Testing for breaks in the weighting matrix

Ana Angulo[a] , Peter Burridge[b] and Jesús Mur[a]

[a] *Department of Economic Analisis, University of Zaragoza, Spain*

[b] *Department of Economic and Related Studies, University of York, UK*

July 14, 2016

Abstract

The weighting matrix is a key element in the specification of a spatial model. Typically, this matrix is fixed a priori by the researcher, which is not always satisfactory. Theoretical justification for the chosen matrix tends to be very vague, and the selection problem is seldom reconsidered. However, several recent proposals advocate a more data-driven approach. In fact, if we have panel data, the weighting matrix can be estimated from the data; this facilitates the development of statistical procedures for testing various hypotheses of interest. In the paper, we focus on the assumption of stability, through time, of this matrix by adapting a collection of covariance matrix stability tests, developed in a multivariate context. The tests are compared in a Monte Carlo; two examples illustrate the proposal.

Keywords: *Weights matrix; Estimation of W; Structural breaks; Tests of equality.*

JEL classification: C4, C5, R1.

*Corresponding author: Jesus Mur. Department of Economic Analysis. University of Zaragoza (email: jmur@unizar.es)

1 Introduction

Brun et al. (2005) draw attention to an interesting fact: most of the literature on gravity models finds estimated distance coefficients that increase, rather than decrease, through time. This is unexpected because (p. 103) *"the common perception of globalization is that distance should be becoming less important for international trade, implying decreasing rather than increasing values of the estimated coefficient of distance"*. The puzzle is solved after correcting some mis-specifications in the gravity equation. Then the authors (p. 117) announce the death of distance, estimating an *"11% decrease in the impact of distance on bilateral trade over the 35-year period 1962-1996"*. This is just one example, that makes clear that spatial relations are not static but evolve over time. Not only may the strength of estimated spatial dependence change, but also the pattern of such dependence, often represented by a matrix, $\mathbf{W}$. Another example is the model for household demand of Case (1991), where the weights in the $\mathbf{W}$ matrix depend on incomes of the different districts (so-called income distance), which means weights (i) may be endogeneous and (ii) may change continuously.

The case of an endogenous $\mathbf{W}$ matrix is well-known, though technically difficult. Only recently have Kelejian and Piras (2014) and Qu and Lee (2015) formalized its treatment using algorithms (such as generalized method of moments, GMM, or instrumental variables, IV) with good properties. Ahrens and Bhattacharjee (2015) develop a two-step Lasso algorithm to estimate the weights in a spatial autoregressive panel data model. The Lasso approach allows the authors to circumvent at the same time the problems of endogeneity, associated with the spatial lags of the endogenous variable, and of high dimensionality of the equation. Let us note that they treat the weights of $\mathbf{W}$ as unknown coefficients to be estimated so sparseness of this matrix is a key assumption (more sparseness, fewer parameters to estimate, faster convergence).

We are aware of few papers that deal with changes in the $\mathbf{W}$ matrix across time. Druska and Horrace (2004) in their study of Javan rice farm efficiency, may be the first to work formally with a time-varying $\mathbf{W}$; they use a two seasons model which involves two different pre-specified weight matrices, one for the dry season, the other for the wet season. Obviously, season changes are known in advance, and so are exogenous. Lee and Yu (2012) develop a quasi-maximum likelihood estimation, QML, of dynamic panel data models where spatial weights matrices are also time-varying; they are assumed exogeneous and known. QML estimates are consistent and asymptotically normal when both the number of spatial units and time periods increase. Moreover, Lee and Yu also detect significant biases in the case of substantial misspecification, with a time invariant weighting matrix assumed when the true process has time varying matrices. The magnitude of the bias increases for the estimates of the marginal direct/indirect effects. Angulo *et al.* (2016) focus on the problem of testing for changes in the matrix across time, building the weights by means of exponential and inverse distance decay functions; a likelihood ratio test with good properties, even for very small samples, is obtained. Our goal is along the same lines, trying to provide tools to detect changes in the weighting matrix.

