely used, relies on few parameters and is easy to estimate. These characteristics make it easy for researchers to have their estimates compared with previous studies. Murthy (2002) provides an extensive discussion of the Cobb-Douglas approach and shows that many of the econometric estimation problems can be easily addressed. The most common criticism of the Cobb-Douglas approach is the inflexible function form of the production function. Except for its inflexibility, all other assumptions can be relaxed.

To overcome the flexibility issue, Christensen, Jorgenson and Lau (1971) proposed the translog approach which is a generalization of the Cobb-Douglas form. The translog is estimated by adding the squares and cross products of the logs of all explanatory variables. The advantage of the translog is the flexible functional form, meaning less restrictions on production elasticities. However, it should be bear in mind that it is difficult to interpret and requires estimation of many parameters. Another drawbacks of the translog specification are multicollinearity and degrees of freedom problems due to the inclusion of the second order terms. Finally, the translog parameters do not have a direct interpretation compared with those from Cobb-Douglas, where they are interpreted as elasticities – see Felipe (1998). Consequently, the majority of researchers use the more restrictive Cobb-Douglas form. In this chapter and as a robustness check, we also estimate the translog specification allowing for cross-section dependence through a common factor model, which, to the best of our knowledge, has not been applied in the literature. Finally, the robustness check also covers the issue of spatial dependence, a form of weak dependence that might be affecting the Spanish regions.

The structure of this chapter is as follows. Section 3.2 presents the model for panel data and the data used in this study. In Section 3.3 we present the econometric methodology while the results are presented in Section 3.4. Finally, this chapter concludes with Section 3.5. The database used in this chapter is described in Appendix C.

3.2 Specification of the model and the data

This section presents the production function and the data used in this study. We employ a modified Cobb-Douglas production function, used also by Bajo-Rubio and Díaz-Roldán (2005), that has the following form:

$$Y_{i,t} = A_{i,t}\, F\left(K_{i,t}, G_{i,t}, H_{i,t}, L_{i,t}\right), \tag{3.2.1}$$

where $i = 1, \ldots, N$ represents the cross-section dimension and $t = 1, \ldots, T$ represents the time series dimension. The variable $Y_{i,t}$ is the output that depends on the private capital $(K_{i,t})$, the public capital $(G_{i,t})$, the human capital $(H_{i,t})$ and the labor $(L_{i,t})$. The variable $A_{i,t}$ is the total factor productivity (TFP), which is the part of the output not explained by the inputs. Next, we express the production function in per worker terms, obtaining:

$$Y_{i,t}/L_{i,t} = A_{i,t}/L_{i,t}\, f\left(K_{i,t}/L_{i,t}, G_{i,t}/L_{i,t}, H_{i,t}/L_{i,t}\right). \tag{3.2.2}$$

As it is well known, TFP represents the unobservable part of the production function and usually reflects the technological progress of the respective country or region. Further, if the technology represents the cumulation of the innovations and progress efforts made by economic agents, we should expect the TFP to be an I(1) non-stationary stochastic process. However, since the TFP cannot be measured directly, the empirical researchers estimate it as the residual of the estimated production function. Although intuitive, this approach causes serious econometric and interpretation problems. First, if not appropriately accounted for, the potential stochastic trend of the TFP would imply that the estimation of the production function is, in fact, a spurious regression. Therefore, panel data cointegration test statistics would lead to the conclusion that the variables involved in the production function are not cointegrated. Second, the issue that part of the technology that is available is common to all the economies implies a source of cross-section dependence, which needs to be accounted for in order to obtain meaningful conclusions of the panel cointegration test statistics. As can be seen, the specification of a common factor model can capture this unobservable variable that is difficult to approximate.

We take advantage of the recent developments in the field of non-stationary panel data analysis and decompose the TFP into an unobserved common factor component $F_t'\lambda_i$ – where F_t is a $(r\times 1)$-vector of unobserved common factors, λ_i is a $(r\times 1)$-vector of loadings – and an idiosyncratic component $e_{i,t}$. The common factor approach allows us to capture the effect of common shocks that affect the countries or regions, making it a desirable way to model the cross-section dependence. Therefore, following Costantini and Destefanis (2009) and Banerjee, Eberhardt and Reade (2010), TFP is modeled through the common factor specification given by:

$$A_{i,t}/L_{i,t} = e^{D_{i,t}+F_t'\lambda_i+e_{i,t}}, \tag{3.2.3}$$

where $D_{i,t}$ denotes the deterministic component being either a constant $(D_{i,t}=\mu_i)$ or a linear time trend $(D_{i,t}=\mu_i+\delta_i t)$. Assuming a Cobb-Douglas function and taking the natural logarithm of the variables from (3.2.2) and (3.2.3), we obtain the model:

$$\begin{aligned} y_{i,t} &= a_{i,t}+(\alpha+\beta+\gamma+\delta-1)\,l_{i,t}+\alpha k_{i,t}+\beta g_{i,t}+\delta h_{i,t} && (3.2.4)\\ a_{i,t} &= D_{i,t}+F_t'\lambda_i+e_{i,t}, && (3.2.5)\end{aligned}$$

where $y_{i,t}=\ln(Y_{i,t}/L_{i,t})$, $a_{i,t}=\ln(A_{i,t}/L_{i,t})$, $l_{i,t}=\ln L_{i,t}$, $k_{i,t}=\ln(K_{i,t}/L_{i,t})$, $g_{i,t}=\ln(G_{i,t}/L_{i,t})$ and $h_{i,t}=\ln(H_{i,t}/L_{i,t})$. Note that the model can be written in a single-equation form as:

$$y_{i,t}=D_{i,t}+(\alpha+\beta+\gamma+\delta-1)\,l_{i,t}+\alpha k_{i,t}+\beta g_{i,t}+\delta h_{i,t}+F_t'\lambda_i+e_{i,t}. \tag{3.2.6}$$

Following the existing contributions in the literature, we propose two alternative measures for public capital and human capital. First, we use the total public capital $(g_{i,t})$ or the productive public capital $(gp_{i,t}=\ln(Gp_{i,t}/L_{i,t}))$.[2] Second, the human capital is proxied by the rate of employees with at least secondary school studies over the total number of employees $(h_{i,t})$ and the average number of schooling years

[2]Productive public capital includes road and highways, ports, airports, railroads, water and sewer systems, public electric and gas utilities, and telecommunications.

($hs_{i,t} = \ln(Hs_{i,t}/L_{i,t})$) – see Serrano (1996). The use of these variables defines up to four different model specifications depending on whether total or productive public capital is used and on whether we use $h_{i,t}$ or $hs_{i,t}$ to proxy the human capital.

The data employed in our study contains annual observations for the $N = 17$ Spanish regions observed over the $T = 37$ year period from 1964 to 2000. We collect the data from the BD.MORES database provided by the Spanish Ministry of Economy and Finance and from the Instituto Valenciano de Investigaciones Economicas (IVIE) – for a detailed description of the variables and the sources, see the Appendix C. The Spanish regions are: Andalucía, Aragón, Asturias, Baleares, Canarias, Cantabria, Castilla y León, Castilla-La Mancha, Catalunya, Comunidad Valenciana, Extremadura, Galicia, Madrid, Murcia, Navarra, País Vasco and La Rioja. The picture of the variables can be found in Figure 3.1 which evidences first, the clear trend pattern shown by the variables of the model and, second, the co-movement (cross-section dependence) that seems to be present in their evolution.

3.3 Econometric methodology

In this section we describe the tools that are used throughout the chapter in order to analyze our dataset. The order in which we present the econometric procedures is the one that will be followed when applying them in the empirical estimation of the regional Spanish production function. Since the validity of the panel data unit root, stationarity and cointegration test statistics requires to assess whether the units in the panel data set are cross-section dependent, we first start the discussion describing Pesaran's (2004, 2013) CD test statistic that tests the null hypothesis of cross-section independence against the alternative hypothesis of cross-section dependence. Note that in our case we are analyzing macroeconomic time series of highly economic integrated regions, provided that the regions belong to the same economy – see Figure 3.1. Therefore we can expect the presence of cross-section dependence among the units of the panel. Second, we present the panel unit root and stationarity tests that control for the presence of cross-section dependence in different ways. To be specific, we apply the panel data unit root tests in Bai and Ng (2004), Moon and Perron (2004) and Pesaran (2007), and the panel stationarity tests in Hadri (2000). Finally, we summarize recent developments in panel cointegration testing and estimation that take into consideration the cross-section dependence. In this regard, we first focus on the system-based approach proposed in Chapter 2 – see also Carrion-i-Silvestre and Surdeanu (2011). Then, we consider the single-equation-based procedures proposed in Westerlund (2008) and Banerjee and Carrion-i-Silvestre (2011, 2013) to test whether a cointegration relation among the variables of the model exists. Finally, we proceed to estimate the cointegration relations using the proposals in Bai, Kao and Ng (2009) and Kapetanios, Pesaran and Yamagata (2011). As pointed out below, these statistical procedures allow for cross-section dependence through the specification of a common factor model.

3.3.1 Cross-section dependence

In this subsection we test the null hypothesis of cross-section independence against the alternative hypothesis of cross-section dependence using the approach suggested in Pesaran (2004, 2013). For notational convenience, throughout this and the next section, we will use $y_{i,t}$ as the variable of interest, although the same applies for the other variables of the system – i.e., $k_{i,t}$, $g_{i,t}$, $gp_{i,t}$, $h_{i,t}$, $hs_{i,t}$ and $l_{i,t}$. The test statistic is based on the average of pair-wise Pearson's correlation coefficients $\hat{\rho}_{i,j}$, $i = 1,2,\ldots,N-1,\ j = i+1,2,\ldots,N$ – i.e., we have $n = N(N-1)/2$ correlation coefficients – of the residuals $\varepsilon_{i,t}$ obtained from the following augmented Dickey-Fuller (ADF) type regression equation:[3]

$$\Delta y_{i,t} = \mu_i + \delta_i t + \alpha_{i,0} y_{i,t-1} + \sum_{j=1}^{p_i} \alpha_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t}, \tag{3.3.1}$$

$i = 1,\ldots,N$. Pesaran (2004, 2013) CD test is based on averaging all pair-wise correlation coefficients ($\hat{\rho}_{i,j}$) of the Ordinary Least Squares (OLS) estimated residuals $\hat{\varepsilon}_{i,t}$ in (3.3.1):

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{i,j} \right), \tag{3.3.2}$$

with $i = 1,\ldots,N-1$ and $j = i+1,\ldots,N$. Under the null hypothesis of cross-section independence, the CD statistic of Pesaran (2004, 2013) converges to the standard normal distribution. The simulations that he conducts shows that the CD statistic has the correct size and satisfactory power even in small samples, making it attractive in the empirical research.

[3]Throughout this chapter, the cross-section dependence and order of integration analyses assume a linear time trend as the deterministic component of the auxiliary regressions that are estimated, provided the trending behavior of the variables drawn from Figure 3.1.

3.3.2 Panel data unit root and stationarity test statistics

The panel data unit root tests applied in this chapter are those developed by Bai and Ng (2004), Moon and Perron (2004) and Pesaran (2007). All of them take into account the presence of cross-section dependence by specifying a model of common factors. Of the three approaches the most general one is the one in Bai and Ng (2004) because it allows to test the order of integration of the idiosyncratic and common components in a separate way. For presentation clarity, we first describe Pesaran (2007) approach, then continue with Moon and Perron (2004) approach and finally, we discuss the proposal in Bai and Ng (2004). In addition and as a robustness analysis, we also compute the panel data stationarity test statistics in Hadri (2000).

Pesaran (2007) panel data unit root test statistic

The approach in Pesaran (2007) assumes that the cross-section dependence is driven by one unobservable stationary common factor, which can be proxied using cross-section averages of the units that define the panel data set. For the case of uncorrelated residuals, the starting regression has the following form:

$$\Delta y_{i,t} = \mu_i + \delta_i t + \alpha_i y_{i,t-1} + \lambda_i f_t + \varepsilon_{i,t},$$

where $\Delta y_{i,t} = y_{i,t} - y_{i,t-1}$, f_t denotes the unobserved common factor and $\varepsilon_{i,t}$ is the idiosyncratic error. The common factor f_t can be proxied by the cross-section mean of $y_{i,t}$ (i.e., $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{i,t}$) and its lagged values ($\bar{y}_{t-1}, \bar{y}_{t-2}\ldots$). Pesaran (2007) notes that these cross-section averages are sufficient for eliminating the effect of the common factor. Therefore, after substituting the proxies we obtain the modified cross-sectionally ADF (CADF) regression:

$$\Delta y_{i,t} = \mu_i + \delta_i t + \alpha_{i,0} y_{i,t-1} + \sum_{j=1}^{p} \alpha_{i,j}\Delta y_{i,t-j} + \xi_i \bar{y}_{t-1} + \sum_{j=0}^{p} \eta_{i,j}\Delta\bar{y}_{t-j} + e_{i,t}. \quad (3.3.3)$$

One of the panel unit root statistics proposed by Pesaran (2007) consists of the average of the individual CADF statistics:

$$CIPS(N,T) = N^{-1}\sum_{i=1}^{N} CADF_i,$$

where $CADF_i$ is the cross-sectionally ADF statistic for the i-th cross-section unit given by the t-ratio of the OLS estimate of $\alpha_{i,0}$ in (3.3.3).

Pesaran (2007) also proposes a truncated version of the $CIPS$ test, denoted $CIPS^*$. This statistic is useful when T is small, usually between 10 and 20. The author presents the simulated critical values of $CIPS$ and $CIPS^*$ in his paper. If the value of the $CIPS$ and $CIPS^*$ statistics is smaller than the critical value, then the null hypothesis of panel unit root is rejected. Although Pesaran's (2007) framework allows for only one stationary common factor, Pesaran, Smith and Yamagata (2013) show that it is also valid when there is either more than one common factor and/or the common factors are I(0) or I(1).

Moon and Perron (2004) panel data unit root test statistics

Moon and Perron (2004) propose two statistics that test for the presence of a panel data unit root while accounting for cross-sectional dependence among the units of the panel. Their approach is based on an approximate common factor model and tests for the unit root in the defactored series. The authors consider an autoregressive process in which the error term follows a factor structure:

$$\begin{aligned} y_{i,t} &= \mu_i + \delta_i t + y^0_{i,t} \\ y^0_{i,t} &= \rho_i y^0_{i,t-1} + u_{i,t} \\ u_{i,t} &= F_t' \lambda_i + e_{i,t}. \end{aligned}$$

Moon and Perron (2004) first transform the model in order to eliminate the common factors and obtain a defactored data that has no cross-sectional dependence. In the second step, they construct the panel unit root tests t_a and t_b using the defactored data. Then the null hypothesis $H_0 : \rho_i = 1 \ \forall i$ against the alternative hypothesis $H_1 :$ $\rho_i < 1$ for some i, $i = 1, \ldots, N$, is tested, using the following pooled test statistics:

$$\begin{aligned} t_a &= \frac{T\sqrt{N}\left(\hat{\rho}^*_{pool} - 1\right)}{\sqrt{2\hat{\phi}^4_e / \hat{\omega}^4_e}} \\ t_b &= T\sqrt{N}\left(\hat{\rho}^*_{pool} - 1\right)\sqrt{\frac{1}{NT^2} tr\left(Z_{-1} Q Z'_{-1}\right)}\left(\frac{\hat{\omega}_e}{\hat{\phi}^2_e}\right), \end{aligned}$$

where the terms $\hat{\rho}^*_{pool}$, $\hat{\phi}_e$, $\hat{\omega}_e$, Z_{-1} and Q are defined in Moon and Perron (2004). Under the null hypothesis of panel unit root, both test statistics converge to the standard normal distribution. The null hypothesis of a unit root is rejected if the value of the t_a or t_b statistics is smaller than the critical value drawn from the standard normal distribution. Moon and Perron (2004) also show that estimating the common factors by principal components leads to feasible statistics with the same limiting distribution as if they were observable.

Bai and Ng (2004) panel data unit root test statistics

Bai and Ng (2004) decompose the observable variable $y_{i,t}$ into a deterministic component $D_{i,t}$, a common component $\lambda_i' F_t$ and an idiosyncratic component $e_{i,t}$:

$$\begin{aligned} y_{i,t} &= D_{i,t} + \lambda_i' F_t + e_{i,t} && (3.3.4)\\ (1-L)\, F_{j,t} &= C_j(L)\, w_{j,t}; \quad j = 1\ldots, r && (3.3.5)\\ (1-\rho_i L)\, e_{i,t} &= H_i(L)\, \varepsilon_{i,t}, && (3.3.6) \end{aligned}$$

where $D_{i,t}$ denotes the deterministic part of the model – either a constant or a linear time trend – F_t is a $(r \times 1)$-vector that accounts for the common factors that are present in the panel, and $e_{i,t}$ is the idiosyncratic disturbance term. The $(r \times 1)$-vector of loading parameters λ_i measures the effect that the common factors have on the i-th time series. Unobserved common factors and idiosyncratic disturbance terms are estimated using principal components on the first difference model. The estimation of the number of common factors is obtained using the panel Bayesian information criterion (BIC) in Bai and Ng (2002).

Once both the idiosyncratic and common components have been estimated, we can proceed to test their order of integration using unit root tests. First, it is possible to test whether there are I(0) and/or I(1) common factors (F_t) using the ADF (for the one common factor case) or the MQ test statistics in Bai and Ng (2004) (for the general case where there is more than one common factor). We consider both the parametric ($MQ_f^j(m)$) and non-parametric ($MQ_c^j(m)$) versions of the MQ test statistics, where $j = c$ for the model that includes a constant, $j = \tau$ for the model that includes a linear time trend and m denotes the number of stochastic trends considered under the null hypothesis. The critical values for up to six factors for the MQ tests can be found in Table I of Bai and Ng (2004). For the case of one common factor we use the usual critical values of the Dickey-Fuller test. Therefore, using these statistics we will be able to conclude how many (if any) of the r estimated common factors are I(0) stationary (r_0) and how many are I(1) non-stationary (r_1), so that $r = r_0 + r_1$.

We can test the panel unit root hypothesis focusing on the idiosyncratic shocks ($e_{i,t}$). For this case, Bai and Ng (2004) propose to compute the usual ADF pseudo t-ratio statistic applied to the idiosyncratic component. If the model contains only an intercept, the pseudo t-ratio statistic is denoted as $ADF_{\hat{e}}^{c}$ and its asymptotic distribution coincides with the Dickey-Fuller distribution for the case of no constant. For the model that includes a linear trend the statistic is denoted as $ADF_{\hat{e}}^{\tau}$ and its asymptotic distribution is a function of a Brownian bridge.

As can be seen, this technique can determine the source of the non-stationarity that is present in the observable variable, if this is the case. It is possible that the non-stationarity of the observed variables is the result of the presence of I(1) common factors – or a combination of I(0) and I(1) common factors. This would imply that the panel data set is I(1) non-stationary and that the source of non-stationarity is a common cause for all the units that define the panel. In this case, we should conclude that there are global permanent shocks affecting the whole panel. Another source of panel non-stationarity could be given by the non-stationarity of the idiosyncratic disturbance terms processes, a fact that implies that shocks affecting each time series – i.e., not the global shocks – have a permanent character.

The approach of Bai and Ng (2004) nests the ones in Moon and Perron (2004) and Pesaran (2007). As noted by Bai and Ng (2010), Moon and Perron (2004) and Pesaran (2007) control the presence of cross-section dependence allowing for common factors, although the common factors and idiosyncratic shocks are restricted to have the same order of integration. Therefore, it is not possible to cover situations in which one component (e.g., the common factors) is I(0) and the other component (for example, the idiosyncratic shocks) is I(1), and vice versa. In practical terms, the test statistics in Moon and Perron (2004) and Pesaran (2007) turn out to be statistical procedures to make inference only on the idiosyncratic shocks, where the dynamics of both the idiosyncratic and the common components are restricted to be the same.

Hadri (2000) panel data stationarity test statistics

The panel data stationarity test statistic in Hadri (2000) specifies the null hypothesis that the units in the panel data set are I(0) against the alternative hypothesis that there are some units that are I(1). The test is based on the OLS estimation of the following regression equation:

$$y_{i,t} = D_{i,t} + u_{i,t}, \tag{3.3.7}$$

where $D_{i,t}$ denotes the deterministic component. The estimated residuals from Equation (3.3.7) are used to define the partial sum processes $\hat{S}_{i,t} = \sum_{j=1}^{t} \hat{u}_{i,j}$ for each unit. Using this individual information, Hadri (2000) proposes a panel data stationarity test:

$$LM^j = N^{-1} \sum_{i=1}^{N} \eta_i^j,$$

where $\eta_i^j = \hat{\omega}_i^{-2} T^{-2} \sum_{t=1}^{T} \hat{S}_{i,t}^2$, $i = 1, \ldots, N$, denotes the individual stationarity test statistic proposed in Kwiatkowski, Phillips, Schmidt and Shin (1992) – KPSS henceforth – where $j = c$ for the model that only includes a constant ($D_{i,t} = \mu_i$) and $j = \tau$ for the one that includes a linear time trend ($D_{i,t} = \mu_i + \delta_i t$), with $\hat{\omega}_i^2$ being a consistent estimate of the long-run variance of the error term $u_{i,t}$ – Carrion-i-Silvestre, del Barrio Castro and López-Bazo (2005) suggest to estimate the long-run variance following the procedure described by Sul, Phillips and Choi (2005), using the Quadratic spectral kernel. At this stage, we should mention that it is possible to compute two different LM statistics, depending on whether the long-run variance is allowed to be heterogeneous across i (LM_{HET}^j) or homogeneous for all individuals (LM_{HOM}^j) – in the latter case we use $\hat{\omega}^2 = N^{-1} \sum_{i=1}^{N} \hat{\omega}_i^2$. After standardizing the LM statistic by its mean and variance and assuming that $u_{i,t}$ in (3.3.7) are cross-section independent, the authors derive the new test Z_k^j, $k = \{HOM, HET\}$, that has the following distribution under the null hypothesis of panel stationarity:

$$Z_k^j = \frac{\sqrt{N}(LM_k^j - \overline{\xi^j})}{\overline{\varsigma^j}} \Rightarrow N(0,1),$$

$j = \{c, \tau\}$, $k = \{HOM, HET\}$. The terms $\overline{\xi^j}$ and $\overline{\varsigma^j}$ are the cross-section averages of the mean and the variance of the individual KPSS statistic defined in Hadri (2000), $j = \{c, \tau\}$. Finally, it should be mentioned that the test in Carrion-i-Silvestre, del Barrio Castro and López-Bazo (2005) bases on Hadri's (2000) proposal but uses bootstrapped p-values following the lines given in Maddala and Wu (1999) to deal with the cross-section dependence among the time series in the panel. Note thus that these statistics do not account for the presence of cross-section dependence using a common factor model.

3.3.3 Panel data cointegration test statistics

Carrion-i-Silvestre and Surdeanu (2011) panel data cointegration test statistics

The first category of testing for cointegration in panel data is based on a system-based approach. As mentioned above, the main advantage of the system-based approach is to assess how many cointegrating relation exist among the variables for each individual system. Let us define the vector $X_{i,t} = (y_{i,t}, k_{i,t}, g_{i,t}, h_{i,t}, l_{i,t})'$ that collects the observable variables of our model, for which we define the following VAR representation:

$$\begin{aligned} X_{i,t} &= D_{i,t} + \lambda_i F_t + e_{i,t} && (3.3.8)\\ \left(I_q - L\right) F_t &= C(L) w_t && (3.3.9)\\ (I_k - L) e_{i,t} &= H_i(L) \varepsilon_{i,t}, && (3.3.10) \end{aligned}$$

where $i = 1, \ldots, N$ and $t = 1, \ldots, T$. In this setup $D_{i,t}$ is defined as a $(k \times 1)$-vector that contains the deterministic component of each of the variables in the vector $X_{i,t}$, i.e., $k = 5$ in our case. The term F_t is a $(r \times 1)$-vector of common factors, λ_i is a $(k \times r)$ matrix of factor loadings and $e_{i,t}$ is a $(k \times 1)$-vector that collects the idiosyncratic stochastic term. The estimation of the unobservable common factors is made using the principal component approach suggested in Bai and Ng (2002, 2004). Once the effects of the common factors are removed, cointegration analysis is then performed focusing on both the idiosyncratic and common factor components. This gives us further insight on the cointegration analysis, since the inference on the cointegrating rank can be distorted if common factors are not accounted for in the model – see Chapter 2 and Carrion-i-Silvestre and Surdeanu (2011) for further details.

The determination of the number of stochastic trends in the system relies on a sequential testing procedure that starts assuming that the cointegrating rank is zero – i.e., there are $m = k$ stochastic trends – and, defining the multivariate MSB test statistic $MSB_{j,i}(m)$, $j = \{c, \tau\}$, we can proceed to test whether there are $m = k$ stochastic trends or fewer than k trends. Using the $MSB_{j,i}(m)$ test statistic, $j = \{c, \tau\}$, we can estimate the number of stochastic trends for each individual system using the critical values in Chapter 2 and Carrion-i-Silvestre and Surdeanu (2011).

It is also possible to combine the individual information and define panel data cointegrating rank tests. Assuming the same number of stochastic trends m in all individual systems and that the idiosyncratic component is cross-section independent, it is possible to test the null hypothesis that all N individual systems have

$m_i = m$ stochastic trends against the alternative hypothesis that there are $m_i = m-1$ stochastic trends:

$$\begin{cases} H_0: & m_i = m \text{ stochastic trends} \quad \forall i = 1,\dots,N \\ H_1: & m_i = m-1 \text{ stochastic trends} \quad \forall i = 1,\dots,N. \end{cases} \tag{3.3.11}$$

The first panel data statistic is based on the standardized mean of the individual statistics:

$$PMSB_j^Z(m) = \frac{\sqrt{N}(\overline{MSB}_j(m) - E(MSB_j(m)))}{\sqrt{Var(MSB_j(m))}},$$

where $\overline{MSB}_j(m) = N^{-1}\sum_{i=1}^{N} MSB_{j,i}(m)$, and $E(MSB_j(m))$ and $Var(MSB_j(m))$ are the mean and the variance of the $MSB_j(m)$ statistic, $j = \{c,\tau\}$, given given in Carrion-i-Silvestre and Surdeanu (2011). Under the null hypothesis of m stochastic trends $PMSB_j^Z(m) \Rightarrow N(0,1)$. The remaining tests are based on the combination of the p-values (φ_i) of the individual MSB statistic:

$$\begin{aligned} PMSB_j^F(m) &= -2\sum_{i=1}^{N} \ln \varphi_i \\ PMSB_j^C(m) &= \frac{-2\sum_{i=1}^{N} \ln \varphi_i - 2N}{\sqrt{4N}}, \end{aligned}$$

where under the null hypothesis of m stochastic trends $PMSB_j^F(m) \Rightarrow \chi^2_{2N}$ and $PMSB_j^C(m) \Rightarrow N(0,1)$, $j = \{c,\tau\}$. In this chapter we use another test, $PMSB_j^{CZ}$, in order to test for cointegration among the cross-sections of the panel. The $PMSB_j^{CZ}$ statistic, $j = \{c,\tau\}$, originally proposed by Choi (2001), is based on the p-values of the individual MSB tests and has the following form:

$$PMSB_j^{CZ}(m) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Phi^{-1}(\hat{\varphi}_i),$$

where $\Phi(\cdot)$ denotes the standard Normal cumulative distribution function and $j = \{c,\tau\}$. Although Carrion-i-Silvestre and Surdeanu (2011) do not prove it, they conjecture that the limiting distribution of this statistic is also standard normal, a claim that is supported by their Monte Carlo simulations. Like the previous panel cointegration tests, the null hypothesis of $PMSB_j^{CZ}$, $j = \{c,\tau\}$, is that of no panel cointegration.

Westerlund (2008) panel data cointegration test statistics

The author proposes two panel cointegration tests of the null hypothesis of no cointegration. They allow for cross-sectional dependence that is modeled by a factor model as in Bai and Ng (2004). More exactly, the model is given by:

$$
\begin{aligned}
y_{i,t} &= \mu_i + x_{i,t}'\beta_i + \lambda_i' F_t + e_{i,t} \\
F_{j,t} &= \rho_j F_{j,t-1} + u_{j,t} \quad j = 1, \ldots, r \\
e_{i,t} &= \theta_i e_{i,t-1} + \varepsilon_{i,t}
\end{aligned}
$$

where F_t is a $(r \times 1)$-vector of I(0) stationary common factors, and λ_i is a $(r \times 1)$-vector of factor loadings. To test the null hypothesis of no cointegration is equivalent to running a unit root test on the following regression:

$$\hat{e}_{i,t} = \theta_i \hat{e}_{i,t-1} + \hat{\varepsilon}_{i,t}. \tag{3.3.12}$$

Westerlund (2008) proposes two panel cointegration tests based on the Durbin-Hausman principle applied to (3.3.12). The first statistic, DH_p, is the panel Durbin-Hausman statistic constructed by summing the N individual terms before multiplying them together. It is defined as follows:

$$DH_p = \hat{S}_N \left(\tilde{\theta} - \hat{\theta}\right)^2 \sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{i,t-1}^2, \tag{3.3.13}$$

where $\tilde{\theta}$ and $\hat{\theta}$ denote the pooled instrumental variables (IV) and the pooled OLS estimators and $\hat{S}_N$ is defined in Westerlund (2008). The null hypothesis $H_0 : \theta_i = 1$ $\forall i$ is tested against the alternative hypothesis $H_1 : \theta_i = \theta < 1$ $\forall i$. A rejection of the null hypothesis should indicate that all individuals in the panel are cointegrated.

The second test proposed, DH_g, is the group mean panel statistic and is formulated as:

$$DH_g = \sum_{i=1}^{N} \hat{S}_i \left(\tilde{\theta}_i - \hat{\theta}_i\right)^2 \sum_{t=2}^{T} \hat{e}_{i,t-1}^2. \tag{3.3.14}$$

The null hypothesis $H_0 : \theta_i = 1$ $\forall i$ is tested against the alternative hypothesis $H_1 : \theta_i < 1$ for at least some i. This implies that a rejection of the null should be taken as evidence in favor of cointegration for at least some units. For both tests, we reject the null hypothesis if the value of the test is greater than its respective critical value.

Banerjee and Carrion-i-Silvestre (2011) panel data cointegration test statistics

Banerjee and Carrion-i-Silvestre (2011) propose a panel cointegration test based on the common correlated effects (CCE) estimation approach developed by Pesaran (2006). The idea behind the CCE estimation is relatively simple. Since the cross-section dependence is sometimes caused by unobservable common factors, Pesaran (2006) uses cross-section averages of the dependent and the explanatory variables as proxies for the common factors. Banerjee and Carrion-i-Silvestre (2011) use the following model:

$$y_{i,t} = D_{i,t} + x_{i,t}'\beta + \bar{z}_t'\eta_i + e_{i,t},$$

where $\bar{z}_t' = (\bar{y}_t, \bar{x}_t')'$ is the vector of cross-section means of the dependent and explanatory variables. Following Pesaran (2006), Holly, Pesaran and Yamagata (2010) and Kapetanios, Pesaran and Yamagata (2011), Banerjee and Carrion-i-Silvestre (2011) use the pooled estimator:

$$\hat{\beta}_{CCEP} = \left(\sum_{i=1}^{N} x_i'\bar{M}x_i\right)^{-1} \left(\sum_{i=1}^{N} x_i'\bar{M}y_i\right),$$

where $x_i = (x_{i,1}, x_{i,2}, ..., x_{i,T})'$, $y_i = (y_{i,1}, y_{i,2}, ..., y_{i,T})'$ and the matrix $\bar{M}$ is defined in Holly, Pesaran and Yamagata (2010). In the next step, Banerjee and Carrion-i-Silvestre (2011) define the variable $\tilde{y}_{i,t} = y_{i,t} - x_{i,t}'\hat{\beta}_{CCEP}$ and then estimate the regression below using the OLS estimation procedure:

$$\tilde{y}_{i,t} = D_{i,t} + e_{i,t}. \tag{3.3.15}$$

Both individual ($t_{\hat{\alpha}_{i,0}}$) and panel cointegration ($CADFC_P$) test statistics are based on the OLS residuals $\hat{e}_{i,t}$ from (3.3.15). The individual cointegration test statistic is the pseudo t-ratio of the estimated parameter $\hat{\alpha}_{i,0}$ in the following regression:

$$\Delta\hat{e}_{i,t} = \alpha_{i,0}\hat{e}_{i,t-1} + \sum_{j=1}^{p} \alpha_{i,j}\Delta\hat{e}_{i,t-j} + \zeta_i\bar{\hat{e}}_{t-1} + \sum_{j=0}^{p} \theta_{i,j}\Delta\bar{\hat{e}}_{t-j} + \kappa_{i,t}.$$

Finally, the panel cointegration statistic is defined as:

$$CADFC_P = N^{-1}\sum_{i=1}^{N} t_{\hat{\alpha}_{i,0}}.$$

The critical values for both individual and panel cointegration test statistics are presented in Banerjee and Carrion-i-Silvestre (2011). The null hypothesis of no cointegration is rejected if the value of the corresponding test statistic is smaller than the critical value.

Banerjee and Carrion-i-Silvestre (2013) panel data cointegration test statistics

The next three cointegration testing procedures are based on a single-equation framework. Banerjee and Carrion-i-Silvestre (2013) deal with the following model specification:

$$y_{i,t} = D_{i,t} + x_{i,t}'\beta_i + F_t'\lambda_i + e_{i,t}. \tag{3.3.16}$$

The common factors and factor loadings are estimated using principal components following the approach in Bai and Ng (2004). In order to do so, orthogonal projections on the first difference of Equation (3.3.16) are taken:

$$\begin{aligned} M_i\Delta y_i &= M_i\Delta F\lambda_i + M_i\Delta e_i \\ y_i^* &= f\lambda_i + z_i, \end{aligned} \tag{3.3.17}$$

with $M_i = I_{T-1} - \Delta x_i^d \left(\Delta x_i^{d\prime}\Delta x_i^d\right)^{-1}\Delta x_i^{d\prime}$ being the idempotent matrix, $\Delta x_i^d = [\Delta D_i\ \Delta x_i]$ a matrix that contains the first difference of the deterministic component and the stochastic regressors, $y_i^* = M_i\Delta y_i$, $f = M_i\Delta F$ and $z_i = M_i\Delta e_i$.[4] The estimation of the common factors and factor loadings is done as in Bai and Ng (2004) using principal components. Specifically, the estimated principal components of $f = (f_2, f_3, \ldots, f_T)$, denoted as $\tilde{f}$, are $\sqrt{T-1}$ times the r eigenvectors corresponding to the first r largest eigenvalues of the $(T-1)\times(T-1)$ matrix $y^*y^{*\prime}$. Under the normalization $\tilde{f}\tilde{f}'/(T-1) = I_r$, the estimated loading matrix is $\tilde{\Lambda} = \tilde{f}'y^*/(T-1)$ and the estimated residuals are defined as $\tilde{z}_{i,t} = y_{i,t}^* - \tilde{f}_t'\tilde{\lambda}_i$. Using these estimates, the idiosyncratic disturbance term is recovered and the common factors are computed through cumulation, i.e., $\tilde{e}_{i,t}^* = \sum_{j=2}^{t}\tilde{z}_{i,j}$ and $\tilde{F}_t = \sum_{j=2}^{t}\tilde{f}_j$. Then we proceed to the estimation of the ADF-type regression equation:

$$\Delta\tilde{e}_{i,t}^* = \alpha_{i,0}\tilde{e}_{i,t-1}^* + \sum_{j=1}^{p_i}\alpha_{i,j}\Delta\tilde{e}_{i,t-j}^* + w_{i,t}, \tag{3.3.18}$$

[4] It should be understood that $\Delta x_i^d = \Delta x_i$ for the models that do not include deterministic component or that include a constant term provided that in these cases $\Delta D_i = 0$.

so that the null hypothesis of no cointegration can be tested using the pseudo t-ratio of $\alpha_{i,0}$ ($t_{\tilde{\alpha}_{i,0}}$). Banerjee and Carrion-i-Silvestre (2013) define the panel cointegration test statistic $Z_j = (N^{-1}\sum_{i=1}^{N} t_{\tilde{\alpha}_{i,0}} - \Theta_j^e)(\Psi_j^e/N)^{-1/2}$, where $j = c$ refers to the model that includes a constant and $j = \tau$ to the model that includes a linear time trend, with Θ_j^e and Ψ_j^e the mean and variance of the relevant functionals of Brownian motions.[5] As $T, N \to \infty$, $N/T \to 0$, the Z_j test statistic converges in the limit to a standard normal distribution under the null hypothesis of no panel cointegration. If there is only one common factor, its order of integration can be tested using the ADF-type regression equation in (3.3.18) with $\tilde{e}_{i,t}^*$ replaced by $\tilde{F}_t$. For the case where more than one factor is estimated, the number of stochastic trends among the common factors can be estimated using the MQ test statistics as in Bai and Ng (2004).[6]

3.4 Empirical results

Throughout the section, the deterministic specification is given by a linear time trend. We start the empirical analysis by checking whether cross-section dependence exists among the variables of our model. Note that while it is convenient to think of cross-section independence as the ideal case, in real world this is not likely to hold in most situations. It should be natural to assume that the regions of Spain are dependent of each other. We employ the CD statistic of Pesaran (2004, 2013) and present the results of the statistic for each variable for different augmentation orders ($p = 0, 1, \ldots, 5$) in Table 3.1. The values of the CD test statistic indicate that we can easily reject the null hypothesis of cross-section independence in favor of cross-section dependence for all variables regardless of the augmentation order that is used.

As pointed out in Pesaran (2013), the large values that take the CD test can be an indication that strong dependence is present, which can be captured by the means of a common factor model.

[5]Banerjee and Carrion-i-Silvestre (2013) approximate the moments of the limiting distribution of the statistics by means of Monte Carlo simulation, which are $(\Theta_c^e, \Theta_\tau^e) = (-0.424, -1.535)$ and $(\Psi_c^e, \Psi_\tau^e) = (0.964, 0.341)$.

[6]The limiting distribution of the ADF test statistic when there is one common factor is the one obtained in Dickey and Fuller (1979), so that the standard critical values for the ADF test statistic can be used in this case. The critical values for the MQ test can be found in Table I in Bai and Ng (2004).

3.4.1 Panel data order of integration analysis

Let us first focus on the results obtained using Pesaran's (2007) statistics. The top of the Table 3.2 presents the CIPS(p) test statistic for different augmentation orders ($p = 0, 1, \ldots, 5$) – the truncated version of the test statistic reported the same results. The results indicate that, in almost all cases, the idiosyncratic component of the variables that we consider in the chapter is I(1) – the null hypothesis of unit root is marginally rejected for $y_{i,t}$ with $p = 0$ and $p = 1$ and for $k_{i,t}$ with $p = 1$.[7] In general, these results suggest that the idiosyncratic component of the variables in our model is I(1) non-stationary.

As mentioned above, two of the testing proposals that are applied in this chapter model the cross-section dependence through the specification of an approximate common factor model, for which the number of common factors needs to be estimated. The use of the panel BIC information criterion in Bai and Ng (2002) selects the maximum number of common factors permitted, which is six in our case. This seems a rather typical problem encountered by Sul (2005), Basher and Carrion-i-Silvestre (2007) and Holly, Pesaran and Yamagata (2010), among others. One reasonable explication, sustained by Bai and Ng (2002) as well, is that for small number of cross-sections (less than 20), the number of common factors is difficult to estimate. We then determine the number of I(1) non-stationary common factors (r_1) using the three information criteria (IPC_1, IPC_2 and IPC_3) proposed by Bai (2004) and the MQ test statistics by Bai and Ng (2004) while setting the maximum number of factors at six. The IPC_1 and IPC_2 criteria yield four I(1) non-stationary common factors, IPC_3 criteria suggests three I(1) non-stationary common factors, and the MQ test statistics indicate that all six factors are I(1) non-stationary. Therefore, for the rest of our analysis we calculate the statistics described previously for three, four or six factors in order to obtain robust conclusions.

Table 3.2 presents the panel data unit root test statistics in Moon and Perron (2004). Regardless of the number of factors considered, for the variables $g_{i,t}$ and $h_{i,t}$ the two statistics of Moon and Perron (2004) do not reject the null hypothesis of panel unit root at the 5% level of significance. For the rest of the variables, the results are somewhat mixed, depending on the number of factors taken into account – for more than half of the cases, the idiosyncratic component of these seven variables is I(1). However, we cannot conclude anything about the order

[7] The 5% critical value of the statistic for the case with intercept and time trend is -2.72 – see Table II(c) in Pesaran (2007).

of integration of the common factors from the application of these statistics since Moon and Perron (2004) test statistics focus only on the idiosyncratic component. Note that we wipe out the effect of the common factors, so that we are just focusing on the idiosyncratic disturbance terms – see Bai and Ng (2010). A more informative picture is thus obtained from Bai and Ng's (2004) approach, provided that separate inference can be conducted on the idiosyncratic and the common factor components of the observable variable.