The discussion below is confined to the most common situation in applied work, where the weighting matrix is exogenous (probably depending on physical distance or some other measure of separation). One of the main difficulties for testing hypotheses in relation to **W** is the potential problem of high-dimensionality given that, usually, the cross sectional dimension is larger than the temporal dimension of the data. This prevents the use of standard approaches such as maximum-likelihood, which is badly equipped to deal with singularity of the covariance matrix. Another difficulty is the lack of identification that affects most spatial models, in the sense that is difficult to separate the impact of a spatial coefficient from the effect of the **W** matrix to which it is applied. However, given the unambiguous relation that exists between the weighting matrix and the covariance matrix, the literature devoted to the second appears to offer a reasonable framework for the problem. We review this strand of literature looking for tests that would allow us to test for the constancy, through time, of the weights matrix. The aim is to demonstrate the usefulness of these procedures in applied research, either using a Spatial Error Model, SEM, or a Spatial Lag Model, SLM.

The second section discusses briefly the role of the **W** matrix in a typical spatial model. In Section 3 we review the recent literature devoted to the problem of comparing covariance matrices estimated using different samples. We select a set of equality tests, apparently with good properties, that address the problem. Section 4 presents the results of a Monte Carlo experiment comparing these tests. Section 5 illustrates our proposal with two case studies in which the assumption of constancy in **W** plays a crucial role. Section 6 concludes.

2 The W issue: some stylized facts

Corrado and Fingleton (2012) pose three questions in relation to **W**. For what is it used? How is it built? Finally, do we need it?

The answer to the first question is simple: spatial lags of some variables are necessary because their spatial spillovers are unobservable events in real life but we have to account for their effects in our models. So, we need to speculate about their appropriate form by using proxy variables.

A matrix of weights constitutes a reasonable procedure, though it is not the only solution. Oud and Folmer (2008), for example, introduce the latent variables approach to account for spatial dependence using a structural equations model; spatially lagged variables are represented by latent variables which are measured through a set of proxies. Paci and Usai (2009) extend the use of proxy variables to the problem of measuring unobserved spatial (knowledge) flows produced over space, where geography constitutes an additional source for proxy variables. The HAC method of Kelejian and Prucha (2007) can be seen as another way of dealing with unobserved effects related to the spatial structure.

However, these cases are the exception; in fact, there are not many alternatives to the use of **W**, which must therefore be defined. Roughly speaking, we may distinguish two approaches to the building of **W** (Harris *et al.*, 2011): (i) specifying **W** exogenously; (ii) estimating **W** from data. The exogenous approach is by far the most common and includes, for example, use of a binary contiguity criterion, k-nearest neighbours, kernel functions based on distance, etc.

The second approach uses the topology of the space and the nature of the data, and takes many forms. Most are ad-hoc procedures in which an objective is selected in advance which guides the search. Kooijman (1976) was one of the first to tackle explicitly the question of estimating a **W** matrix. His suggestion, to build the weights so as to maximize the value of Moran's I, is intuitive, but is difficult to implement because of the high number of unknowns. Kooijman's suggestion is related with other more recent proposals such as Griffith (1996), who tries to find a **W** able to absorb the spatial effects from the data. Fernández *et al.* (2009) propose a specification of **W** based on a measure of entropy, the GME approach, while Mur and Paelinck (2010) focus on the maximization of the Complete Correlation Coefficient, the CCC approach. The LSM (local statistical model) of Getis and Aldstadt (2004) is very popular, as is the AMOEBA algorithm of Aldstadt and Getis (2006).

Also included is the most recent work in which $\mathbf{W}$ is directly estimated from the data. For example, Benjanuvatra and Burridge (2015) present a QML algorithm to estimate the weights in $\mathbf{W}$ using a single cross-section, under the assumption that these weights are a function of the distance between the locations, known up to a parameter that may be estimated.