Table 3.2 summarizes the results from the application of the approach in Bai and Ng (2004), reporting the ADF statistic for the idiosyncratic component of each variable and the MQ test statistics on the estimated common factors.[8] We first look at the results of the ADF statistic and we see that the null hypothesis of panel unit root cannot be rejected at the 5% level of significance for the idiosyncratic component of $y_{i,t}$, $h_{i,t}$ and $hs_{i,t}$ for any number of factors considered. The rest of the variables are I(1) for the case of three and four factors, and I(0) when six factors are taken into account. As for the common factors, the values of the $MQ_f^\tau(m)$ and $MQ_c^\tau(m)$ statistics characterize the six, four or three estimated common factors as stochastic trends – see the critical values for these statistics in Table I in Bai and Ng (2004). Therefore, we can infer that the seven observable panels of variables are I(1). When we consider three or four factors, the source of non-stationarity is of global and idiosyncratic nature for all seven variables. However, when we consider six factors, the source of non-stationarity comes from a global nature for the $k_{i,t}$, $g_{i,t}$, $l_{i,t}$ and $gp_{i,t}$ panels and of a global and idiosyncratic nature for the $y_{i,t}$, $h_{i,t}$ and $hs_{i,t}$ panels.

We complement the analysis of the stochastic properties through the computation of the panel stationary test of Hadri (2000) assuming that the long-run variance is either homogeneous or heterogeneous. The results of the panel data stationary tests of Hadri (2000) and bootstrapped critical values computed as in Carrion-i-Silvestre, del Barrio Castro and López-Bazo (2005) are presented in Table 3.3. It is easy to see that with the exception of the calculated value for $l_{i,t}$, all values of the stationarity test are greater than the 95% bootstrapped critical values. Therefore, we reject the null hypothesis of I(0) for six out of seven panels of variables at the 5% level of significance – this conclusion is reached regardless of the way in which the long-run variance is estimated. As noted in the previous section, while the stationarity test of Hadri (2000) allows for cross-section dependence when computing

[8]Following Perron and Ng (1998), the maximum number of lags that are used to compute the ADF statistic is set at $T^{1/3}$.

the empirical distribution using bootstrap, it does not accommodate for common factors.

To sum up, after analyzing the results from several types of panel data unit root and stationarity tests, we can infer that, in general, the panels of variables are I(1) and we can proceed with the panel cointegration analysis.

3.4.2 Testing for panel data cointegration

The model for the production function involves five observable variables that are driven by global and idiosyncratic stochastic trends. Since we consider two measures of public capital and human capital, we analyze four different combinations of variables. The first combination consists of $y_{i,t}$, $k_{i,t}$, $g_{i,t}$, $h_{i,t}$ and $l_{i,t}$ – hereafter, we denote this model specification as Combination 1. The variables that we test secondly are $y_{i,t}$, $k_{i,t}$, $gp_{i,t}$, $h_{i,t}$ and $l_{i,t}$ (Combination 2). The third combination consists of $y_{i,t}$, $k_{i,t}$, $g_{i,t}$, $hs_{i,t}$ and $l_{i,t}$ (Combination 3) and the last one consists of $y_{i,t}$, $k_{i,t}$, $gp_{i,t}$, $hs_{i,t}$ and $l_{i,t}$ (Combination 4). The presence of unit roots in these panels of variables implies that the estimation of the model that links these macroeconomic aggregates needs to restore on the use of panel cointegration analysis. Thus, we should test whether cointegration is present among these variables, accounting for the feature that global stochastic trends are present. This is addressed using both system-based and single-equation procedures.

System-based panel data cointegration analysis

The individual MSB based statistic and its respective cointegration rank for the first combination of variables are shown in the upper part of Table 3.4. The most common selected rank for the individual Spanish regions is one, suggesting the existence of one cointegrating relation among the variables of the model. For almost half of the regions, the univariate statistic detects no cointegration at all. For two regions, namely Andalucía and La Rioja, the rank is two indicating two cointegrating relations. Overall, the results are mixed indicating the low power of the univariate statistic.

The bottom part of Table 3.4 presents the $PMSB_{\tau}^{Z}$, $PMSB_{\tau}^{F}$ and $PMSB_{\tau}^{C}$ panel data statistics presented in Chapter 2 and in Carrion-i-Silvestre and Surdeanu (2011) for Combination 1. As mentioned before, the number of common factors is estimated using the panel BIC criterion and the number of estimated common factors equals the maximum factors allowed. Since the panel cointegrating test $PMSB_{\tau}^{CZ}$, also used in this chapter, is based on the same MSB statistic as the previous three panel statistics, we present its results in the same table with the Carrion-i-Silvestre and Surdeanu (2011) statistics. The panel cointegration ranks are reported in the last column of the bottom part of Table 3.4. All panel data statistics strongly reject the null hypothesis of no cointegration at the 5% level of significance. Moreover, with the exception of the $PMSB_{\tau}^{Z}$ test, all panel data cointegration test statistics indicate that the cointegration rank is two. This result implies the existence of two cointegrating relations between output, physical capital, human capital, public capital and labor. Table 3.5 presents both the univariate and panel data cointegration statistics for Combination 2. At the univariate level, cointegrating rank is 0 for seven regions, 1 for six regions and 2 for four regions. The panel data results are similar to the ones from Combination 1. Specifically, three panel data statistics detect two cointegration relations while only one statistic detects one cointegration relation. The results of the cointegration tests both for univariate and panel data for the third combination are presented in Table 3.6. The univariate statistic indicates the absence of any cointegrating relation for seven regions, one cointegrating relation for nine regions and two cointegrating relations for one region (Canarias). At the panel level, the results indicate the existence of one common cointegrating relation. Finally, the results from the univariate and panel data cointegration statistics for Combination 4 of variables are presented in Table 3.7. There are seven regions for which the univariate statistic does not detect any cointegration between the variables. The univariate cointegrating rank is one for nine regions and two for only one region (Galicia). The results from the panel cointegration tests indicate the existence of one cointegration relation between this combination of variables.

Overall, we can infer that the results from the individual MSB based statistic are mixed. Approximately half of the time the test statistic detects no cointegration at all and this can be due to the low power of the univariate test. However, the panel cointegration statistics indicate with overwhelming evidence that there exist at least one common cointegrating relation.

Single-equation-based panel data cointegration analysis

This section examines the results from the single-equation-based framework of Westerlund (2008) and Banerjee and Carrion-i-Silvestre (2011, 2013). Let us first focus on the Westerlund (2008) approach and we present the panel Durbin-Hausman cointegration test statistics of Westerlund (2008) and their respective p-values for each combination of variables in Table 3.8. We can easily see that the values of the DH_p indicate that the null hypothesis of no cointegration is strongly rejected at the 5% level. However, the opposite conclusion is found if we compute the DH_g panel data test statistic. This contradiction between these statistics might be due, first, to the different heterogeneity degree that is assumed for the parameter of interest – the DH_p imposes an homogeneous parameter for all units, whereas the DH_g considers heterogeneous parameters when computing the test statistic – and, second, to the assumption that the common factors have to be I(0). The later has been shown to be a problematic assumption, provided the evidence of I(1) common factors as mentioned above.

The results from the panel cointegration test statistic in Banerjee and Carrion-i-Silvestre (2011) appear in Table 3.9. At the individual level, we are able to reject the null hypothesis of no cointegration for only few regions – the results from the individual statistics are not shown in order to save space but they are available upon request. Therefore, at the individual level, for the majority of Spanish regions there is not enough evidence that the variables cointegrate, regardless of the combination of variables that is used. Let us now turn our attention to the panel statistic $CADFC_P$ presented for up to $p = 10$ lags. For $p = 0, 5, 6, 7, 8, 9$, the panel statistic detects no cointegration at any acceptable levels of significance, regardless of the combination of variables that is used. For $p = 1, 2, 3$, the $CADFC_P$ statistic is able to reject the null hypothesis of no cointegration at either the 5% level of significance for every combination. For $p = 4$ and $p = 10$, we obtain mixed results – the $CADFC_P$ test statistic finds evidence in favor of no cointegration for the first two combinations of variables and the opposite conclusion for the last two combinations of variables. Although these results might seem contradictory with the evidence provided by the Z_c test statistic of Banerjee and Carrion-i-Silvestre (2013), it is important to note that the $CADFC_P$ tends to show size distortion problems (under-rejection of the null hypothesis) when the common factors and the idiosyncratic component have different orders of integration – see Banerjee and Carrion-i-Silvestre (2011) for further details.

Next, we focus on the Banerjee and Carrion-i-Silvestre (2013) approach where the common factors are estimated using principal components. In this case, with the maximum number of common factors set at six, the panel BIC information criterion in Bai and Ng (2002) selects two, three or four common factors depending on the combination of variables that we consider. In all cases, the estimated common factors are characterized as I(1) stochastic processes – see Table 3.10. The panel ADF test statistic computed using the idiosyncratic disturbance terms (Z_c test statistic) leads to the rejection of the null hypothesis of spurious regression for all four combinations of variables. Therefore, we conclude that once the presence of common factors is accounted for, there exists a long-run relation among the variables of all four combinations considered. Note that these results imply that the observable economic variables of the model do not cointegrate alone, they take part of a cointegration relation that includes the presence of global stochastic trends. This result is in line with the theoretical arguments that claim that the TFP is an I(1) stochastic process.

When we compare the results of the single-equation-based and system-based cointegration analyses, we conclude that the results are similar at the individual level, provided that little evidence is found in favor of cointegration. However, if we focus on the panel data test statistics, we obtain mixed results. The evidence drawn by the test statistics in Westerlund (2008) depends on the test statistic that is used, although the analysis is conditional on the assumption that the common factors are I(0), a requirement that is not met in our case. In general, Banerjee and Carrion-i-Silvestre (2011) test statistic finds evidence of cointegration when including a small number of lags and no cointegration for larger number of lags. Finally, the test statistics proposed in Chapter 2 and in Carrion-i-Silvestre and Surdeanu (2011) and in Banerjee and Carrion-i-Silvestre (2013) are able to reject the null hypothesis of no cointegration with overwhelming evidence for every combination of variables that is used.

3.4.3 The estimation of the Cobb-Douglas production function

Once the presence of a long-run relation among the different combination of variables that we have considered has been established, we proceed to estimate the panel cointegration relation allowing for common factors. There are few theoretical proposals in the literature that fit our requirements. First, we apply the continuously-updated and fully-modified (CupFM) and the continuously-updated and bias-corrected (CupBC) estimators proposed in Bai, Kao and Ng (2009), which rely on the use of principal components to jointly estimate the cointegrating vector, the factor loadings and the common factors of the model specification. Both estimation procedures render consistent and efficient estimates of the cointegrating vector regardless of whether we have I(0) and/or I(1) common factors. Second, we also use the pooled CCE estimator in Pesaran (2006) that, as established in Kapetanios, Pesaran and Yamagata (2011), produces a consistent estimator of the cointegrating vector. In this case, the common factors are proxied by the use of cross-section averages of the variables of the model.

Although both approaches render consistent estimates of the parameters, the one of Bai, Kao and Ng (2009) uses an efficient estimation procedure, which takes into account the possibility that there might be endogenous regressors in the equation that is estimated. On the contrary, the estimator in Kapetanios, Pesaran and Yamagata (2011) assumes that the stochastic regressors are weakly exogenous, a situation that might not hold in our case. Taking into account this feature, it is worth pointing out that it is possible to conduct statistical inference on the parameters estimated by any of these procedures.[9]

Table 3.11 reports the estimation of the Cobb-Douglas production function in (3.2.6) for the different combinations and estimation procedures. The results for the CupFM estimator are presented in the first four columns. The next four columns show the results for the CupBC estimator while the last four columns of Table 3.11 report the CCEP estimator of Kapetanios, Pesaran and Yamagata (2011). For Combinations 1 and 2, the number of common factors is estimated according to the information criteria in Bai and Ng (2002), which gives two common factors. For Combinations 3 and 4, the estimated number of common factors is three.

The estimated coefficients represent the elasticity of output with respect to physical capital, public capital, human capital and labor. The coefficient of $k_{i,t}$ indicates that the elasticity of output with respect to the physical capital ranges from 0.206 (CupBC) to 0.402 (CCEP), lower than the values commonly found in the empirical literature for the Spanish regions.[10] The lower values for this parameter shows the risk in which practitioners can incur if common factors are not taken into account when estimating production functions, i.e., the possibility of obtaining biased and inconsistent estimates of the parameters.

The coefficient of $g_{i,t}$ indicates that the elasticity of output with respect to the total public capital ranges from 0.118 (CCEP) to 0.173 (CupFM). Similarly, the coefficient of $gp_{i,t}$ shows that the elasticity of output with respect to the productive public capital varies from 0.111 (CCEP) to 0.153 (CupFM).[11] These findings re-

[9]We thank Chihwa Kao, Takashi Yamagata and Mauro Costantini for providing the Gauss code.

[10]For example, the values obtained by Serrano (1996) range from 0.38 to 0.45, those obtained by Bajo-Rubio and Díaz-Roldán (2005) range from 0.59 to 0.68 while that obtained by Márquez, Ramajo and Hewings (2011) is 0.31. Note that specification of the variables, the model, the data and estimation techniques differ from one study to another.

[11]The estimate of Bajo-Rubio and Díaz-Roldán (2005) is 0.09, and the one by Márquez, Ramajo

veal that each of the two types of public capital exerts a positive effect on Spanish regional productivity.

As for the coefficients for the human capital, when we consider the share of the employed population with secondary and university education ($h_{i,t}$), the elasticity of output with respect to the human capital takes on values between 0.213 (CupFM) and 0.272 (CupBC). Similar results are obtained when we measure the human capital as the average years of schooling, with values that vary from 0.205 (CupFM) to 0.339 (CupBC).[12] It is interesting to note that, on one hand, the CupFM and CupBC estimators yield similar results and all the estimated parameters of the human capital are statistically significant. On the other hand, the CCEP estimator indicates that human capital does not contribute significantly to explaining the Spanish regional output, regardless of the combination of variables that is used. This result might be due to the fact that, as mentioned above, the CupFM and CupBC estimators are efficient estimators – i.e., they take into account the fact that some of the stochastic regressors in the cointegration relationship might be endogenous – whereas the CCEP estimator assumes weak exogeneity of the stochastic regressors. Therefore, in this regard the efficient estimators would be preferable provided that they cover broader situations.

Finally, we should highlight the negative and highly significant coefficient for $l_{i,t}$ that is obtained by the three panel cointegration estimators that are considered in this chapter, which indicates that the constant returns to scale assumption cannot be accepted. That is, the negative sign indicates diminishing returns to scale on the factors that have been considered in the model.

Overall, when we estimate the parameters of the Cobb-Douglas production function, the CupFM and CupBC estimators show that all variables contribute significantly to explaining the Spanish regional output. The CCEP estimator shows that only half the estimated parameters are statistically significant, although this conclusion bases on the assumption that the stochastic regressors are weakly exogenous, a feature that might not hold in our case.

---

and Hewings (2011) is 0.10.

[12]The results are in agreement with those of Serrano (1996), who obtained a value of 0.216, or Bajo-Rubio and Díaz-Roldán (2005), who obtained a value of 0.14.

3.4.4 Robustness analysis

Translog production function

In this section we estimate the parameters of the production function using the normalized translog representation. Our choice behind normalization is based on the work of Friedlaender and Spady (1981). They interpreted the translog as a second-order Taylor's series approximation to some true production function. Further, the authors implied that the translog will give a valid quadratic approximation of the true production function only at the sample mean. The translog representation has become increasingly popular in recent years – see López-Bazo and Moreno (2008) and Pablo-Romero and Gómez-Calero (2013). We normalize the data by dividing each series by its mean before the logarithmic transformation is applied – see Felipe (1998) and Sauer and Tchale (2006). To the best of our knowledge, none of the empirical studies studying the translog production function controlled for the cross-section dependence through the specification of a common factor model, as we do in this chapter.

The normalized translog production model has the following form:

$$\begin{aligned} y_{i,t} &= a_{i,t} + \beta_k k_{i,t} + \beta_g g_{i,t} + \beta_h h_{i,t} + \beta_l l_{i,t} + 0.5\beta_{kk} k_{i,t}^2 + 0.5\beta_{gg} g_{i,t}^2 + \\ &\quad + 0.5\beta_{hh} h_{i,t}^2 + 0.5\beta_{ll} l_{i,t}^2 + \beta_{kg} k_{i,t} g_{i,t} + \beta_{kh} k_{i,t} h_{i,t} + \beta_{kl} k_{i,t} l_{i,t} + \\ &\quad + \beta_{gh} g_{i,t} h_{i,t} + \beta_{gl} g_{i,t} l_{i,t} + \beta_{hl} h_{i,t} l_{i,t} \qquad (3.4.1) \\ a_{i,t} &= D_{i,t} + F_t' \lambda_i + e_{i,t}, \end{aligned}$$

where $y_{i,t} = \ln(\frac{Y_{i,t}/Y_{i,\bullet}}{L_{i,t}/L_{i,\bullet}})$, $a_{i,t} = \ln(A_{i,t}/L_{i,t})$, $l_{i,t} = \ln(L_{i,t}/L_{i,\bullet})$, $k_{i,t} = \ln(\frac{K_{i,t}/K_{i,\bullet}}{L_{i,t}/L_{i,\bullet}})$, $g_{i,t}$ $= \ln(\frac{G_{i,t}/G_{i,\bullet}}{L_{i,t}/L_{i,\bullet}})$ and $h_{i,t} = \ln(\frac{H_{i,t}/H_{i,\bullet}}{L_{i,t}/L_{i,\bullet}})$. The variables $Y_{i,\bullet}$, $K_{i,\bullet}$, $G_{i,\bullet}$, $H_{i,\bullet}$ and $L_{i,\bullet}$ are the mean values of $Y_{i,t}$, $K_{i,t}$, $G_{i,t}$, $H_{i,t}$ and $L_{i,t}$ respectively.

For Combinations 1, 2 and 4, the number of common factors estimated according to the information criteria in Bai and Ng (2002) is one. For Combination 3 we obtain three factors. Since the translog regression is linear in parameters we can apply standard panel data estimation techniques. Table 3.12 shows the CupFM and CupBC estimators for the translog specification. We can easily see that the results from these two estimators are similar. At the 1% level of significance, the first-order coefficients from the translog specification have the same signs and similar values with those from the Cobb-Douglas specification. The translog coefficients range from 0.300 to 0.534 for $k_{i,t}$, 0.145 to 0.157 for $g_{i,t}$, 0.081 to 0.085 for $gp_{i,t}$, 0.679 to

0.723 for $h_{i,t}$, 0.585 to 0.639 for $hs_{i,t}$ and -0.093 to -0.367 for $l_{i,t}$. At the 10% and even at 5% levels of significance, we obtain a few significant coefficients that have the opposite sign of what we obtain at 1% level of significance. For example, in the CupFM estimation, the coefficient of $g_{i,t}$ for Combination 1 (-0.279) or that of $gp_{i,t}$ for Combination 2 (-0.250) are negative and significant. The quadratic terms are only half the time significant. Growiec, Pajor, Pelle and Predki (2011) obtained the same results in their study on 19 OECD countries and state that "at this point we cannot infer if the departures from the Cobb-Douglas benchmark are economically important or not".

The interaction terms in the translog production function can be explained as follows: if the coefficient is positive then those two variables are complementary inputs and if the coefficient is negative then the two variables are considered substitutes. We see that the interaction terms between $k_{i,t}$ and $l_{i,t}$, $g_{i,t}$ and $hs_{i,t}$, $gp_{i,t}$ and $hs_{i,t}$, and $h_{i,t}$ and $l_{i,t}$ are positive indicating that these are complementary factors. The interaction terms between physical capital and both forms of human capital are negative and highly significant. Therefore, this indicates that these terms are substitutes. Evans, Green and Murinde (2002) obtained similar results and indicated that there is evidence against the embodiment and learning-by-doing hypotheses. The rest of the interaction coefficients are negative, suggesting a lack of complementarity.

Next, we compare the parameters estimates of the production function using both Cobb-Douglas and normalized translog representations. It is well known that the translog parameters do not have a direct interpretation compared with those from the Cobb-Douglas production function, which are interpreted as elasticities – see Felipe (1998). Therefore, in order to make the results comparable, we use the estimated translog regression coefficients to calculate the elasticity of production – which is computed by taking the partial derivative of the Equation (3.4.1) with respect to each explanatory variable. For example, for Combination 1, the partial elasticity of $y_{i,t}$ with respect to $k_{i,t}$ is calculated as $-0.199 + 2 * 0.5 * 0.043 * k_{i,t} + 0.347 * g_{i,t} - 0.393 * h_{i,t} + 0.551 * l_{i,t}$. We can easily see that the partial elasticity depends both on the information of the i-th individual and on t. Therefore, in order to obtain a summarized measure, we replace $k_{i,t}$, $g_{i,t}$, $h_{i,t}$ and $l_{i,t}$ by their (time and cross-section) means.[13] Table 3.13 presents the results of partial elasticities of output with respect to each variable for the normalized translog production function. After comparing the results from the Cobb-Douglas representation (Table 3.11) with

[13]The mean values of $k_{i,t}$, $g_{i,t}$, $gp_{i,t}$, $h_{i,t}$, $hs_{i,t}$ and $l_{i,t}$ are available upon request.

the results from the translog representation (Table 3.13), we see that they are similar with very few exceptions (negative sign for $hs_{i,t}$ in Combination 3 and positive sign for $l_{i,t}$ in Combinations 1 and 2). The rest of the estimated coefficients are very similar in both Cobb-Douglas and normalized translog production functions.