More flexible approaches to $\mathbf{W}$ are possible if repeated information about the interactions is available. The initial suggestion of Meen (1996) is to set the problem in a multivariate framework: (i) a SUR model is estimated where each equation corresponds to a region (assuming necessary homogeneity and aggregation restrictions), then (ii) spatial error dependence, if it exists, will be captured by the covariance matrix of the SUR residuals. As indicated by Meen (1996, p 360), "*the advantage is that there is no need to specify in advance the form of the spatial error dependence through a weights matrix*". He goes a step further suggesting a LS regression of the SUR residuals in each region on the residuals of the other regions:

$$\hat{u}_{rt} = \sum_{\substack{j=1 \\ j \neq r}}^{n} \omega_{rj} \hat{u}_{jt} + \varepsilon_{rt}, \qquad t = 1, ..., T. \tag{1}$$

Then the t-ratios can be used to assess if the sequence of weights, ω_{rj}, is statistically significant in each equation. Unfortunately, there are strong endogeneity problems in (1).

Battacharjee and Jensen-Butler (2013) consider a panel data model with SEM errors:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{x}_t \boldsymbol{\beta} + \mathbf{u}_t \\ \mathbf{u}_t &= \mathbf{R}_n \mathbf{W}_n \mathbf{u}_t + \boldsymbol{\varepsilon}_t, \qquad t = 1, 2, ..., T \\ E\{\boldsymbol{\varepsilon}_t\} &= \mathbf{0}, \qquad Var\{\boldsymbol{\varepsilon}_t\} = \boldsymbol{\Sigma}_n = diag\{\sigma_1^2, \sigma_2^2, ..., \sigma_n^2\} \\ E\{\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^{'}\} &= \mathbf{0}, \qquad t \neq s. \end{aligned} \tag{2}$$

Here, $\mathbf{y}_t$ is an $(n \times 1)$ vector of observations of the endogenous variable and $\mathbf{x}_t$ an $(n \times k)$ matrix of exogenous regressors. $\mathbf{R}_n$ is an $(n \times n)$ diagonal matrix of spatial dependence coefficients $\{\rho_i, i = 1, ..., n\}$ and $\mathbf{W}_n$ the unknown weight matrix with zeroes in its main diagonal; the nonsingularity condition, $|\mathbf{I} - \mathbf{R}_n \mathbf{W}_n| \neq 0$ applies. Moreover, n is fixed and T is allowed to increase. Like Meen (1996), they suggest using a consistent estimator of the covariance matrix of the error terms obtained from the LS residuals, $\hat{\mathbf{\Gamma}} = T^{-1}\sum_{t=1}^{T}\hat{\mathbf{u}}_t\hat{\mathbf{u}}_t'$, in order to produce unique consistent estimates of matrices $\mathbf{R}_n$, $\mathbf{\Sigma}_n$ and $\mathbf{W}_n$, via the relation, $\hat{\mathbf{\Gamma}}_n = \hat{\mathbf{A}}_n\hat{\mathbf{\Sigma}}_n\hat{\mathbf{A}}_n'$, where $\hat{\mathbf{A}}_n = (\mathbf{I}_n - \hat{\mathbf{R}}_n\hat{\mathbf{W}}_n)^{-1}$. There are identification problems in (2) solved by assuming that (i) the spatial autoregressive parameter is the same for all regions, so that $\mathbf{R}_n\mathbf{W}_n = \rho\mathbf{W}_n$ and (ii) the spatial weights matrix is symmetric.

These two conditions are also essential in the proposal of Beenstock and Felsenstein (2012) for estimating $\mathbf{W}_n$ in a pure panel data SLM model with unobserved random effects:

$$\mathbf{y}_t = \boldsymbol{\eta}_n + \mathbf{R}_n\mathbf{W}_n\mathbf{y}_t + \boldsymbol{\varepsilon}_t \tag{3}$$

$\boldsymbol{\eta}_n$ being an $(n \times 1)$ vector of unobserved random effects assumed orthogonal to the idiosyncratic error terms, $\boldsymbol{\varepsilon}_t$.