Spatial dependence

So far, we assumed that the cross-section dependence among the Spanish regions is captured through the specification of a model of unobserved common factors. Examples of such common shocks are oil price, stock market or technological shocks. The econometric techniques that have been applied have considered this form of strong dependence when estimating the parameters of the model in order to get consistent and efficient estimates. However, it is possible that the Spanish regions might be affected by the presence of weak dependence – for instance, it would be possible that one Spanish region is affected by its neighbors due, for instance, to the existence of spillover effects. Therefore, it makes sense to consider the tools developed by the spatial econometrics as a way to model the weak dependence that might be affecting the Spanish regions – see Chudik, Pesaran and Tosetti (2011) and Banerjee and Carrion-i-Silvestre (2011) for the discussion about the distinction between weak and strong dependence.

The spatial dependence in econometric studies is carried out by defining a weight matrix, W, which indicates whether any pair of regions share a common border. If region i and j share a common border, then $W(i,j) = 1$ and zero otherwise. The testing for spatial dependence is typically done by maximum likelihood technique or generalized method of moments (Pesaran and Tosetti (2011)).

We follow Holly, Pesaran and Yamagata (2010) and proceed to estimate the model in (3.2.6) using the pooled CCE estimator developed by Pesaran (2006). The residual term given by $u_{i,t} = F_t' \lambda_i + e_{i,t}$ is estimated as $\hat{u}_{i,t} = \tilde{F}_t' \tilde{\lambda}_i + \tilde{e}_{i,t}$, where $\tilde{F}_t$ denotes the $(r \times 1)$-vector of common factors, $\tilde{\lambda}_{i,j}$ the $(r \times 1)$-vector of factor loadings and $\tilde{e}_{i,t}$ the idiosyncratic component, which are obtained from the estimation of (3.2.6). As in Holly, Pesaran and Yamagata (2010), we perform the OLS regression of $\hat{u}_{i,t}$ on the estimated factors $\tilde{F}_t$ and obtain the idiosyncratic components $\tilde{e}^*_{i,t}$. For each idiosyncratic disturbance term we specify:

$$\tilde{e}^*_{i,t} = \Gamma \sum_{j=1}^{N} w_{i,j} \tilde{e}^*_{j,t} + v_{i,t},$$

where Γ is the spatial autoregressive parameter, $w_{i,j}$ is the (i, j) element of the spatial weight matrix W and $v_{i,t} \sim iid\left(0, \sigma_v^2\right)$. We then calculate the log likelihood function:

$$L = -\left(\frac{NT}{2}\right)\ln(\sigma_v^2) + T\ln|I_N - \Gamma W| - \frac{1}{2\sigma_v^2}\sum_{t=1}^{T}(\tilde{e}_t^* - \Gamma W\tilde{e}_t^*)'(\tilde{e}_t^* - \Gamma W\tilde{e}_t^*),$$

where $\tilde{e}_t^* = \left(\tilde{e}_{1,t}^*, \tilde{e}_{2,t}^*, ..., \tilde{e}_{N,t}^*\right)'$. Since Baleares and Canarias are islands, they have no neighbors and we eliminate their data for this analysis. Thus, we made the calculation considering $N = 15$ regions and $T = 37$ years. The maximum likelihood estimates of Γ are presented in Table 3.14. It is easy to see that the results are mixed and vary depending on the numbers of factors considered. When we consider two or three factors for all the combinations of variables or six factors for Combinations 1 and 2, the estimates are significant at the 5% level. This indicates that, even after controlling for the strong cross-section dependence, there exists spatial (weak) dependence among the Spanish regions. However and regardless of the presence of weak dependence, the estimation of the production function that has been conducted in the previous sections is consistent and efficient, since the procedures that have been applied allow for this type of cross-section dependence.

3.5 Conclusion

This chapter reexamines the evidence of cointegration among the output, physical capital, human capital, public capital, and labor. We consider annual data for seventeen Spanish regions observed over the period from 1964 to 2000.

The empirical analyses that focus on the estimation of Spanish production functions usually assume cross-section independence, which is a restrictive assumption especially at the regional level. Our empirical analysis shows that the variables involved in the model are I(1) non-stationary, so that the application of panel data cointegration techniques are required to obtain consistent estimates of the parameters of interest. The study takes advantage of the recently developed non-stationary panel data analysis methodologies, in both single-equation and system-equations based framework, that are general enough to permit cross-section dependence across the units of the panel.

The results reveal evidence of panel data cointegration among the variables of the model up to the presence of I(1) non-stationary common factors. Consequently, the observable economic variables alone do not generate an equilibrium relationship, we need to consider the, otherwise, expected global stochastic common trends that defines the TFP of the regions. The procedures applied in this chapter detect one or two cointegration relations among output, physical capital, human capital, public capital (all in per worker terms) and labor, depending on the combination of variables that is used.

We estimate the Spanish regional production function using Bai, Kao and Ng (2009) and Kapetanios, Pesaran and Yamagata (2011) panel data cointegration estimators. The results indicate that physical capital, human capital, public capital (all in per worker terms) affect positively the Spanish productivity, whereas the negative coefficient that has been obtained for the labor indicates the existence of decreasing returns to scale. We also show that the Spanish regions suffer from weak spatial dependence even after controlling for the strong cross-section dependence, although the conclusions that we have obtained are robust to the presence of this form of cross-section dependence and also to the functional form that is adopted for the production function.

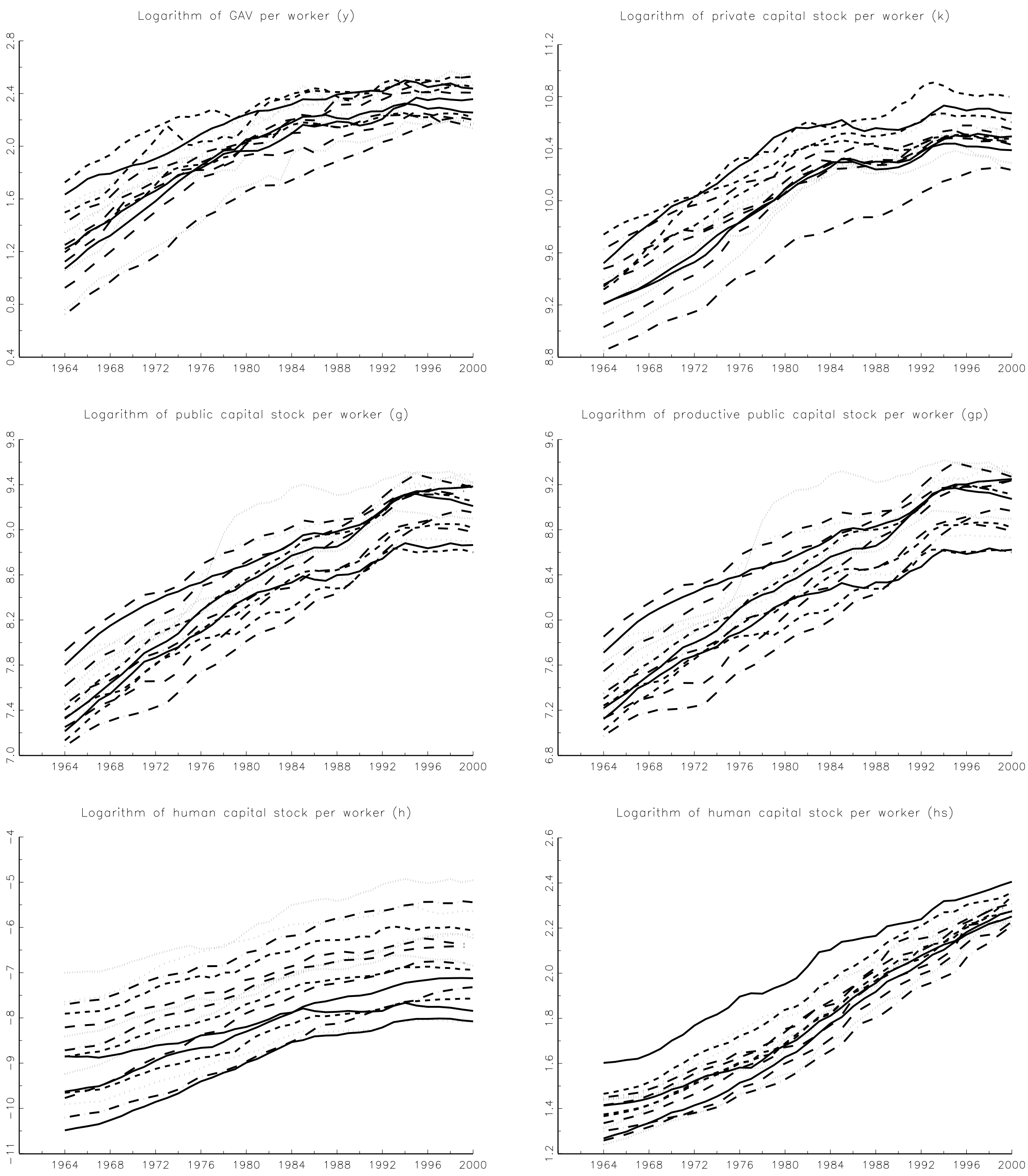

Figure 3.1: Time series variables of the seventeen Spanish regions

Table 3.1: Pesaran (2004, 2013) cross-section dependence tests

	$y_{i,t}$	$k_{i,t}$	$g_{i,t}$	$h_{i,t}$	$l_{i,t}$	$gp_{i,t}$	$hs_{i,t}$
CD(0)	8.505	29.191	27.233	21.061	36.083	26.488	16.978
CD(1)	9.205	19.378	19.309	18.386	21.965	17.727	17.661
CD(2)	8.291	18.669	17.501	18.359	21.724	16.010	18.370
CD(3)	7.984	19.232	16.918	18.278	22.240	15.741	15.687
CD(4)	8.307	18.403	16.121	17.858	21.629	15.019	15.075
CD(5)	7.592	18.361	16.776	17.354	21.528	15.632	15.243

$y_{i,t}$ is the logarithm of GVA per worker; $k_{i,t}$ denotes the logarithm of private capital per worker; $g_{i,t}$ and $gp_{i,t}$ are the logarithms of two forms of public capital (both in per worker terms): the former represents the total public capital while the latter represents the productive public capital; $h_{i,t}$ and $hs_{i,t}$ are the logarithms of two forms of human capital (both in per worker terms): the former represents the rate of employees with at least secondary school studies while the latter represents the average number of schooling years; $l_{i,t}$ denotes the logarithm of labor.

Table 3.2: Panel data unit root tests

		Pesaran (2007)						
		$y_{i,t}$	$k_{i,t}$	$g_{i,t}$	$h_{i,t}$	$l_{i,t}$	$gp_{i,t}$	$hs_{i,t}$
	CADF(0)	-2.730	-2.687	-1.934	-2.293	-2.677	-1.844	-2.122
	CADF(1)	-2.916	-2.817	-2.197	-2.553	-2.716	-2.065	-2.713
	CADF(2)	-2.522	-2.407	-1.922	-2.123	-2.222	-1.809	-2.276
	CADF(3)	-2.416	-2.306	-1.929	-2.065	-2.039	-1.846	-1.990
	CADF(4)	-2.254	-2.088	-2.097	-1.646	-1.817	-2.075	-1.641
	CADF(5)	-2.271	-2.170	-1.804	-1.614	-1.648	-1.805	-2.156
		Moon and Perron (2004)						
		$y_{i,t}$	$k_{i,t}$	$g_{i,t}$	$h_{i,t}$	$l_{i,t}$	$gp_{i,t}$	$hs_{i,t}$
6 factors	t_a	-1.267	-0.844	-0.146	-0.955	-1.414	-0.301	-4.707
	p-value	(0.103)	(0.199)	(0.442)	(0.170)	(0.079)	(0.382)	(0.000)
	t_b	-1.225	-0.743	-0.119	-0.818	-1.379	-0.253	-4.307
	p-value	(0.110)	(0.229)	(0.453)	(0.207)	(0.084)	(0.400)	(0.000)
4 factors	t_a	-1.621	-0.397	-0.793	-0.928	-1.719	-0.515	-2.470
	p-value	(0.052)	(0.346)	(0.214)	(0.177)	(0.043)	(0.303)	(0.007)
	t_b	-1.604	-0.354	-0.771	-0.837	-2.006	-0.478	-1.762
	p-value	(0.054)	(0.362)	(0.220)	(0.201)	(0.022)	(0.316)	(0.039)
3 factors	t_a	-1.761	-1.985	-1.073	-1.176	-2.015	-1.531	-1.539
	p-value	(0.039)	(0.024)	(0.142)	(0.120)	(0.022)	(0.063)	(0.062)
	t_b	-1.610	-1.886	-1.212	-1.082	-2.511	-1.759	-1.220
	p-value	(0.054)	(0.030)	(0.113)	(0.140)	(0.006)	(0.039)	(0.111)
		Bai and Ng (2004)						
		$y_{i,t}$	$k_{i,t}$	$g_{i,t}$	$h_{i,t}$	$l_{i,t}$	$gp_{i,t}$	$hs_{i,t}$
6 factors	$ADF_{\hat{e}}^{\tau}$	0.410	-1.770	-3.513	-0.219	-1.774	-2.707	2.210
	p-value	(0.659)	(0.038)	(0.000)	(0.413)	(0.038)	(0.003)	(0.986)
	$MQ_f^{\tau}(6)$	-31.907	-30.099	-27.514	-34.730	-30.349	-26.869	-31.790
	$MQ_c^{\tau}(6)$	-35.406	-27.884	-23.257	-33.837	-25.857	-24.339	-35.729
4 factors	$ADF_{\hat{e}}^{\tau}$	-0.733	-0.540	-0.026	-0.593	1.466	-0.313	0.199
	p-value	(0.232)	(0.294)	(0.490)	(0.276)	(0.929)	(0.377)	(0.579)
	$MQ_f^{\tau}(4)$	-27.996	-27.129	-13.887	-21.437	-24.762	-13.888	-5.911
	$MQ_c^{\tau}(4)$	-22.761	-26.854	-20.255	-20.019	-20.248	-19.498	-7.870
3 factors	$ADF_{\hat{e}}^{\tau}$	0.360	-1.221	0.544	-0.373	1.855	1.231	0.199
	p-value	(0.641)	(0.111)	(0.707)	(0.355)	(0.968)	(0.891)	(0.579)
	$MQ_f^{\tau}(3)$	-17.847	-24.356	-9.758	-20.501	-23.963	-8.660	-5.911
	$MQ_c^{\tau}(3)$	-15.914	-26.647	-13.875	-19.851	-18.138	-13.758	-7.870

See Table 3.1 for the description of variables.

Table 3.3: Hadri (2000) panel stationarity tests

Variable	Long-run variance	Z_j^{τ} Statistic	Bootstrapped critical values 90%	95%	99%
$y_{i,t}$	Homogeneous	17.359	6.894	8.911	12.403
	Heterogeneous	26.744	8.590	11.034	17.607
$k_{i,t}$	Homogeneous	38.494	7.444	9.836	15.575
	Heterogeneous	44.574	9.428	12.136	20.319
$g_{i,t}$	Homogeneous	21.396	5.737	7.207	11.470
	Heterogeneous	21.762	7.102	8.738	13.455
$h_{i,t}$	Homogeneous	11.421	6.895	8.791	13.027
	Heterogeneous	19.596	8.606	11.441	17.954
$l_{i,t}$	Homogeneous	5.456	5.890	7.363	10.961
	Heterogeneous	6.868	7.830	9.678	14.177
$gp_{i,t}$	Heterogeneous	17.534	5.520	6.998	10.596
	Heterogeneous	13.752	7.051	8.833	12.781
$hs_{i,t}$	Heterogeneous	9.220	6.019	7.663	11.580
	Heterogeneous	12.081	7.617	9.732	15.976

See Table 3.1 for the description of variables.

Table 3.4: Individual and panel data cointegration tests of Carrion-i-Silvestre and Surdeanu (2011). Results for the Combination 1

	Individual statistic					
Region	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
Andalucía	0.017**	0.021**	0.070	0.081	0.258	2
Aragón	0.029	0.036	0.048	0.060	0.257	0
Asturias	0.025**	0.034	0.037	0.038	0.162	1
Baleares	0.028	0.033	0.034	0.047	0.150	0
Canarias	0.036	0.058	0.070	0.111	0.120	0
Cantabria	0.026**	0.029	0.035	0.050	0.122	1
Castilla y León	0.024**	0.033	0.059	0.091	0.062	1
Castilla-La Mancha	0.016**	0.033	0.068	0.073	0.107	1
Catalunya	0.030	0.038	0.045	0.132	0.105	0
Comunidad Valenciana	0.026	0.027	0.056	0.055	0.198	0
Extremadura	0.017**	0.033	0.100	0.099	0.207	1
Galicia	0.021**	0.041	0.051	0.086	0.306	1
Madrid	0.029	0.029	0.044	0.070	0.193	0
Murcia	0.026	0.026	0.062	0.073	0.087	0
Navarra	0.019**	0.033	0.047	0.104	0.166	1
País Vasco	0.025**	0.037	0.048	0.078	0.061	1
La Rioja	0.021**	0.027**	0.050	0.063	0.057	2

	Panel statistics					
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
$PMSB_{\tau}^{Z}$	-6.807**	-1.493	2.948	1.189	-0.611	1
$PMSB_{\tau}^{F}$	134.862**	61.071**	15.798	19.579	29.308	2
$PMSB_{\tau}^{C}$	12.231**	3.283**	-2.207	-1.749	-0.569	2
$PMSB_{\tau}^{CZ}$	-7.434**	-2.170**	2.840	1.701	0.016	2

m represents the number of stochastic trends. ** denotes that the test is significant at the 5% level. The variables consisting of Combination 1 are $y_{i,t}$, $k_{i,t}$, $g_{i,t}$, $h_{i,t}$ and $l_{i,t}$. See Table 3.1 for the description of variables.

Table 3.5: Individual and panel data cointegration tests of Carrion-i-Silvestre and Surdeanu (2011). Results for the Combination 2

	Individual statistic					
Region	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
Andalucía	0.019**	0.023**	0.071	0.071	0.103	2
Aragón	0.028	0.037	0.062	0.067	0.226	0
Asturias	0.027	0.033	0.039	0.039	0.112	0
Baleares	0.029	0.032	0.036	0.051	0.055	0
Canarias	0.030	0.067	0.079	0.116	0.122	0
Cantabria	0.024**	0.028	0.034	0.104	0.120	1
Castilla y León	0.026	0.044	0.061	0.089	0.055	0
Castilla-La Mancha	0.014**	0.031	0.070	0.087	0.090	1
Catalunya	0.029	0.038	0.058	0.095	0.089	0
Comunidad Valenciana	0.026**	0.027**	0.036	0.074	0.167	2
Extremadura	0.018**	0.034	0.115	0.124	0.159	1
Galicia	0.022**	0.037	0.052	0.113	0.167	1
Madrid	0.027	0.027	0.048	0.076	0.210	0
Murcia	0.025**	0.025**	0.072	0.072	0.095	2
Navarra	0.018**	0.037	0.082	0.088	0.163	1
País Vasco	0.025**	0.037	0.053	0.066	0.064	1
La Rioja	0.020**	0.025**	0.050	0.068	0.078	2

	Panel statistics					
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
$PMSB_{\tau}^{Z}$	-7.207**	-1.006	4.683	1.901	-1.624	1
$PMSB_{\tau}^{F}$	137.166**	70.158**	14.216	15.236	36.389	2
$PMSB_{\tau}^{C}$	12.511**	4.385**	-2.399	-2.275	0.290	2
$PMSB_{\tau}^{CZ}$	-7.889**	-2.290**	3.849	2.444	-1.107	2

m represents the number of stochastic trends. ** denotes that the test is significant at the 5% level. The variables consisting of Combination 2 are $y_{i,t}$, $k_{i,t}$, $gp_{i,t}$, $h_{i,t}$ and $l_{i,t}$. See Table 3.1 for the description of variables.