3 The Covariance and the Weighting matrix

As is so often the case, tests for the constancy of effects go hand-in-hand with the development of estimation methods. Bhattacharjee and Jensen-Butler (2013) extend the *identity test* of Ledoit and Wolf (2002), LW, to the case of testing for a given driver of spatial diffusion; specifically, their hypothesis is that:

$$H_0 : \mathbf{W} = \mathbf{W}_0 \qquad vs \qquad H_A : \mathbf{W} \neq \mathbf{W}_0. \tag{4}$$

For the case of model (2) the test is constructed by comparing two estimated covariance matrices. The first is a restricted estimator, efficient only under the null hypothesis while it is biased under the alternative:

$$\hat{\mathbf{\Gamma}}_{W_0} = (\mathbf{I} - \hat{\mathbf{R}}_{n,r}\mathbf{W}_0)'^{-1}\hat{\mathbf{\Sigma}}_{n,r}(\mathbf{I} - \hat{\mathbf{R}}_{n,r}\mathbf{W}_0)^{-1}. \tag{5}$$

$\hat{\mathbf{R}}_{n,r}$ and $\hat{\mathbf{\Sigma}}_{n,r}$ are the corresponding restricted estimates. The second, unrestricted estimator is consistent under both null and alternative, but inefficient under the null.

$$\hat{\mathbf{\Gamma}} = (\mathbf{I} - \hat{\mathbf{R}}_{n,u}\hat{\mathbf{W}}_n)'^{-1}\hat{\mathbf{\Sigma}}_{n,u}(\mathbf{I} - \hat{\mathbf{R}}_{n,u}\hat{\mathbf{W}}_n)^{-1}. \tag{6}$$

$\hat{\mathbf{R}}_{n,u}$, $\hat{\mathbf{\Sigma}}_{n,u}$ and $\hat{\mathbf{W}}_n$ are the unrestricted estimates of $\mathbf{R}$, $\mathbf{\Sigma}$ and $\mathbf{W}$.

The LW test is consistent (*under* $T \to \infty$) and has good behaviour for samples of small to moderate size, even for the more difficult case of high-dimensionality ($n > T$); see Ledoit and Wolf (2002) for the details. Unfortunately, the LW test is not adequate to test for breaks in the weighting matrix because $\mathbf{W}_0$ should be specified in advance.

In line with most previous studies, our suggestion is to test for the existence of a breakpoint in $\mathbf{W}$ by using consistent estimates of the corresponding covariance matrix, $\mathbf{\Gamma}$. Let us assume, for the moment, that we know the data generating process, i.e. a static spatial model

$$S_y(\rho_y; \mathbf{W}_y)\mathbf{y}_t = \boldsymbol{\mu} + S_x(\rho_x; \mathbf{W}_x)\mathbf{x}_t\boldsymbol{\beta} + S_u(\rho_u; \mathbf{W}_u)\mathbf{u}_t; \quad t = 1, 2, 3, ...T \qquad (7)$$