Table 3.6: Individual and panel data cointegration tests of Carrion-i-Silvestre and Surdeanu (2011). Results for the Combination 3

	Individual statistic					
Region	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
Andalucía	0.021**	0.030	0.030	0.320	0.108	1
Aragón	0.025**	0.031	0.052	0.082	0.101	1
Asturias	0.025**	0.029	0.036	0.075	0.295	1
Baleares	0.031	0.033	0.062	0.071	0.282	0
Canarias	0.025**	0.027**	0.070	0.100	0.221	2
Cantabria	0.023**	0.035	0.066	0.101	0.101	1
Castilla y León	0.028	0.032	0.086	0.086	0.082	0
Castilla-La Mancha	0.033	0.037	0.037	0.118	0.112	0
Catalunya	0.032	0.047	0.064	0.106	0.091	0
Comunidad Valenciana	0.024**	0.031	0.045	0.088	0.371	1
Extremadura	0.021**	0.028	0.115	0.173	0.229	1
Galicia	0.030	0.028	0.053	0.053	0.146	0
Madrid	0.031	0.044	0.051	0.082	0.202	0
Murcia	0.022**	0.030	0.062	0.062	0.161	1
Navarra	0.018**	0.036	0.082	0.106	0.504	1
País Vasco	0.034	0.036	0.037	0.073	0.179	0
La Rioja	0.019**	0.033	0.045	0.079	0.081	1
	Panel statistics					
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
$PMSB_{\tau}^{Z}$	-5.300**	-1.496	4.214	4.846	0.611	1
$PMSB_{\tau}^{F}$	120.704**	45.451	18.137	9.435	21.565	1
$PMSB_{\tau}^{C}$	10.514**	1.389	-1.924	-2.979	-1.508	1
$PMSB_{\tau}^{CZ}$	-6.125**	-1.605	3.325	3.973	1.193	1

m represents the number of stochastic trends. ** denotes that the test is significant at the 5% level. The variables consisting of Combination 3 are $y_{i,t}$, $k_{i,t}$, $g_{i,t}$, $hs_{i,t}$ and $l_{i,t}$. See Table 3.1 for the description of variables.

Table 3.7: Individual and panel data cointegration tests of Carrion-i-Silvestre and Surdeanu (2011). Results for the Combination 4

	Individual statistic					
Region	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
Andalucía	0.018**	0.031	0.032	0.370	0.095	1
Aragón	0.025**	0.033	0.059	0.087	0.087	1
Asturias	0.025**	0.029	0.038	0.069	0.303	1
Baleares	0.030	0.034	0.065	0.077	0.256	0
Canarias	0.026	0.028	0.075	0.104	0.235	0
Cantabria	0.021**	0.036	0.082	0.097	0.204	1
Castilla y León	0.027	0.031	0.086	0.087	0.082	0
Castilla-La Mancha	0.032	0.037	0.038	0.116	0.099	0
Catalunya	0.032	0.042	0.069	0.116	0.075	0
Comunidad Valenciana	0.024**	0.032	0.037	0.096	0.346	1
Extremadura	0.021**	0.029	0.117	0.189	0.221	1
Galicia	0.025**	0.028**	0.048	0.050	0.257	2
Madrid	0.041	0.044	0.051	0.098	0.203	0
Murcia	0.018**	0.037	0.043	0.064	0.155	1
Navarra	0.018**	0.036	0.080	0.126	0.536	1
País Vasco	0.033	0.035	0.035	0.070	0.212	0
La Rioja	0.018**	0.031	0.044	0.083	0.066	1
	Panel statistics					
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	Rank
$PMSB_{\tau}^{Z}$	-5.618**	-1.276	4.342	5.826	0.917	1
$PMSB_{\tau}^{F}$	127.145**	38.686	18.590	9.104	22.399	1
$PMSB_{\tau}^{C}$	11.295**	0.568	-1.869	-3.019	-1.407	1
$PMSB_{\tau}^{CZ}$	-6.326**	-1.096	3.275	4.292	1.325	1

m represents the number of stochastic trends. ** denotes that the test is significant at the 5% level. The variables consisting of Combination 4 are $y_{i,t}$, $k_{i,t}$, $gp_{i,t}$, $hs_{i,t}$ and $l_{i,t}$. See Table 3.1 for the description of variables.

Table 3.8: Panel data cointegration statistic of Westerlund (2008)

	Combination 1	Combination 2	Combination 3	Combination 4
DH_p	4.219	4.421	2.226	2.322
	(0.000)	(0.000)	(0.013)	(0.010)
DH_g	0.490	0.728	0.086	0.961
	(0.312)	(0.233)	(0.466)	(0.168)

The numbers in parentheses are the p-values. See Table 3.10 for the description of the Combinations 1 to 4.

Table 3.9: Panel data cointegration test statistic of Banerjee and Carrion-i-Silvestre (2011)

	$CADFC_P$			
p	Combination 1	Combination 2	Combination 3	Combination 4
0	-2.062	-2.057	-2.125	-2.121
1	-2.307*	-2.300*	-2.385**	-2.378**
2	-2.347**	-2.338**	-2.465**	-2.456**
3	-2.461**	-2.459**	-2.635**	-2.630**
4	-2.106	-2.122	-2.249*	-2.260*
5	-1.802	-1.823	-1.925	-1.940
6	-1.748	-1.768	-1.842	-1.857
7	-1.593	-1.621	-1.710	-1.735
8	-1.608	-1.650	-1.558	-1.583
9	-1.539	-1.621	-1.004	-1.025
10	-1.734	-1.804	-2.326**	-2.210*

** denotes that the test is significant at the 5% level and * denotes that the test is significant at the 10% level. p is the number of lags. See Table 3.10 for the description of the Combinations 1 to 4.

Table 3.10: Banerjee and Carrion-i-Silvestre (2013) panel data cointegration test statistic

	Combination 1					Combination 2			
	Test	$\hat{r}$	$\hat{r}_1^{NP}$	$\hat{r}_1^{P}$		Test	$\hat{r}$	$\hat{r}_1^{NP}$	$\hat{r}_1^{P}$
Z_c	-2.863	2	2	2	Z_c	-2.609	3	3	3
$MQ_c^c(2)$	-10.031				$MQ_c^c(3)$	-23.306			
$MQ_f^c(2)$	-8.447				$MQ_f^c(3)$	-24.115			
	Combination 3					Combination 4			
	Test	$\hat{r}$	$\hat{r}_1^{NP}$	$\hat{r}_1^{P}$		Test	$\hat{r}$	$\hat{r}_1^{NP}$	$\hat{r}_1^{P}$
Z_c	-2.578	4	4	4	Z_c	-2.797	3	3	3
$MQ_c^c(4)$	-21.995				$MQ_c^c(3)$	-19.798			
$MQ_f^c(4)$	-20.807				$MQ_f^c(3)$	-17.876			

The numbers in parentheses are the p-values. The dependent variable is $y_{i,t}$. The exogenous variables for the Combination 1 are $k_{i,t}$, $g_{i,t}$, $h_{i,t}$ and $l_{i,t}$. The exogenous variables for the Combination 2 are $k_{i,t}$, $gp_{i,t}$, $h_{i,t}$ and $l_{i,t}$. The exogenous variables for the Combination 3 are $k_{i,t}$, $g_{i,t}$, $hs_{i,t}$ and $l_{i,t}$. The exogenous variables for the Combination 4 are $k_{i,t}$, $gp_{i,t}$, $hs_{i,t}$ and $l_{i,t}$. See Table 3.1 for the description of variables.

Table 3.11: Estimates of the panel cointegrating vector using the Cobb-Douglas production function

	CupFM Combinations				CupBC Combinations				CCEP Combinations			
	1	2	3	4	1	2	3	4	1	2	3	4
$k_{i,t}$	0.223	0.241	0.340	0.365	0.206	0.219	0.279	0.299	0.257	0.248	0.402	0.388
	(7.312)	(8.095)	(14.534)	(16.199)	(6.479)	(7.205)	(12.077)	(13.157)	(1.600)	(1.553)	(2.111)	(2.187)
$g_{i,t}$	0.173	-	0.149	-	0.164	-	0.145	-	0.120	-	0.118	-
	(7.326)		(8.929)		(6.752)		(8.701)		(3.760)		(2.741)	
$gp_{i,t}$	-	0.153	-	0.130	-	0.144	-	0.126	-	0.111	-	0.112
		(7.185)		(8.593)		(6.524)		(8.280)		(4.103)		(4.298)
$h_{i,t}$	0.218	0.213	-	-	0.272	0.271	-	-	0.084	0.089	-	-
	(10.939)	(10.298)			(13.839)	(13.672)			(1.144)	(1.132)		
$hs_{i,t}$	-	-	0.205	0.211	-	-	0.339	0.333	-	-	0.039	0.046
			(6.742)	(6.853)			(11.238)	(10.940)			(0.249)	(0.283)
$l_{i,t}$	-0.123	-0.133	-0.208	-0.209	-0.068	-0.064	-0.234	-0.233	-0.564	-0.619	-0.542	-0.619
	(-4.614)	(-4.991)	(-6.422)	(-6.368)	(-2.523)	(-2.426)	(-8.104)	(-7.804)	(-2.564)	(-2.865)	(-2.512)	(-3.028)

The numbers in parentheses are the t-statistics. The dependent variable is $y_{i,t}$. See Table 3.10 for the description of the Combinations 1 to 4 and Table 3.1 for the description of variables. For the Combinations 1 and 2, the number of factors estimated according to Bai and Ng (2002) is two. For the Combinations 3 and 4, the number of estimated factors is three.

Table 3.12: Estimates of the panel cointegrating vector using the translog production function

	CupFM				CupBC			
	Combinations				Combinations			
	1	2	3	4	1	2	3	4
$k_{i,t}$	-0.199	-0.229	0.534	0.332	-0.278	-0.283	0.458	0.300
	(-1.317)	(-1.575)	(15.576)	(6.672)	(-1.846)	(-1.962)	(13.756)	(5.971)
$g_{i,t}$	-0.279	-	0.157	-	-0.193	-	0.145	-
	(-2.004)		(7.942)		(-1.379)		(7.847)	
$gp_{i,t}$	-	-0.250	-	0.085	-	-0.166	-	0.081
		(-1.961)		(2.738)		(-1.299)		(2.586)
$h_{i,t}$	0.723	0.704	-	-	0.706	0.679	-	-
	(6.134)	(5.891)			(6.037)	(5.736)		
$hs_{i,t}$	-	-	-0.081	0.585	-	-	-0.037	0.639
			(-1.781)	(9.620)			(-0.794)	(10.158)
$l_{i,t}$	0.316	0.306	-0.093	-0.360	0.257	0.264	-0.063	-0.367
	(1.891)	(1.883)	(-2.765)	(-7.511)	(1.536)	(1.629)	(-1.875)	(-7.490)
$k_{i,t}^2$	0.043	0.279	0.777	-0.157	0.289	0.499	0.983	-0.359
	(0.154)	(1.099)	(3.652)	(-0.605)	(1.057)	(2.029)	(4.870)	(-1.406)
$g_{i,t}^2$	0.050	-	-0.316	-	0.018	-	0.085	-
	(0.233)		(-1.989)		(0.084)		(0.526)	
$gp_{i,t}^2$	-	0.087	-	-0.402	-	0.061	-	-0.528
		(0.448)		(-1.588)		(0.314)		(-2.076)
$h_{i,t}^2$	0.376	0.346	-	-	0.362	0.328	-	-
	(4.078)	(3.679)			(3.980)	(3.545)		
$hs_{i,t}^2$	-	-	-0.580	-1.477	-	-	0.704	-2.307
			(-1.320)	(-2.054)			(1.673)	(-3.122)
$l_{i,t}^2$	0.234	0.296	1.508	1.509	0.038	0.073	1.102	1.689
	(0.524)	(0.673)	(4.389)	(2.829)	(0.084)	(0.167)	(3.402)	(3.158)
$k_{i,t} * g_{i,t}$	0.347	-	-0.320	-	0.278	-	-0.421	-
	(1.647)		(-1.872)		(1.366)		(-2.512)	
$k_{i,t} * gp_{i,t}$	-	0.212	-	-0.045	-	0.142	-	-0.094
		(1.112)		(-0.213)		(0.770)		(-0.449)
$k_{i,t} * h_{i,t}$	-0.393	-0.433	-	-	-0.448	-0.465	-	-
	(-3.302)	(-3.734)			(-3.794)	(-4.077)		
$k_{i,t} * hs_{i,t}$	-	-	-0.622	-0.319	-	-	-0.572	0.097
			(-2.874)	(-0.896)			(-2.691)	(0.278)
$k_{i,t} * l_{i,t}$	0.551	0.571	1.361	1.489	0.402	0.380	0.807	1.456
	(1.790)	(2.004)	(6.381)	(6.258)	(1.313)	(1.338)	(3.765)	(6.102)
$g_{i,t} * h_{i,t}$	-0.315	-	-	-	-0.253	-	-	-
	(-2.948)				(-2.364)			
$g_{i,t} * hs_{i,t}$	-	-	1.104	-	-	-	0.485	-
			(6.117)				(2.814)	
$g_{i,t} * l_{i,t}$	-0.911	-	-0.239	-	-0.773	-	0.135	-
	(-3.705)		(-1.215)		(-3.138)		(0.697)	

Continues.

Table 3.12: Estimates of the panel cointegrating vector using the translog production function

	CupFM Combinations				CupBC Combinations			
	1	2	3	4	1	2	3	4
$gp_{i,t} * h_{i,t}$	-	-0.263	-	-	-	-0.207	-	-
		(-2.674)				(-2.095)		
$gp_{i,t} * hs_{i,t}$	-	-	-	0.503	-	-	-	0.727
				(1.423)				(2.034)
$gp_{i,t} * l_{i,t}$	-	-0.958	-	-1.175	-	-0.799	-	-0.923
		(-4.149)		(-4.109)		(-3.463)		(-3.238)
$h_{i,t} * l_{i,t}$	0.264	0.248	-	-	0.202	0.196	-	-
	(1.871)	(1.799)			(1.427)	(1.426)		
$hs_{i,t} * l_{i,t}$	-	-	-0.786	0.112	-	-	-1.260	-0.390
			(-2.734)	(0.234)			(-4.346)	(-0.813)

The numbers in parentheses are the t-statistics. The dependent variable is $y_{i,t}$. For the Combinations 1, 2 and 4 the number of factors estimated according to Bai and Ng (2002) is one. For the Combination 3 the number of estimated factors is three. See Table 3.1 for the description of variables.

Table 3.13: Elasticities of the panel cointegrating vector using the translog production function

	CupFM Combinations				CupBC Combinations			
	1	2	3	4	1	2	3	4
$k_{i,t}$	0.283	0.313	0.550	0.356	0.274	0.299	0.475	0.325
$g_{i,t}$	0.126	-	0.174	-	0.136	-	0.138	-
$gp_{i,t}$	-	0.089	-	0.126	-	0.103	-	0.131
$h_{i,t}$	0.279	0.294	-	-	0.276	0.289	-	-
$hs_{i,t}$	-	-	-0.168	0.604	-	-	-0.095	0.640
$l_{i,t}$	0.048	0.059	-0.121	-0.310	0.065	0.080	-0.084	-0.328

See Table 3.1 for the description of variables.

Table 3.14: Spatial MLE estimates

Nr. of factors	Combination 1	Combination 2	Combination 3	Combination 4
1	-0.010	0.005	-0.051	-0.045
2	0.167**	0.176**	0.136**	0.143**
3	0.185**	0.183**	0.142**	0.151**
4	0.004	0.003	-0.045	-0.048
5	-0.020	-0.022	-0.041	-0.044
6	-0.220**	-0.218**	-0.038	-0.037

** denotes significance at the 5% level. The dependent variable is $y_{i,t}$. See Table 3.10 for the description of the Combinations 1 to 4.

Chapter 4

Conclusions and proposed future research

This thesis consists of three self-contained chapters on non-stationary panel data analysis. All three essays concentrate on both cointegration and unit root analysis in panel data while allowing for the presence of cross-section dependence through the specification of an approximate common factor model.

In the first chapter we propose several panel data unit root tests that allow for multiple structural breaks and common factors to control for the presence of cross-section dependence. The test statistics are based on the use of GLS detrending procedure and the structural breaks are allowed under both the null and the alternative hypotheses. The model specification considers both known and unknowns breaks. This study derives the limiting distribution of the individual unit root test statistics for the idiosyncratic disturbance term and the common factors. Further, we also show that panel data unit root test statistics can be defined through the combination of the individual test statistics of the idiosyncratic component. The performance of the statistics that have been proposed is evaluated using a Monte Carlo simulation experiment. The simulations show that the test statistics perform well for the cases of known structural breaks. When the location of the structural breaks is not known *a priori* the panel statistics suffer for under-rejection when the time series dimension is large. Finally, we apply the proposed tests to a panel data set of annual real per capita GDP over the period 1870-2008 for 19 OECD developed countries. All panel statistics rejected the null hypothesis of panel data unit root in favor of I(0) stationarity for the idiosyncratic component of the real per capita GDP. However, all unit root tests for the common factor cannot reject the null hypothesis of unit root.

In the second chapter we propose a new test statistic to estimate the cointegrating rank both in a unit-by-unit analysis and in a panel data framework. This setup allows us to cover strong cross-section dependence cases, i.e., cases where the time series of one individual system are cointegrated with times series of other individual system (cross-cointegration), as well as cases where the factors appear by construction due to model specification. The performance of the proposed tests is investigated with Monte Carlo simulations. In general, the panel data based MSB statistic provides better estimation of the number of stochastic trends present in each individual system than the univariate one. Moreover, the simulations reveal that the existence of common factors can lead to misleading conclusions even if the analysis is carried out at a unit-by-unit basis. This is relevant from an empirical point of view considering that, in most cases, the cointegration analysis is conducted by focusing on one country whose economic system is related to that of other countries or ruled by international organizations such as in the case of the European Union. Therefore, the theoretical proposal presented in this chapter has also a significant empirical contribution. We then illustrate the application of the techniques to two popular empirical models: the money demand model and the monetary exchange model. The statistics applied to the money demand model, which consists of annual observations that covers the period 1957 to 1996 for nineteen countries, detect two stochastic trends. For the monetary exchange model, which consists of quarterly observations that covers the first quarter of 1973 up to the first quarter of 1997 for nineteen countries, the statistics detect three stochastic trends.

In the third chapter of this dissertation we present an empirical application. More exactly, we investigate the cointegrating relation between the output, physical capital, human capital, public capital, and labor. We consider annual data for seventeen Spanish regions observed over the period from 1964 to 2000. The empirical analysis shows that the variables involved in the model are I(1) non-stationary, so that the application of panel data cointegration techniques are required to obtain consistent estimates of the parameters of interest. The results reveal evidence of panel data cointegration among the variables of the model up to the presence of I(1) non-stationary common factors. Consequently, the observable economic variables alone do not generate an equilibrium relationship, we need to consider the, otherwise, expected global stochastic common trends that defines the TFP of the regions. The procedures applied in this chapter detect one or two cointegration relations among output, physical capital, human capital, public capital (all in per worker terms) and labor, depending on the combination of variables that is used. We then estimate the Spanish regional production function using Bai, Kao and Ng (2009) and Kapetanios, Pesaran and Yamagata (2011) panel data cointegration estimators. The results indicate that physical capital, human capital, public capital (all in per worker terms) affect positively the Spanish productivity, whereas the negative coefficient that has been obtained for the labor indicates the existence of decreasing returns to scale. We also show that the Spanish regions suffer from weak spatial dependence even after controlling for the strong cross-section dependence, although the conclusions that we have obtained are robust to the presence of this form of cross-section dependence and also to the functional form that is adopted for the production function.

The main contribution of this thesis is allowing for the presence of cross-section dependence through the specification of an approximate common factor model. We take advantage of the recently developed non-stationary panel data analysis methodologies, in both single-equation and system-equations based framework, that are general enough to permit cross-section dependence across the units of the panel.

Since the analysis of non-stationary panel data is still developing, we propose some directions for future research. For example, the proposed panel unit root tests based on the use of GLS detrending procedure in the first chapter can be extended to cover the case when the breaks are not common to all cross-sections. Although we allow mild heterogeneity of the break points across the units of the panel, in the limit we impose that the break points are common. Future research could focus on testing while relaxing this assumption. Also, it would be interesting in a future study to extend the cointegration panel tests proposed in the second chapter to cover the case of structural breaks, both known and unknown. In the second chapter, the statistics do not take into consideration the structural breaks. Therefore, this extension would make an original future work. And finally, we would like to explore the fractional integration in non-stationary panel data in the presence of structural breaks and cross-section dependence. So far we assume that a process can be I(0) stationary versus I(1) non-stationary but in the real world this might not be the case – see for example, real exchange rates, CPI or unemployment that show fractional values of integration. Therefore, we could develop more realistic panel data models that would help the empirical researchers make better economic decisions.

Appendix A

Mathematical Appendix for the First Chapter

A.1 Proof of Theorem 1.1

The proof focuses on Model II, since it is the most general model specification that we consider. The proof for the other models follows this one. The GLS detrended variable can be written as:

$$\begin{aligned}
\tilde{y}_i &= y_i - z_i\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) y_i^{\bar{\alpha}} \\
&= y_i - z_i\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right)\left(y_i - \bar{\alpha}_i y_{i,-1}\right) \\
&= y_i - A_i\left(y_i - \bar{\alpha} y_{i,-1}\right) \\
&= F\delta_i - A_i F^{\bar{\alpha}}\delta_i + e_i - A_i e_i^{\bar{\alpha}},
\end{aligned}$$

where $A_i = z_i\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right)$ and $D_T = diag\{D_{1,T}, D_{2,T}\} = diag(1, \ldots, 1, T^{-1/2}, \ldots, T^{-1/2})$. For subsequent developments, we define the partitioned vector of regressors $z_{i,t}\left(\lambda_i^0\right)$ as $z_{i,t}\left(\lambda_i^0\right) = \left(z'_{i,t,1}\left(\lambda_i^0\right), z'_{i,t,2}\left(\lambda_i^0\right)\right)'$. The term $z_{i,t,1}\left(\lambda_i^0\right)$ captures the $m+1$ regressors corresponding to the constant and the impulse dummy variables, while $z_{i,t,2}\left(\lambda_i^0\right)$ collects the $m+1$ trending regressors. Further, we also define $z_{i,t}^{\bar{\alpha}}\left(\lambda_i^0\right) = (z_{i,t,1}^{\bar{\alpha}\prime}\left(\lambda_i^0\right), z_{i,t,2}^{\bar{\alpha}\prime}\left(\lambda_i^0\right))'$ as the quasi-differenced $z_{i,t}\left(\lambda_i^0\right)$.

Taking the first difference we obtain the usual common factor representation

$$x_i = f\delta_i + \xi_i,$$

where $x_i = \Delta\tilde{y}_i$,
$f = \Delta F - \Delta z_i\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) F^{\bar{\alpha}} \delta_i$ and
$\xi_i = \Delta e_i - \Delta z_i\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}}$.

We can apply the method of principal components as in Bai and Ng (2004) and estimate f_t, δ_i and $\xi_{i,t}$ that are used to construct the unit root statistic which is based on the cumulative sum of the residuals

$$\tilde{e}_{i,t} = \sum_{j=2}^{t} \tilde{\xi}_{i,j}$$

Let us first focus on the idiosyncratic component $\xi_{i,t}$. Subtracting $x_{i,t} = \tilde{f}_t'\tilde{\delta}_i + \tilde{\xi}_{i,t}$ from $x_{i,t} = f_t'\delta_i + \xi_{i,t}$ yields

$$\tilde{\xi}_{i,t} = \xi_{i,t} + f_t'\delta_i - \tilde{f}_t'\tilde{\delta}_i.$$

Following Bai and Ng (2004) and Bai and Carrion-Silvestre (2009), we can rewrite this equation as

$$\begin{aligned}
\tilde{\xi}_{i,t} &= \xi_{i,t} + f_t' H H^{-1}\delta_i - \tilde{f}_t' H^{-1}\delta_i + \tilde{f}_t' H^{-1}\delta_i - \tilde{f}_t'\tilde{\delta}_i \\
&= \xi_{i,t} + \left(H' f_t - \tilde{f}_t\right)' H^{-1}\delta_i - \tilde{f}_t'\left(\tilde{\delta}_i - H^{-1}\delta_i\right) \\
&= \xi_{i,t} - \eta_t' H^{-1}\delta_i - \tilde{f}_t'\kappa_i, \qquad \text{(A.1.1)}
\end{aligned}$$

where $\eta_t = \left(\tilde{f}_t - H' f_t\right)$ and $\kappa_i = \left(\tilde{\delta}_i - H^{-1}\delta_i\right)$.

Using Lemma 3 and C1 from Bai and Ng (2004) and Theorem 2 from Bai and Carrion-i-Silvestre (2009), with the condition that $N, T \to \infty$, we obtain

$$\begin{aligned}
T^{-1/2}\left\|\sum_{j=2}^{t} \eta_j' H^{-1}\delta_i\right\| &= O_p\left(C_{NT}^{-1}\right) \\
T^{-1/2}\left\|\sum_{j=2}^{t} \tilde{f}_j \kappa_i\right\| &= O_p\left(C_{NT}^{-1}\right),
\end{aligned}$$

where $C_{NT} = \min\left[\sqrt{N}, \sqrt{T}\right]$.

We define the cumulative sum residuals as $\tilde{e}_{i,t} = \sum_{j=2}^{t} \tilde{\xi}_{i,j}$ and rewrite the previous equation as

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

Note that $\xi_{i,t} = \Delta e_{i,t} - \Delta z_{i,t}\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}}$. Thus,

$$\begin{aligned}
T^{-1/2}\sum_{j=2}^{t}\xi_{i,j} &= T^{-1/2}\sum_{j=2}^{t}\Delta e_{i,j} - T^{-1/2}\sum_{j=2}^{t}\Big[\Delta z_{i,t}\left(\lambda_i^0\right) \\
&\quad D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}}\Big] \\
&= T^{-1/2} e_{i,t} - T^{-1/2} e_{i,1} \\
&\quad - T^{-1/2} z_{i,t}\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}} \\
&\quad + T^{-1/2} z_{i,1}\left(\lambda_i^0\right) D_{i,T}\left(D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_i^{\bar{\alpha}}\left(\lambda_i^0\right) D_{i,T}\right)^{-1} D_{i,T} z_i^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}}.
\end{aligned}$$

The second terms is $o_p(1)$ since $T^{-1/2} e_{i,1} \xrightarrow{p} 0$. Further, $T^{-1/2} z_{i,t}\left(\lambda_i^0\right)\left(z_{i,1}^{\bar{\alpha}\prime}\left(\lambda_i^0\right) z_{i,1}^{\bar{\alpha}}\left(\lambda_i^0\right)\right)^{-1} z_{i,1}^{\bar{\alpha}\prime}\left(\lambda_i^0\right) e_i^{\bar{\alpha}} \xrightarrow{p} 0\,\forall t$ and $T^{-1/2} D_{i,T,2} z_{i,2}^{\bar{\alpha}}\left(\lambda_i^0\right) \to z_2(r)$ uniformly in $r \in [0,1]$, where $z_2(s) = \left(s, \left(s-\lambda_1^0\right) 1\left(s > \lambda_1^0\right), \ldots, \left(s-\lambda_m^0\right) 1\left(s > \lambda_m^0\right)\right)'$. Taking into account these elements, we can see that as $N, T \to \infty$ with $N/T \to 0$, we obtain

$$T^{-1/2}\sum_{j=2}^{t}\xi_{i,j} \Rightarrow \sigma_i\left[W_{i,c}(s) - z_2(s) A\left(\lambda^0\right)^{-1}\bar{V}_i\left(\lambda^0\right)\right],$$

where

$$A(\lambda^0) = \begin{bmatrix} a\left(\lambda_0^0,\lambda_0^0\right) & a\left(\lambda_0^0,\lambda_1^0\right) & \cdots & a\left(\lambda_0^0,\lambda_m^0\right) \\ & a\left(\lambda_1^0,\lambda_1^0\right) & \cdots & a\left(\lambda_1^0,\lambda_m^0\right) \\ & & \ddots & \vdots \\ & & & a\left(\lambda_m^0,\lambda_m^0\right) \end{bmatrix},$$

with elements defined as

$$a\left(\lambda_i^0,\lambda_j^0\right) = \frac{1}{6}\left(1-\lambda_j^0\right)\left[6\bar{c}\left(\lambda_i^0-1\right) + \bar{c}^2\left(\lambda_j^0\left(3\lambda_i^0-1\right) - 3\lambda_i^0 - \lambda_j^{0^2} + 2\right) + 6\right],$$

with $\lambda_i^0 < \lambda_j^0$, $\forall i,j = 0,1,\ldots,m$, where $\lambda_0^0 = 0$. Finally, $\bar{V}_i\left(\lambda^0\right) = \left(V_i\left(\lambda_0^0\right),\ldots,V_i\left(\lambda_m^0\right)\right)'$ with

$$\begin{aligned}
V_i\left(\lambda_j^0\right) &= \left(1+\lambda_j^0\bar{c}\right)\left(\left(W_i(1) - W_i(\lambda_j^0)\right) + (c-\bar{c})\int_{\lambda_j^0}^{1} W_{i,c}(s)\,ds\right) \\
&\quad - \bar{c}\int_{\lambda_j^0}^{1} s\,dW_i(s) - (c-\bar{c})\,\bar{c}\int_{\lambda_j^0}^{1} sW_{i,c}(s)\,ds.
\end{aligned}$$

Consequently,

$$\begin{aligned} T^{-1/2}\tilde{e}_{i,t} &\Rightarrow \sigma_i \left[W_{i,c}(s) - z_2(s) A\left(\lambda^0\right)^{-1} \bar{V}_i\left(\lambda^0\right)\right] \\ &\equiv \sigma_i V_{i,c,\bar{c}}\left(s,\lambda^0\right), \end{aligned}$$

and by the Functional Central Limit Theorem (FCLT) we have that

$$MSB_i^{GLS} \Rightarrow \left(\int_0^1 V_{i,c,\bar{c}}\left(s,\lambda^0\right)^2 ds\right)^{1/2}.$$

The limit distribution of the other two test statistics follows easily from the developments above. Further, note that the limiting distribution for Model I is the same as the ones derived for Model II since the impulse dummies – i.e., the elements collected in $z_{i,1}^{\bar{\alpha}}$ – are asymptotically negligible, as shown above.

Let us now focus on the estimated common factors $\tilde{F}_t$. From $\tilde{f}_t = Hf_t + v_t$ we have in the limit

$$\begin{aligned} \tilde{F}_t &= H\sum_{s=2}^{t} f_s + \sum_{s=2}^{t} v_s \\ &= H\sum_{s=2}^{t}\left(\Delta F_s - \Delta z_s\left(\lambda^0\right) D_T\left(D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right) z^{\bar{\alpha}}\left(\lambda^0\right) D_T\right)^{-1}\right. \qquad \text{(A.1.2)} \\ &\quad \left. D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right) F^{\bar{\alpha}}\bar{\delta}\right) + \sum_{s=2}^{t} v_s \\ &= H\left[F_t - F_1 - z_t\left(\lambda^0\right) D_T\left(D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right) z^{\bar{\alpha}}\left(\lambda^0\right) D_T\right)^{-1} D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right) F^{\bar{\alpha}}\bar{\delta}\right. \\ &\quad \left. + z_1\left(\lambda^0\right) D_T\left(D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right) z^{\bar{\alpha}}\left(\lambda^0\right) D_T\right)^{-1} D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right) F^{\bar{\alpha}}\bar{\delta}\right] \\ &\quad + V_t, \end{aligned}$$

where we have defined $z_t\left(\lambda^0\right)$ as the vector of regressors defined by the use of $T_j^0 = E\left(T_{i,j}^0\right) = N^{-1}\sum_{i=1}^{N} T_{i,j}^0$, $j = 0, 1, \ldots, m$. Further, note that in (A.1.2) we have

$\bar{\delta}$, provided that it holds that

$$
\begin{aligned}
N^{-1}\sum_{i=1}^{N}\tilde{F}_t &= N^{-1}\sum_{i=1}^{N}\Bigg[H\sum_{s=2}^{t}\left(\Delta F_s-\Delta z_s\left(\lambda^0\right)\right. \\
&\qquad \left. D_T\left(D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right)z^{\bar{\alpha}}\left(\lambda^0\right)D_T\right)^{-1}D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right)F^{\bar{\alpha}}\delta_i\right)+\sum_{s=2}^{t}v_s\Bigg] \\
\tilde{F}_t &= H\sum_{s=2}^{t}\left(\Delta F_s-\Delta z_s\left(\lambda^0\right)D_T\left(D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right)z^{\bar{\alpha}}\left(\lambda^0\right)D_T\right)^{-1}D_T z^{\bar{\alpha}\prime}\left(\lambda^0\right)F^{\bar{\alpha}}\bar{\delta}\right) \\
&\quad +\sum_{s=2}^{t}v_s.
\end{aligned}
$$

Then, we can define $\tilde{F}_t^d = HF_t^d + V_t^d$, where the d superscript denotes that the variable has been detrended. In our case, the detrending is based on the use of the quasi-GLS procedure described above, so that we define $\tilde{F}_t^d = \tilde{F}_t - z_t'\left(\lambda^0\right)\hat{\psi}$ where $z_t\left(\lambda^0\right)$ is the vector of regressors defined by the use of $T_j^0 = E\left(T_{i,j}^0\right) = N^{-1}\sum_{i=1}^{N}T_{i,j}^0$ and $\hat{\psi}$ is the matrix of parameters that has been obtained using the GLS detrending procedure. Using these elements and the developments above, we have

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

and the proof of the limiting distribution of the M-type unit root test statistics for the case of just one common factor $(r=1)$ follows the one for the idiosyncratic component replacing $e_{i,t}$ by $\tilde{F}_t^d$.

When we have more than one common factor $(r>1)$ we can proceed to apply the MQ tests in Bai and Ng (2004) in order to know how many I(1) and I(0) common factors we have. In this case, the limiting distribution of the MQ test statistics is given in Bai and Ng (2004) but using GLS detrended-based Brownian motions functionals instead of OLS-detrended Brownian motions ones.

Appendix B

Mathematical Appendix for the Second Chapter

B.1 Proof of Theorem 2.1

B.1.1 The intercept case

Let us first concentrate on the part of the proof that deals with the idiosyncratic component. Note that the estimated difference of the idiosyncratic stochastic term is:

$$\hat{z}_{i,t} = z_{i,t} + \lambda_i f_t - \hat{\lambda}_i \hat{f}_t.$$

Following Bai and Ng (2004), we can express the model as

$$\begin{aligned} \hat{z}_{i,t} &= z_{i,t} + \lambda_i H^{-1} H f_t - \lambda_i H^{-1} \hat{f}_t + \lambda_i H^{-1} \hat{f}_t - \hat{\lambda}_i \hat{f}_t \\ &= z_{i,t} + \lambda_i H^{-1} \left(H f_t - \hat{f}_t \right) - \left(\hat{\lambda}_i - \lambda_i H^{-1} \right) \hat{f}_t \\ &= z_{i,t} + \lambda_i H^{-1} v_t - d_i \hat{f}_t, \end{aligned} \tag{B.1.1}$$

where $v_t = \left(H f_t - \hat{f}_t \right)$ and $d_i = \left(\hat{\lambda}_i - \lambda_i H^{-1} \right)$. Let us define the partial sum process using the estimated residuals as $\hat{e}_{i,t} = \sum_{j=2}^{t} \hat{z}_{i,j} = \sum_{j=2}^{t} \left([M\ \Delta \hat{e}_i]_j \right)'$, where $[\cdot]_j$ denotes the j-th row of the matrix between brackets. The idiosyncratic disturbance terms can be expressed as $\hat{e}_{i,t} = e_{i,t} + A_{i,t}$ with $A_{i,t} = -e_{i,1} + \lambda_i H^{-1} V_t - d_i \hat{F}_t$. Denote by $\hat{e}_{i,t}(l)$ the l-th element of the $(k \times 1)$-vector $\hat{e}_{i,t}$, $l = 1, \ldots, k$. If $\hat{e}_{i,t}(l) \sim I(0)$ then

$T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^{2}(l)=O_p\left(T^{-1}\right)$, whereas if $\hat{e}_{i,t}(l)\sim I(1)$ then we have

$$\begin{aligned}T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^{2}(l) &= T^{-2}\sum_{t=2}^{T}e_{i,t}^{2}(l)+2T^{-2}\sum_{t=2}^{T}e_{i,t}(l)A_{i,t}(l)+T^{-2}\sum_{t=2}^{T}A_{i,t}^{2}(l)\\ &= I+II+III.\end{aligned}$$

Part I is $O_p(1)$ and, from Bai and Ng (2004), III is $O_p\left(C_{NT}^{-2}\right)$ for all i, with $C_{NT}=\min\left[\sqrt{N},\sqrt{T}\right]$. Let us now focus on II:

$$\begin{aligned}T^{-2}\sum_{t=2}^{T}e_{i,t}(l)A_{i,t}(l) &= -T^{-2}\sum_{t=2}^{T}e_{i,t}(l)e_{i,1}(l)+T^{-2}\sum_{t=2}^{T}e_{i,t}(l)\lambda_i(l)H^{-1}V_t\\ &\quad -T^{-2}\sum_{t=2}^{T}e_{i,t}(l)d_i(l)\hat{F}_t\\ &= a+b+c,\end{aligned}$$

where $\lambda_i(l)$ and $d_i(l)$ denote the l-th row of the $(k\times q)$ matrices λ_i and d_i. Element a is $O_p\left(T^{-1/2}\right)$, while b is

$$\begin{aligned}\|b\| &\leq \left(T^{-2}\sum_{t=2}^{T}\|e_{i,t}(l)\lambda_i(l)\|^2\right)^{1/2}\left(T^{-2}\sum_{t=2}^{T}\left\|H^{-1}V_t\right\|^2\right)^{1/2}\\ &= O_p(1)O_p\left(N^{-1/2}\right)=O_p\left(C_{NT}^{-1}\right).\end{aligned}$$

Finally, element c

$$\begin{aligned}\|c\| &\leq \left(T^{-2}\sum_{t=2}^{T}\|e_{i,t}(l)d_i(l)\|^2\right)^{1/2}\left(T^{-2}\sum_{t=2}^{T}\left\|\hat{F}_t\right\|^2\right)^{1/2}\\ &= O_p\left(T^{-1/2}\right)O_p(1)=O_p\left(T^{-1/2}\right),\end{aligned}$$

given that $d_i(l)=O_p\left(\min\left[N,T^{-1/2}\right]\right)$ – see Lemma 1(c) in Bai and Ng (2004). Taking all these elements together we have that $II=O_p\left(T^{-1/2}\right)+O_p\left(C_{NT}^{-1}\right)+O_p\left(T^{-1/2}\right)=O_p\left(C_{NT}^{-1}\right)$. Consequently, if $\hat{e}_{i,t}(l)\sim I(1)$

$$T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^{2}(l)=T^{-2}\sum_{t=2}^{T}e_{i,t}^{2}(l)+O_p\left(C_{NT}^{-1}\right).$$

For subsequent results we note that Bai and Ng (2010) show that averaging part II across N gives $N^{-1}\sum_{i=1}^{N}T^{-2}\sum_{t=2}^{T}e_{i,t}(l)A_{i,t}(l)=O_p\left(C_{NT}^{-2}\right)$ provided that $N/T\to 0$.

We can define the orthogonal matrix $A=[A_1:A_2]$ with A_1 a $(k\times r)$ matrix and A_2 a $(k\times m)$ matrix, $m=k-r$, such that the first r elements of the rotated vector $\hat{e}_{i,t}^{A}=A'\hat{e}_{i,t}=\left((A_1'\hat{e}_{i,t})',(A_2'\hat{e}_{i,t})'\right)'$ are I(0) and the other m elements are I(1). Using this rotation we have that $T^{-1/2}A_1'\hat{e}_{i,t}=o_p(1)$ – provided that $A_1'\hat{e}_{i,t}$ defines the stationary relations – and $T^{-1/2}A_2'\hat{e}_{i,t}=O_p(1)$ – given that $A_2'\hat{e}_{i,t}$ defines the I(1) stochastic trends.[1] Therefore, under the null hypothesis that there are m stochastic trends we have that as $T\to\infty$

$$\begin{aligned} T^{-1}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}} &= T^{-2}\hat{e}_i^{A_2\prime}\hat{e}_i^{A_2}=T^{-2}\left(\hat{e}_i^{A_2C'}\right)'\hat{e}_i^{A_2C'} \\ &\Rightarrow \Omega_{22,i}^{1/2}\int_0^1 W_i(s)W_i(s)'ds\,\Omega_{22,i}^{1/2}, \end{aligned}$$

where $W_i(s)$ denotes an m-vector of independent standard Brownian motions with $W_i(0)=0$. Therefore, using these elements the limiting distribution of the multivariate MSB statistic is given by:

$$\begin{aligned} MSB_{\mu,i}(m) &= \eta^{\min}\left(T^{-1}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}}\hat{\Omega}^{-1}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}\right) \\ &= \eta^{\min}\left(T^{-1}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}\right) \\ &\Rightarrow \eta^{\min}\left(\int_0^1 W_i(s)W_i(s)'ds\right), \end{aligned}$$

provided that $\hat{\Omega}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}\xrightarrow{p}\Omega_{\Delta e_i^{A_2}\Delta e_i^{A_2}}$, where $\xrightarrow{p}$ denotes convergence in probability.