where $\mathbf{y}_t$ and $\mathbf{u}_t$ are $(n \times 1)$ vectors and $\mathbf{x}_t$ a $(n \times k)$ matrix of k covariates, $\mathbf{x}_t = (x_{1t}, x_{2t}, ..., x_{kt})$; $\boldsymbol{\beta}$ is a vector of k parameters and $\boldsymbol{\mu}$ a $(n \times 1)$ vector of unobserved individual effects ($\boldsymbol{\mu} \sim i.i.d.\,(\boldsymbol{\mu}_0, \mathbf{\Sigma}_\mu)$). The terms in vector $\mathbf{u}_t$, are spatially and serially uncorrelated, centered on zero and have a diagonal covariance matrix, $\mathbf{\Lambda}$, that is constant through time. Each of the k covariates, $x_i, i = 1, 2, .., k$, conform to a stationary process with mean vector $\boldsymbol{\eta}_{x_i}$ and covariance matrix $\mathbf{\Sigma}_{x_i}$. Moreover, $S_g(\rho_g; \mathbf{W}_g)$, S_g in short, refers to the spatial structure related to g, where ρ_g is a spatial dependence parameter and $\mathbf{W}_g$ an $(n \times n)$ matrix. This structure can be a spatial moving average, an autoregressive process or another spatial mechanism that depends on a spatial driver. The weighting matrices may coincide ($\mathbf{W}_g = \mathbf{W}, \forall\ g$), as usual in applied work, or they may be different. Let us assume that the covariates, unobserved effects and error terms are all independent, $Cov(\boldsymbol{\mu}_t; x_{it}) = Cov(\boldsymbol{\mu}_t; u_t) = Cov(\mathbf{u}_t; x_{it}) = \mathbf{0}; i = 1, 2, .., k; \forall t$. Under these circunstances, the covariance matrix of vector $\mathbf{y}$ is

$$V(\mathbf{y}_t) = \mathbf{\Gamma} = S_y^{-1}\left[\Sigma_\mu + S_x\left(\sum_{i=1}^{k}\beta_i^2\mathbf{\Sigma}_{x_i}\right)S_x^{'} + S_u\mathbf{\Lambda}S_u^{'}\right]S_y^{'-1}; \forall t \qquad (8)$$

Our focus is on the weights matrix. It is clear that, assuming stability in the other elements, we should observe changes in $\mathbf{\Gamma}$ if $\mathbf{W}$ itself also changes. The contrary does not apply given that the covariance matrix depends on many other terms (variances, slope coefficients, etc.). This is a limitation of the procedure.

Assume that a sequence of $\{t = 1, 2, ..., T\}$ time series observations is available for vector $\mathbf{y}_t$. The null hypothesis states that the covariance matrix, $\mathbf{\Gamma}$, is the same for the whole period, whereas the alternative introduces a breakpoint in period T_b, $1 < T_b < T$. In that case, the sample should be divided into two different samples $\{t = 1, ..., T_b\}$ and $\{t = T_b + 1, ..., T\}$, because different weights matrices apply in each period, $\mathbf{W}_1$ and $\mathbf{W}_2$, which in turn produce different covariance matrices, $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$. We want to test if there is a break in T_1.

$$H_0 : \mathbf{\Gamma}_1 = \mathbf{\Gamma}_2 = \mathbf{\Gamma} \qquad vs \qquad H_A : \mathbf{\Gamma}_1 \neq \mathbf{\Gamma}_2. \tag{9}$$

The likelihood ratio test (Anderson, 2003), LR, is one of the most popular tests in this field, for which we need (i) the assumption of a correct, known specification[1], (ii) the assumption of normality and (iii) knowledge about the breakpoints. The LR for the null hypothesis of (9) does not depend on the first order moments:

$$LR = -2 \ln \frac{\left|\widehat{\mathbf{\Gamma}}\right|^{-T/2}}{\left|\widehat{\mathbf{\Gamma}}_1\right|^{-T_1/2} \left|\widehat{\mathbf{\Gamma}}_2\right|^{-T_2/2}} \rightarrow \chi^2 (\frac{n(n+1)}{2}) \tag{10}$$

as $T \rightarrow \infty$ for n fixed.

$\widehat{\mathbf{\Gamma}}$ and $\widehat{\mathbf{\Gamma}}_i$; $i = 1, 2$ are the sample covariance matrices under the null and alternative hypotheses. The statistic of (10) is easily extended to the case of d known breakpoints. The LR is severely affected by the *high-dimensionality problem* which means that: (i)- the LR test degenerates when $n \geqslant T_i$, $i = 1, 2$, because the sample covariance matrices are singular; (ii)- if both dimensions, T and n, are large ($T, n \rightarrow \infty$, but $n < T_i$) the percentage of false rejections increases dramatically. Ledoit and Wolf (2002) refer to the ratio $\frac{n}{T_i}$ as the *concentration* c_i index, which should be considerably lower than 1 (according to their simulation results, less than 0.05). Bai *et al.* (2009) later on show that the LR statistic drifts to infinity almost surely, given that $n \rightarrow \infty$, independently of the *concentration* c_i.