Note that we do not observe A_2, which is estimated as the matrix of the eigenvectors associated with the largest m eigenvalues of $T^{-2}\sum_{t=1}^{T}\hat{e}_{i,t}\hat{e}_{i,t}'$. Based on Stock and Watson (1988), Bai and Ng (2004), pp. 1165, show that $\hat{A}_2\xrightarrow{p}A_2C'$ for some matrix C, so that $\hat{A}_2'\hat{e}_{i,t}=CA_2'\hat{e}_{i,t}+\left(\hat{A}_2-A_2C'\right)'\hat{e}_{i,t}=CA_2'\hat{e}_{i,t}$, given that $\left(\hat{A}_2-A_2C'\right)\xrightarrow{p}0$. Let us define $\Sigma_{22,i}=C\Omega_{22,i}C'$ and $\hat{\Sigma}_{22,i}$ a consistent estimate of $\Sigma_{22,i}$. Then $T^{-1}\hat{\Sigma}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}Q_{\hat{e}_i^{\hat{A}_2}\hat{e}_i^{\hat{A}_2}}\hat{\Sigma}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}$ has the same eigenvalues as $T^{-1}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}$.

The proof for the MSB statistic computed on the estimated common factors

[1] It should be understood that when $m=k$ then $A=A_2$.

resembles the developments above using the result that $T^{-1/2}\hat{F}_t = H\ T^{-1/2}F_t + O_p\left(C_{NT}^{-1}\right)$ – see Lemma B.2(i) in Bai and Ng (2004, pp. 1158) – so that

$$T^{-2}\sum_{t=2}^{T}\hat{F}_t^2 = H^2\ T^{-2}\sum_{t=2}^{T}F_t^2 + O_p\left(C_{NT}^{-1}\right),$$

and

$$\begin{aligned} MSB_{\mu}^{F}(q_1) &= \eta^{\min}\left(T^{-1}Q_{\hat{F}^{A_2}\hat{F}^{A_2}}\hat{\Omega}^{-1}_{\Delta\hat{F}^{A_2}\Delta\hat{F}^{A_2}}\right) \\ &\Rightarrow \eta^{\min}\left(\int_0^1 W(s)\,W(s)'\,ds\right), \end{aligned}$$

where $W(s)$ is a $(q_1\times 1)$ vector of independent standard Brownian motions. As before, we do not observe A_2, which is estimated as the matrix of the eigenvectors associated with the largest q_1 eigenvalues of $T^{-2}\sum_{t=1}^{T}\hat{F}_t\hat{F}_t'$. Since $\hat{A}_2 \xrightarrow{p} A_2C'$ for some matrix C, so that $\hat{A}_2'\hat{F}_t = CA_2'\hat{F}_t + \left(\hat{A}_2 - A_2C'\right)'\hat{F}_t = CA_2'\hat{F}_t$, given that $\left(\hat{A}_2 - A_2C'\right) \xrightarrow{p} 0$. Then $T^{-1}Q_{\hat{F}^{\hat{A}_2}\hat{F}^{\hat{A}_2}}\hat{\Omega}^{-1}_{\Delta\hat{F}^{\hat{A}_2}\Delta\hat{F}^{\hat{A}_2}}$ has the same eigenvalues as T^{-1} $Q_{\hat{F}^{A_2}\hat{F}^{A_2}}\hat{\Omega}^{-1}_{\Delta\hat{F}^{A_2}\Delta\hat{F}^{A_2}}$.

B.1.2 The linear time trend case

As above, we first focus on the idiosyncratic component. The proof for this deterministic component follows the one for the intercept case, but now $\hat{e}_{i,t} = e_{i,t} - \frac{t-1}{T-1}e_{i,T} + A_{i,t}$ with $A_{i,t} = -e_{i,1} + \frac{t-1}{T-1}e_{i,1} + \lambda_i H^{-1}V_t - d_i\hat{F}_t = -\frac{T-t}{T-1}e_{i,1} + \lambda_i H^{-1}V_t - d_i\hat{F}_t$. If $\hat{e}_{i,t}(l) \sim I(0)$ then $T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^2(l) = O_p\left(T^{-1}\right)$, whereas if $\hat{e}_{i,t}(l)\sim I(1)$ then we have

$$\begin{aligned} T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^2(l) &= T^{-2}\sum_{t=2}^{T}\left(e_{i,t}(l) - \frac{t-1}{T-1}e_{i,T}(l)\right)^2 \\ &\quad +2T^{-2}\sum_{t=2}^{T}\left(e_{i,t}(l) - \frac{t-1}{T-1}e_{i,T}(l)\right)A_{i,t}(l) \\ &\quad +T^{-2}\sum_{t=2}^{T}A_{i,t}^2(l) \\ &= I + II + III. \end{aligned}$$

Part I is $O_p(1)$ and, as before, III is $O_p\left(C_{NT}^{-2}\right)$ for all i. Part II is given by

$$
\begin{aligned}
II &= -T^{-2}\sum_{t=2}^{T}\frac{T-t}{T-1}e_{i,t}(l)\,e_{i,1}(l)+T^{-2}\sum_{t=2}^{T}e_{i,t}(l)\,\lambda_i(l)\,H^{-1}V_t \\
&\quad -T^{-2}\sum_{t=2}^{T}e_{i,t}(l)\,d_i(l)\,\hat{F}_t+T^{-2}\sum_{t=2}^{T}\frac{T-t}{T-1}\frac{t-1}{T-1}e_{i,T}(l)\,e_{i,1}(l) \\
&\quad -T^{-2}\sum_{t=2}^{T}\frac{t-1}{T-1}e_{i,T}(l)\,\lambda_i(l)\,H^{-1}V_t+T^{-2}\sum_{t=2}^{T}\frac{t-1}{T-1}e_{i,T}(l)\,d_i(l)\,\hat{F}_t \\
&= a+b+c+d+e+f.
\end{aligned}
$$

The first component is $a = O_p\left(T^{-1/2}\right) = O_p\left(C_{NT}^{-1}\right)$, and from the previous proof, $b = O_p\left(C_{NT}^{-1}\right)$ and $c = O_p\left(C_{NT}^{-1}\right)$. Consider d

$$
\begin{aligned}
d &= (T-1)^{-2}T^{-2}\sum_{t=2}^{T}\left(-t^2+(T+1)t-T\right)e_{i,T}(l)\,e_{i,1}(l) \\
&= O_p\left(T^{-1/2}\right)+O_p\left(T^{-1/2}\right)+O_p\left(T^{-3/2}\right) = O_p\left(C_{NT}^{-1}\right).
\end{aligned}
$$

Component e is given by

$$
\begin{aligned}
\|e\| &\leq \left(T^{-2}\sum_{t=2}^{T}\left\|\frac{t-1}{T-1}e_{i,T}(l)\,\lambda_i(l)\right\|^2\right)^{1/2}\left(T^{-2}\sum_{t=2}^{T}\left\|H^{-1}V_t\right\|^2\right)^{1/2} \\
&= O_p(1)\,O_p\left(N^{-1/2}\right) = O_p\left(C_{NT}^{-1}\right).
\end{aligned}
$$

Finally, component f is

$$
\begin{aligned}
\|f\| &\leq \left(T^{-2}\sum_{t=2}^{T}\left\|\frac{t-1}{T-1}e_{i,T}(l)\,d_i(l)\right\|^2\right)^{1/2}\left(T^{-2}\sum_{t=2}^{T}\left\|\hat{F}_t\right\|^2\right)^{1/2} \\
&= O_p\left(T^{-1/2}\right)O_p(1) = O_p\left(C_{NT}^{-1}\right).
\end{aligned}
$$

Therefore, part $II = O_p\left(C_{NT}^{-1}\right)$, so that if $\hat{e}_{i,t}(l) \sim I(1)$

$$
T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^2(l) = T^{-2}\sum_{t=2}^{T}\left(e_{i,t}(l)-\frac{t-1}{T-1}e_{i,T}(l)\right)^2+O_p\left(C_{NT}^{-1}\right).
$$

As already mentioned above, Bai and Ng (2010) show that averaging part II across N gives $N^{-1}\sum_{i=1}^{N}T^{-2}\sum_{t=2}^{T}e_{i,t}(l)\,A_{i,t}(l) = O_p\left(C_{NT}^{-2}\right)$ provided that $N/T \to 0$.

As before, we define $\hat{e}_{i,t}^A = A'\hat{e}_{i,t} = \left((A_1'\hat{e}_{i,t})',(A_2'\hat{e}_{i,t})'\right)'$, with $T^{-1/2}A_1'\hat{e}_{i,t} =$

$o_p(1)$ and $T^{-1/2}A_2'\hat{e}_{i,t} = O_p(1)$ provided that $A_2'\hat{e}_{i,t}$ defines the I(1) stochastic trends. Then,

$$T^{-1/2}A_2'\hat{e}_{i,t} \Rightarrow \Omega_{22,i}^{1/2}\left(W_i(s) - sW_i(1)\right)',$$

which implies that

$$\begin{aligned} T^{-1}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}} &= T^{-2}\hat{e}_i^{A_2'}\hat{e}_i^{A_2} \\ &\Rightarrow \Omega_{22,i}^{1/2}\int_0^1 V_i(s)V_i(s)'ds\,\Omega_{22,i}^{1/2}, \end{aligned}$$

where $V_i(s) = W_i(s) - sW_i(1)$ is a vector of independent Brownian bridges. Therefore,

$$MSB_{\tau,i}(m) \Rightarrow \eta^{\min}\left(\int_0^1 V_i(s)V(s)_i'ds\right),$$

given that $\hat{\Omega}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}} \xrightarrow{p} \Omega_{\Delta e_i^{A_2}\Delta e_i^{A_2}}$. As for proof for the intercept case, we have to bear in mind that we do not observe A_2, but we estimate A_2 as the matrix of the eigenvectors associated with the largest m eigenvalues of $T^{-2}\sum_{t=1}^T \hat{e}_{i,t}\hat{e}_{i,t}'$. Given that $\left(\hat{A}_2 - A_2C'\right) \xrightarrow{p} 0$, we obtain that $T^{-1}\hat{\Sigma}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}} Q_{\hat{e}_i^{\hat{A}_2}\hat{e}_i^{\hat{A}_2}} \hat{\Sigma}^{-1/2}_{\Delta\hat{e}_i^{\hat{A}_2}\Delta\hat{e}_i^{\hat{A}_2}}$ has the same eigenvalues as
$T^{-1}\hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}} Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}} \hat{\Omega}^{-1/2}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}$.

The proof for the MSB statistic computed on the estimated common factors is similar given the result that $T^{-1/2}\hat{F}_t = H\,T^{-1/2}\left[F_t - \frac{F_T - F_1}{T-1}(t-1)\right] + O_p\left(C_{NT}^{-1}\right)$ – see Bai and Ng (2004, pp. 1172) – so that

$$T^{-2}\sum_{t=2}^T \hat{F}_t^2 = H^2\,T^{-2}\sum_{t=2}^T\left[F_t - \frac{F_T - F_1}{T-1}(t-1)\right]^2 + O_p\left(C_{NT}^{-1}\right),$$

and

$$\begin{aligned} MSB_\tau^F(q_1) &= \eta^{\min}\left(T^{-1}Q_{\hat{F}^{A_2}\hat{F}^{A_2}}\hat{\Omega}^{-1}_{\Delta\hat{F}^{A_2}\Delta\hat{F}^{A_2}}\right) \\ &\Rightarrow \eta^{\min}\left(\int_0^1 V(s)V(s)'ds\right), \end{aligned}$$

where $V(s) = W(s) - sW(1)$ is a $(q_1 \times 1)$ vector of independent Brownian bridges. We estimate A_2 as the matrix of the eigenvectors associated with the largest q_1 eigenvalues of $T^{-2}\sum_{t=1}^T \hat{F}_t\hat{F}_t'$. Since $\hat{A}_2 \xrightarrow{p} A_2C'$ for some matrix C, so that $\hat{A}_2'\hat{F}_t = CA_2'\hat{F}_t + \left(\hat{A}_2 - A_2C'\right)'\hat{F}_t = CA_2'\hat{F}_t$, given that $\left(\hat{A}_2 - A_2C'\right) \xrightarrow{p} 0$. Then $T^{-1}Q_{\hat{F}^{\hat{A}_2}\hat{F}^{\hat{A}_2}}$ $\hat{\Omega}^{-1}_{\Delta\hat{F}^{\hat{A}_2}\Delta\hat{F}^{\hat{A}_2}}$ has the same eigenvalues as $T^{-1}Q_{\hat{F}^{A_2}\hat{F}^{A_2}}\hat{\Omega}^{-1}_{\Delta\hat{F}^{A_2}\Delta\hat{F}^{A_2}}$.

B.2 Proof of Theorem 2.2

We focus on the MSB statistic computed for the idiosyncratic component, but the proof entirely applies to the statistic computed using the common factors. Let us consider the null hypothesis that there are $m = k - m$ stochastic trends. From the proof of Theorem 2.1 we have that $T^{-1}A_1'e_i'e_iA_1 = O_p(1)$, $T^{-1}A_1'e_i'e_iA_2 = O_p(1)$ and $T^{-2}A_2'e_i'e_iA_2 = O_p(1)$, so that $T^{-2}A_1'e_i'e_iA_1 = O_p\left(T^{-1}\right)$ and $T^{-2}A_1'e_i'e_iA_2 = O_p\left(T^{-1}\right)$. Consequently, under the alternative hypothesis that there are $l < m$ stochastic trends the rank of the matrix $T^{-1}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}}$ will be $l < m$. Using these elements, we can see that the cross-products involving I(0) stochastic processes in $T^{-1}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}}$ tend to zero at rate $O_p\left(T^{-1}\right)$.

Let us now focus on the estimate of the long-run covariance matrix. Note that under both the null and the alternative hypotheses $T^{-1}\hat{\varepsilon}_i'\hat{\varepsilon}_i = O(1)$, with $T^{-1}\hat{\varepsilon}_i'\hat{\varepsilon}_i \xrightarrow{p} \Sigma_{\varepsilon_i}$. Since all roots of the determinant of $\left(I - \hat{\Gamma}_{i,p_i}(L)\right)$ lie outside the unit circle interval, we can define $\hat{\Xi}_{i,\infty}(L) = \left(I - \hat{\Gamma}_{i,p_i}(L)\right)^{-1}$, with $\Xi_{i,\infty}(L) = (I + \Xi_{i,1}L + \Xi_{i,2}L^2 + \cdots)$ and where the sequence of matrix coefficients $\{\Xi_{i,s}\}_{s=0}^{\infty}$ is absolutely summable. Then, $\Xi_{i,\infty}(1) < \infty$ so that $\hat{\Omega}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}} \xrightarrow{p} \Xi_{i,\infty}'(1)\Sigma_{\varepsilon_i}\Xi_{i,\infty}(1) = \Omega_{\Delta e_i^{A_2}\Delta e_i^{A_2}}$. Therefore, the long-run covariance matrix estimator converges to a positive definite matrix under both the null and the alternative hypotheses – note that this result can be seen as the generalization of the one in Perron and Ng (1996) and Stock (1999).

Finally, note that under the alternative hypothesis that there are $l < m$ stochastic trends, $rank\left(T^{-1}Q_{\hat{e}_i^{A_2}\hat{e}_i^{A_2}}\hat{\Omega}^{-1}_{\Delta\hat{e}_i^{A_2}\Delta\hat{e}_i^{A_2}}\right) = l$, where the elements that cause rank deficiency tend to zero at rate $O_p\left(T^{-1}\right)$. This proves the consistency of the MSB statistic under the alternative hypothesis.

B.3 Proof of Theorem 2.3

Let us first focus on the intercept model specification. Note that for any real symmetric $(k \times k)$ matrix, say $B_i = \frac{1}{T^2}e_i'e_i$, it is true that $\eta_i^{\min}(B_i) \leq T^{-2}\sum_{t=2}^{T} e_{i,t}^2(l)$ for $l = 1, \ldots, k$. Using this inequality we have

$$\hat{\eta}_{i,1} - \eta_{i,1} \leq T^{-2}\sum_{t=2}^{T} \hat{e}_{i,t}^2(l) - T^{-2}\sum_{t=2}^{T} e_{i,t}^2(l) = O_p\left(C_{NT}^{-1}\right), \tag{B.3.1}$$

where $\hat{\eta}_{i,1} = \eta^{\min}\left(\frac{1}{T^2}\hat{e}_i'\hat{e}_i\right)$ and $\eta_{i,1} = \eta^{\min}\left(\frac{1}{T^2}e_i'e_i\right)$, given that $T^{-2}\sum_{t=2}^{T}\hat{e}_{i,t}^2(l) = T^{-2}\sum_{t=2}^{T}e_{i,t}^2(l) + O_p\left(C_{NT}^{-1}\right)$. Therefore,

$$\hat{\eta}_{i,1} \leq \eta_{i,1} + O_p\left(C_{NT}^{-1}\right). \tag{B.3.2}$$

The same result is achieved for the time trend specification, but with $T^{-2}\sum_{t=2}^{T}e_{i,t}^2(l)$ replaced by $T^{-2}\sum_{t=2}^{T}\left(e_{i,t}(l) - \frac{t-1}{T-1}e_{i,T}(l)\right)^2$.

Bai and Ng (2010) show that averaging the $O_p\left(C_{NT}^{-1}\right)$ component in (B.3.1) and (B.3.2) across N produces a term that is $O_p\left(C_{NT}^{-2}\right)$ provided that $N/T \to 0$ – see the order of magnitude of part II in Equation (A.1) of Bai and Ng (2010). Consequently, under the null hypothesis that there are m stochastic trends

$$\begin{aligned}
PMSB_j^Z(m) &= \frac{\sqrt{N}(\overline{MSB}_j(m) - E(MSB_j(m)))}{\sqrt{Var(MSB_j(m))}} \\
&= \frac{\sqrt{N}\left[N^{-1}\sum_{i=1}^{N}\left(\eta_{i,1} + O_p\left(C_{NT}^{-1}\right)\right) - E(MSB_j(m))\right]}{\sqrt{Var(MSB_j(m))}} \\
&= \frac{\sqrt{N}\left[N^{-1}\sum_{i=1}^{N}\eta_{i,1} - E(MSB_j(m))\right]}{\sqrt{Var(MSB_j(m))}} + \frac{\sqrt{N}}{\min[N,T]} \\
&= \frac{\sqrt{N}\left[N^{-1}\sum_{i=1}^{N}\eta_{i,1} - E(MSB_j(m))\right]}{\sqrt{Var(MSB_j(m))}} + o_p(1),
\end{aligned}$$

for $j = \{\mu, \tau\}$. Therefore, as $T \to \infty$, $N \to \infty$, with $N/T \to 0$, and assuming finite second moments of the random variables characterized as Brownian motion functionals $\Upsilon \equiv \left(\eta^{\min}\left(\int_0^1 W_i(s)W_i(s)'ds\right), \eta^{\min}\left(\int_0^1 V_i(s)V_i(s)'ds\right)\right)'$, $PMSB_j^Z(m) \Rightarrow N(0,1)$, $j = \{\mu, \tau\}$, by the Lindberg-Levy Central Limit Theorem.

Let us now focus on the panel data statistics that pool the p-values of the individual statistics. To avoid unnecessary confusions, we avoid the use of the subscript i that denotes the time series in the panel data here. First, we note that the p-value is given by the cumulated distribution function, which using Imhof's (1961) formula it can be expressed as – see Tanaka (1996), eq. (6.13):

$$F(\hat{\eta}_1) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{\theta}\operatorname{Im}\left[e^{-i\theta\hat{\eta}_1}\phi_j(\theta)\right]d\theta, \tag{B.3.3}$$

where $i = \sqrt{-1}$, $\operatorname{Im}[\cdot]$ takes the imaginary part of the argument, and $\phi_j(\theta)$ denotes the characteristic function, $j = \{\mu, \tau\}$, where in our case $\phi_\mu(\theta) = \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-1/2}$

and $\phi_\tau(\theta) = \left(\cos\sqrt{2i\theta}\right)^{-1/2}$ – see Tanaka (1996), pp. 111 and 112. Further,

$$\begin{aligned} F(\hat{\eta}_1) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{1}{\theta}\phi_j(\theta)\sin\theta\hat{\eta}_1 d\theta \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{1}{\theta}\phi_j(\theta)\sin\left[\theta\left(\eta_1 + O_p\left(C_{NT}^{-1}\right)\right)\right] d\theta. \end{aligned}$$

Note that we can write

$$\begin{aligned} &\sin\left[\theta\eta_1 + \theta O_p\left(C_{NT}^{-1}\right)\right] = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!}\left[\theta\eta_1 + \theta O_p\left(C_{NT}^{-1}\right)\right]^{2n-1} \\ = &\left[\theta\eta_1 + \theta O_p\left(C_{NT}^{-1}\right)\right] + \frac{-1}{3!}\left[\theta\eta_1 + \theta O_p\left(C_{NT}^{-1}\right)\right]^3 \\ &+ \frac{1}{5!}\left[\theta\eta_1 + \theta O_p\left(C_{NT}^{-1}\right)\right]^5 + \cdots \\ = &\left[\theta\eta_1 + \theta O_p\left(C_{NT}^{-1}\right)\right] + \frac{-1}{3!}\left[(\theta\eta_1)^3 + 3(\theta\eta_1)^2\theta O_p\left(C_{NT}^{-1}\right)\right] \\ &+ \frac{1}{5!}\left[(\theta\eta_1)^5 + 5(\theta\eta_1)^4\theta O_p\left(C_{NT}^{-1}\right)\right] + \cdots \\ = &\sin(\theta\eta_1) + O_p\left(C_{NT}^{-1}\right)\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{(2n-1)!}\left[(2n-1)(\theta\eta_1)^{n-1}\theta\right] \\ = &\sin(\theta\eta_1) + O_p\left(C_{NT}^{-1}\right)O(1), \end{aligned}$$

so that

$$\begin{aligned} F(\hat{\eta}_1) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{1}{\theta}\phi_j(\theta)\sin\theta\eta_1 d\theta + O_p\left(C_{NT}^{-1}\right) \\ &= \varphi + O_p\left(C_{NT}^{-1}\right), \end{aligned}$$

where φ denotes the p-value of the statistic. The statistic given in (2.4.3) takes the form

$$PMSB_j^C(m) = \frac{-2\sum_{i=1}^N \ln\hat{\varphi}_i - 2N}{\sqrt{4N}}.$$

The term $\sum_{i=1}^N \ln\hat{\varphi}_i$ that appears in the numerator of the statistic

$$\begin{aligned} \sum_{i=1}^N \ln\hat{\varphi}_i &= \ln\left(\prod_{i=1}^N\left(\varphi_i + O_p\left(C_{NT}^{-1}\right)\right)\right) \\ &= \ln\left(\prod_{i=1}^N \varphi_i + O_p\left(C_{NT}^{-1}\right)\right), \end{aligned}$$

which in the limit converges to $\ln\left(\prod_{i=1}^{N}\varphi_i\right) = \sum_{i=1}^{N}\ln\varphi_i$. Consequently, as $T\to\infty$, $N\to\infty$ and $N/T\to 0$, the $PMSB_j^C(m)$ statistic converges to the standard Normal distribution. Theorem 2.3 has been proved.

Appendix C

Appendix for the Third Chapter

The description of the variables that are used in the database is the following:

- $Y_{i,t}$ = the output, measured by GVA at factor cost of region i in the year t at 1980 constant prices, from the BD.MORES database, Spanish Ministry of Finance and Public Administrations.

- $K_{i,t}$ = the stock of private capital of region i in the year t at 1980 constant prices, from the Stock de Capital database, IVIE.

- $G_{i,t}$ = the stock of total public capital of region i in the year t at 1980 constant prices, from the Stock de Capital database, IVIE.

- $GP_{i,t}$ = the stock of productive public capital of region i in the year t at 1980 constant prices, from the Stock de Capital database, IVIE.

- $H_{i,t}$ = the stock of human capital, measured as a share of the employed population with secondary and university education of region i in the year t, from the Stock de Capital Humano database, IVIE.

- $HS_{i,t}$ = the stock of human capital, measured as an average years of schooling, from the Stock de Capital Humano database, IVIE, and Serrano (1996).

- $L_{i,t}$ = labor, measured as the employed population of region i in the year t, from the Stock de Capital Humano database, IVIE.

Bibliography

[1] Amsler, C. and J. Lee (1995) An LM Test for a Unit-Root in the Presence of a Structural Change, *Econometric Theory*, 11, 359-368.

[2] Andrews, D.W.K. (2005) Cross-Section Regression with Common Shocks, *Econometrica*, 73, 1551-1585.

[3] Aschauer, D. (1989) Is Public Expenditure Productive?, *Journal of Monetary Economics*, 23, 177-200.

[4] Bai, J. (2004) Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data, *Journal of Econometrics*, 122, 137-183.

[5] Bai, J. and J.L. Carrion-i-Silvestre (2009) Structural Changes, Common Stochastic Trends, and Unit Roots in Panel Data, *Review of Economic Studies*, 76, 471-501.

[6] Bai, J. and J.L. Carrion-i-Silvestre (2013) Testing Panel Cointegration with Unobservable Dynamic Common Factors that are Correlated with the Regressors, *Econometrics Journal*, 16, 222-249.

[7] Bai, J., Kao, C. and S. Ng (2009) Panel Cointegration with Global Stochastic Trends, *Journal of Econometrics*, 149, 82-99.

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

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

[10] Bai, J. and S. Ng (2010) Panel Unit Root Tests with Cross-Section Dependence: A Further Investigation, *Econometric Theory* , 26, 1088-1114.

[11] Bajo-Rubio, O. and C. Díaz-Roldán (2005) Optimal Endowments of Public Capital: An Empirical Analysis for the Spanish Regions, *Regional Studies*, 39, 297-304.

[12] Baltagi, B.H. and N. Pinnoi (1995) Public Capital Stock and State Productivity Growth: Further Evidence from an Error Component Model, *Empirical Economics*, 20, 351-359.

[13] Banerjee, A. and J.L. Carrion-i-Silvestre (2011) Testing for Panel Cointegration using Common Correlated Effects Estimators, Working Paper.

[14] Banerjee, A. and J.L. Carrion-i-Silvestre (2013) Cointegration in Panel Data with Breaks and Cross-Section Dependence, *Journal of Applied Econometrics*, forthcoming.

[15] Banerjee, A., Eberhardt, M. and J.J. Reade (2010) Panel Estimation for Worriers, Working Paper, University of Birmingham.

[16] Banerjee, A., Lumsdaine, R.L. and J.H. Stock (1992) Recursive and Sequential Tests of the Unit Root and Trend-Break Hypothesis: Theory and International Evidence, *Journal of Business and Economic Statistics*, 10, 271-287.

[17] Banerjee, A., Marcellino, M. and C. Osbat (2004) Some Cautions on the Use of Panel Methods for Integrated Series of Macroeconomic Data, *Empirical Economics*, 7, 322-340.

[18] Banerjee, A., Marcellino, M. and C. Osbat (2005) Testing for PPP: Should We Use Panel Methods?, *Empirical Economics*, 30, 77-91.

[19] Banerjee, A. and M. Wagner (2009) Panel Methods to Test for Unit Roots and Cointegration, in *Palgrave Handbook of Econometrics, Vol. 2: Applied Econometrics*, Mills TC, Patterson K (eds), Palgrave Macmillan.

[20] Basher, S.A. and J.L. Carrion-i-Silvestre (2007) Another Look at the Null of Stationary Real Exchange Rates: Panel Data with Structural Breaks and Cross-section Dependence, IREA Working Paper, University of Barcelona.

[21] Berndt, E.R. and B. Hansson (1992) Measuring the Contribution of Public Infrastructure Capital in Sweden, *Scandinavian Journal of Economics*, 94, 151-168.

[22] Breitung, J. (2005) A Parametric Approach to the Estimation of Cointegration Vectors in Panel Data, *Econometric Reviews*, 24, 151-173.

[23] Breitung, J. and M.H. Pesaran (2008) Unit Roots and Cointegration in panels, in Matyas, L., and P. Sevestre, *The Econometrics of Panel Data*, Boston: Kluwer Academic Publishers, 279-322.

[24] Carrion-i-Silvestre, J.L., del Barrio Castro, T. and E. López-Bazo (2001) Level Shifts in a Panel Data Based Unit Root Test. An Application to the Rate of Unemployment, *Proceedings of the 2001 North American Econometric Society*.

[25] Carrion-i-Silvestre, J.L., del Barrio Castro, T. and E. López-Bazo (2005) Breaking the Panels. An Application to Real per Capita GDP, *Econometrics Journal*, 8, 159-175.

[26] Carrion-i-Silvestre, J.L., Kim, D. and P. Perron (2009) GLS-based Unit Root Tests with Multiple Structural Breaks both Under the Null and the Alternative Hypotheses, *Econometric Theory*, 25, 1754-1792.

[27] Carrion-i-Silvestre, J.L. and L. Surdeanu (2011) Panel Cointegration Rank Testing with Cross-Section Dependence, *Studies in Nonlinear Dynamics & Econometrics*, 15, article 4.

[28] Choi, I. (2001) Unit Root Tests for Panel Data, *Journal of International Money and Finance*, 20, 249-272.

[29] Christensen, L.R., Jorgenson, D.W. and L.J. Lau (1971) Conjugate Duality and the Transcendental Logarithmic Production Function, *Econometrica*, 39, 255-256.

[30] Christiano, L.J. (1992) Searching for a Break in GNP, *Journal of Business & Economic Statistics*, 10, 237-250.

[31] Chudik, A., Pesaran, M. H. and E. Tosetti (2011) Weak and Strong Cross Section Dependence and Estimation of Large Panels, *Econometrics Journal*, 14, C45-C90.

[32] Clemente, J., Montañés, A. and M. Reyes (1998) Testing for a Unit Root in Variables with a Double Change in the Mean, *Economics Letters*, 59, 175-182.

[33] Cobb, C.W. and P.H. Douglas (1928) A Theory of Production, *American Economic Review*, 18, 139-165.

[34] Costantini, M. and S. Destefanis (2009) Cointegration Analysis for Cross-Sectionally Dependent Panels: The Case of Regional Production Functions, *Economic Modelling*, 26, 320-327.

[35] Dalamagas, B. (1995) A Reconsideration of the Public Sector's Contribution to Growth, *Empirical Economics*, 20, 385-414.

[36] Dawson, J.W. and M.C. Strazicich (2010) Time-Series Tests of Income Convergence with Two Structural Breaks: Evidence from 29 Countries, *Applied Economics Letters*, 17, 909-912.

[37] Dickey, D.A. and W.A. Fuller (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association*, 74, 427-431.

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

[39] Evans, A.D., Green, C.J. and V. Murinde (2002) Human Capital and Financial Development in Economic Growth: New Evidence Using the Translog Production Function, *International Journal of Finance and Economics*, 7, 123-140.

[40] Evans, P. and G. Karras (1994) Is Government Capital Productive? Evidence from a Panel of Seven Countries, *Journal of Macroeconomics*, 16, 271-279.

[41] Felipe, J. (1998) On the Interpretation of Coefficients in Multiplicative-Logarithmic Functions: a Reconsideration, *Applied Economics Letters*, 5, 397-400.

[42] Friedlaender, A.F. and R.H. Spady (1981) *Freight Transport Regulation: Equity, Efficiency and Competition in the Rail and Trucking Industries*, MIT Press, Cambridge, MA.

[43] García-Milá, T. and T. McGuire (1992) The Contribution of Publicly Provided Inputs to States Economics, *Regional Science and Urban Economics*, 22, 229-41.

[44] García-Milá T., McGuire T. and R. Porter (1996) The Effect of Public Capital in State-Level Production Functions Reconsidered, *The Review of Economics and Statistics*, 78, 177-180.

[45] Gengenbach, C., Palm, F.C. and J.P. Urbain (2006) Cointegration Testing in Panels with Common Factors, *Oxford Bulletin of Economics and Statistics*, 68, 683-719.

[46] Gengenbach, C., Palm, F.C. and J.P. Urbain (2010) Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling, *Econometric Reviews*, 29, 111-145.

[47] Gengenbach, C., Westerlund, J. and J.P. Urbain (2008) Panel Error Correction Testing with Global Stochastic Trends, METEOR Working Paper, 51.

[48] Gonzalo, J. and C.W.J. Granger (1995) Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Business & Economic Statistics*, 13, 27-35.

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

[50] Growiec, J., Pajor, A., Pelle, D. and A. Predki (2011) The Shape of Aggregate Production Functions: Evidence from Estimates of the World Technology Frontier, Working Paper, National Bank of Poland.

[51] Hadri, K. (2000) Testing for Stationarity in Heterogeneous Panel Data, *Econometrics Journal*, 3, 148-161.

[52] Harris, R.D.F. and E. Tzavalis (1999) Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed, *Journal of Econometrics*, 91, 201-226.

[53] Hoffman, D.L., Rasche, R.H., and M.A. Tieslau (1995) The Stability of Long-Run Money Demand in Five Industrial Countries, *Journal of Monetary Economics*, 35, 317-339.

[54] Holly, S., Pesaran, M.H. and T. Yamagata (2010) A Spatio-Temporal Model of House Prices in the USA, *Journal of Econometrics*, 158, 160-173.

[55] Holtz-Eakin, D. (1994) Public Sector Capital and the Productivity Puzzle, *Review of Economics and Statistics*, 76, 12-21.

[56] Im, K.S., Lee, J. and M. Tieslau (2005) Panel LM Unit Root Test with Level Shifts, *Oxford Bulletin of Economics and Statistics*, 67, 393-419.

[57] Im, K.S., Lee, J. and M. Tieslau (2012) Panel LM Unit Root Tests with Level and Trend Shifts, Working Paper.

[58] Kapetanios, G. (2005) Unit-Root Testing Against the Alternative Hypothesis of up to m Structural Breaks, *Journal of Time Series Analysis*, 26, 123-133.

[59] Kapetanios, G., Pesaran, M.H. and T. Yamagata (2011) Panels with Non-stationary Multifactor Error Structures, *Journal of Econometrics*, 160, 326-348.

[60] Kejriwal, M. and C. Lopez (2010) Unit Roots, Level Shifts and Trend Breaks in Per Capita Output: A Robust Evaluation, Working Paper, University of Cincinnati.

[61] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Y. Shin (1992) Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, *Journal of Econometrics*, 54, 159-178.

[62] Lee, C. and J.L. Wu (2012) A Panel Unit-Root Test with Smooth Breaks and Cross-Sectional Dependence, Working Paper.

[63] Lee, J. and M.C. Strazicich (2003) Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks, *The Review of Economics and Statistics*, 85, 1082-1089.

[64] Lopez, C. (2009) A Panel Unit Root Tests With Good Power in Small Samples, *Econometric Reviews*, 28, 295-313.

[65] López-Bazo, E. and R. Moreno (2008) Does Human Capital Stimulate Investment in Physical Capital?: Evidence from a Cost System Framework, *Economic Modelling*, 25, 1295-1305.

[66] Lumsdaine, R. L. and D. Papell (1997) Multiple Trend Breaks and the Unit-Root Hypothesis, *Review of Economics and Statistics*, 79, 212-218.

[67] Maddala, G. S. and S. Wu (1999) A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

[68] Maddison, A. (2009) Statistics on World Population, GDP and Per Capita GDP 1-2008 AD, http://www.ggdc.net/MADDISON/oriindex.htm

[69] Mark, N.C. and D. Sul (2001) Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel, *Journal of International Economics*, 53, 29-52.

[70] 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, 655-680.

[71] Márquez, M.A., Ramajo, J. and G.J.D. Hewings (2011) Public Capital and Regional Economic Growth: a SVAR Approach for the Spanish Regions, Working Paper.

[72] Merriman, D. (1990) Public Capital and Regional Output, *Regional Science and Urban Economics*, 20, 437-458.

[73] Moon, H.R. and B. Perron (2004) Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics*, 122, 81-126.

[74] Moon, H.R., Perron, B. and P.C.B. Phillips (2007) Incidental Trends and the Power of Panel Unit Root Tests, *Journal of Econometrics*, 141, 416-459.

[75] Munnell, A. (1990) How Does Public Infrastructure Affect Regional Economic Performance?, *New England Economic Review, Federal Reserve Bank of Boston*, September-October, 11-32.

[76] Murthy, K.V.B. (2002) Arguing a Case for the Cobb-Douglas Production Function, *Review of Commerce Studies*, 20-21, 75-91.

[77] Ng, S. and P. Perron (1995) Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag, *Journal of the American Statistical Association*, 90, 269-281.

[78] Ng, S. and P. Perron (2001) Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power, *Econometrica*, 69, 1519-1554.

[79] Ohara, H.I. (1999) A Unit Root Test with Multiple Trend Breaks: A Theory and Application to US and Japanese Macroeconomic Time-Series, *The Japanese Economic Review*, 50, 266-290.

[80] Otto, G. and G. Voss (1996) Public Capital and Private Production in Australia, *Southern Economic Journal*, 62, 723-39.

[81] Pablo-Romero, M. del P. and M. de la P. Gómez-Calero (2013) A Translog Production Function for the Spanish Provinces: Impact of the Human and Physical Capital in Economic Growth, *Economic Modelling*, 77-87.

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

[83] Perron, P. (1997) Further Evidence on Breaking Trend Functions in Macroeconomic Variables, *Journal of Econometrics*, 80, 355-385.

[84] Perron, P. (2006) Dealing with Structural Breaks, in *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory*, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278-352.

[85] Perron, P. and S. Ng (1996) Useful Modifications to Some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties, *Review of Economic Studies*, 63, 435-463.

[86] Perron, P. and S. Ng (1998) An Autoregressive Spectral Density Estimator At Frequency Zero For Nonstationarity Tests, *Econometric Theory*, 14, 560-603.

[87] Perron, P. and T.J. Vogelsang (1992) Nonstationarity and Level Shifts with an Application to Purchasing Power Parity, *Journal of Business and Economic Statistics*, 10, 301-320.

[88] Pesaran, M.H. (2004) General Diagnostic Tests for Cross Section Dependence in Panels, CESifo Working Paper, 1229.

[89] Pesaran, M.H. (2006) Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure, *Econometrica*, 74, 967-1012.

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

[91] Pesaran, M.H. (2013) Testing Weak Cross-sectional Dependence in Large Panels, *Econometric Reviews*, forthcoming.

[92] Pesaran, M.H., V. Smith and T. Yamagata (2013) Panel Unit Root Tests in the Presence of a Multifactor Error Structure, *Journal of Econometrics*, 175, 94-115.

[93] Pesaran, M.H. and E. Tosetti (2011) Large Panels with Common Factors and Spatial Correlations, *Journal of Econometrics*, 161, 182-202.

[94] Ploberger, W. and P.C.B. Phillips (2004) Optimal Testing for Unit Roots in Panel Data, Working Paper, University of Rochester.

[95] Qu, Z. and P. Perron (2007) A Modified Information Criterion for Cointegration Tests Based on a VAR Approximation, *Econometric Theory*, 23, 638-685.

[96] Rapach, D. and M. Wohar (2004) Testing the Monetary Model of Exchange Rate Determination: a Closer Look at Panels, *Journal of International Money and Finance*, 23, 867-895.

[97] Saikkonen, P. and H. Lutkepohl (2000) Testing for the Cointegrating Rank of a VAR Process with an Intercept, *Econometric Theory*, 16, 373-406.

[98] Sauer, J. and H. Tchale (2006) Alternative Soil Fertility Management Options in Malawi - An Economic Analysis, *American Agricultural Economics Association*, Long Beach, CA.

[99] Serrano, L. (1996) Indicadores de Capital Humano y Productividad, *Revista de Economía Aplicada*, 10, 177-190.

[100] Serrano, L. (1997) Productividad del Trabajo y Capital Humano en la Economía Española, *Moneda y Crédito*, 205, 79-101.

[101] Sosvilla-Rivero, S. and J. Alonso (2005) Estimación de una Función de Producción MRW para la Economía Española, 1910-1995, *Investigaciones Economicas*, Fundación SEPI, 29, 3, 609-624.

[102] Stock, J.H. (1999) A Class of Tests for Integration and Cointegration, in R. F. Engle and H. White (eds.) *Cointegration, Causality and Forecasting. A Festchrift in Honour of Clive W. F. Granger*, Oxford University Press.

[103] Stock, J.H. and M.C. Watson (1988) Testing for Common Trends, *Journal of the American Statistical Association*, 83, 1097-1107.

[104] Stock, J.H. and M.C. Watson (1993) A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems, *Econometrica*, 61, 783-820.

[105] Sul, D. (2005) New Panel Unit Root Tests under Cross Section Dependence for Practitioners, Working Paper.

[106] Sul, D., Phillips, P.C.B. and C. Choi (2005) Prewhitening Bias in HAC Estimation, *Oxford Bulletin of Economics and Statistics*, 67, 517-546.

[107] Tam, P.S. (2006) Breaking Trend Panel Unit Root Tests, Working Paper, University of Macau.

[108] Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, New York: John Wiley & Sons.

[109] Toda, H.Y. (1995) Finite Sample Performance of Likelihood Tests for Cointegrating Ranks in Vector Autoregressions, *Econometric Theory* , 11, 1015-1032.

[110] Westerlund, J. (2008) Panel Cointegration Tests of the Fisher Hypothesis, *Journal of Applied Econometrics*, 23, 193-233.

[111] Westerlund, J. (2012) Testing for Unit Roots in Panel Time-Series Models with Multiple Level Breaks, *The Manchester School*, 80, 671-699.

[112] Westerlund, J. (2013) Asymptotic Redundance Results for the Use of GLS Detrending in Panel Unit Root Testing, Working Paper, Deakin University.

[113] Wylie, P. (1996) Infrastructure and Canadian Economic Growth, 1946-91, *Canadian Journal of Economics*, 29, 350-55.

[114] Zivot, E. and D.W.K. Andrews (1992) Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Journal of Business & Economic Statistics*, 10, 251-270.