[1] Actually this assumption is not strictly necessary. The LR test, as well as the following tests, is obtained from a covariance matrix which can be the residual covariance matrix or the covariance matrix of $\mathbf{y}$. According to our experience, knowledge about the process improves power/size of the tests.

This observation is important because spatial data usually involve a large n and a finite T. As a remedy to the degeneracy problem, Schott (2007) introduces a Wald test, which only requires the covariance matrix to be non-singular under the null (that is $T > n$).

Bai *et al.* (2009) obtain the corrections needed for the LR statistics, in mean and variance, to behave properly for the case where $(T, n \to \infty)$ with concentration c_i indices lower than 1 in both subsamples. They show that if variables are *i.i.d.*, under the null hypothesis the LR ratio of (10), corrected by a quantity proportional to the dimension of the matrices, asymptotically converges to a Normal distribution with finite first and second order moments; that is:

$$T_N = \frac{(LR - nF_{T_1;T_2}) - m(f)}{v(f)^{1/2}} \to N(0,1) \tag{11}$$

where $nF_{\alpha_1;\alpha_2}$ is the Marchenko-Pastur law of indices α_1 and α_2; $m(f)$ and $v(f)$ are the mean and variance, respectively, of the normal random variable $(LR - nF_{T_1;T_2})$; see Theorem 4.1 and expressions (4.2)-(4.7) in Bai *et al* (2009). Note that the result of (11) applies for general *i.i.d.* distributions, having a fourth order moment; the T_N test can be considered as a generalized pseudo-likelihood ratio test for non-Gaussian data.

In another strand of literature, Schott (2007), Srivastava (2007), Srivastava and Yanagihara (2010), Li and Chen (2012) and Srivastava *et al.* (2014) use the squared Frobenius norm to measure the distance between the null and the alternative hypotheses. The resulting statistics rely on estimates of the distance between the two covariance matrices:

$$tr\left(\mathbf{\Gamma}_1 - \mathbf{\Gamma}_2\right)^2 = tr\mathbf{\Gamma}_1^2 + tr\mathbf{\Gamma}_2^2 - 2tr\mathbf{\Gamma}_1\mathbf{\Gamma}_2 \tag{12}$$

Again, high dimensionality is a serious problem because the sample spectral moments, $tr\widehat{\mathbf{\Gamma}}_i^j;\ i = 1, 2$, are poor estimates of the corresponding moments $tr\mathbf{\Gamma}_i^j;\ i = 1, 2$.

Schott (2007) obtains the so called t_{Tn} statistic which is an unbiased measure of the distance in (12). Assuming the data to be *i.i.d.* normal, the sample covariance matrix conforms to a Wishart distribution, $T\hat{\mathbf{\Gamma}}_i \sim W_n(\mathbf{\Gamma}_i; T)$, and the t_{Tn} statistic (defined in expression 1, p. 6536, of Schott, 2007) converges, under the null, to a normal distribution with mean 0 and a well-defined finite variance (θ^2 detailed in expression 4, p. 6538). Asymptotic normality of t_{Tn} holds if T_i and n approach infinity at a regular rate (that is,

$\lim_{h\to\infty} \frac{n}{T_{ih}} = b_i$ where $b_i \in [0;\infty)$ and $b_i > 0$ for at least one i). Again the test can be extended to the case of comparing d breakpoints. The Schott test has been refined recently by Srivastava *et al.* (2014) into the T_3 test, which is based on more efficient sampling estimates of spectral moments; also, it is robust to departure from normality.

Srivastava (2007) uses a lower bound of the Frobenius norm: $tr\left(\mathbf{\Gamma}_1 - \mathbf{\Gamma}_2\right)^2 \geqslant \left(\sqrt{tr\mathbf{\Gamma}_1^2} - \sqrt{tr\mathbf{\Gamma}_2^2}\right)^2 \geqslant 0$, testing that $tr\mathbf{\Gamma}_1^2 - tr\mathbf{\Gamma}_2^2 = 0$. Like Schott (2007), he considers consistent estimates of the traces of both matrices and a consistent estimate of the variance of the difference of the traces (more details in Srivastava, 2007). The statistic obtained, called T_2, is asymptotically distributed N(0,1) under the same conditions as in Schott's t_{Tn}, that is $(T;n) \to \infty$, and *i.i.d.* normality. In Srivastava and Yanagihara (2010) the T_2 statistic is transformed into a standardized version, called Q_2, defined as $tr\mathbf{\Gamma}_1^2/\left(tr\mathbf{\Gamma}_1\right)^2 - tr\mathbf{\Gamma}_2^2/\left(tr\mathbf{\Gamma}_2\right)^2$. Both test statistics, Q_2 and T_2, can also be extended to the d breakpoints case.

Li and Chen (2012) address the '*large n, small T*' question using a different, more direct approach. They propose to streamline terms in the traces, $tr\widehat{\mathbf{\Gamma}}_i^j;\ i = 1,2$, so as to make unbiased estimators by using U statistics. In fact, the statistic that they propose is an estimate of the *trace* of (12), $L_{T_1;T_2} = A_{T_1} + A_{T_2} - 2C_{T_1;T_2}$ where A_{Ti} and $C_{T_1;T_2}$ are unbiased estimates of the respective *traces* (details in Li and Chen, 2012, p. 910-911). Under the null, $L_{T_1;T_2}$ is centred on zero, has a well-defined finite variance, $\sigma^2_{T_1;T_2}$, and supports a CLT so that $L_{T_1;T_2} = \frac{T_{T_1;T_2}}{\sigma_{T_1;T_2}} \overset{D}{\to} N(0,1)$. Note that the $L_{T_1;T_2}$ of Li and Chen (2012) does not require the assumption of normality but is restricted to the case of a single breakpoint.

Finally, the approach of García (2012) is based on the eigenvectors of the two covariance matrices. The idea is that, under the null, the eigenvectors obtained for either of the two sample covariance matrices will explain a similar amount of variation in either of the two samples. Denote by $\mathbf{E}_1$ and $\mathbf{E}_2$ the orthonormal matrices of eigenvectors corresponding to $\hat{\mathbf{\Gamma}}_1$ and $\hat{\mathbf{\Gamma}}_2$, which may be combined with the matrices of data observed in both subsamples, $\mathbf{y}_j$ (or the residuals, $\hat{\mathbf{u}}_j$, if a SEM model has been previously estimated) to form the four matrices:

$$\mathbf{P}_{jk} = [\mathbf{p}_{1jk}, \mathbf{p}_{2jk}, ..., \mathbf{p}_{njk}] = \mathbf{y}_j\mathbf{E}_k; \quad j,k = 1,2 \tag{13}$$

The eigenvalues of $\hat{\mathbf{\Gamma}}_1$ and $\hat{\mathbf{\Gamma}}_2$ are calculated as the sums of squares of elements of $\mathbf{p}_{jj}$, $(j = 1, 2)$. Now define the sums of squares of the elements of the columns of $\mathbf{P}_{jk}$ to be $\{v_{ijk}, i = 1, 2, ..., n\}$ and a new measure of distance between the two covariance matrices:

$$S_1 = 2\sum_{i=1}^{n}[(v_{i11} - v_{i21})^2 + (v_{i12} - v_{i22})^2] \tag{14}$$

No distribution has been provided for S_1. Garcia (2012) proposes the following resampling permutation approach:

(i) Obtain the statistic S_1 for the observed data.

(ii) Resample with replacement, under the null of (9); take 50% of the resampled observations from the first sub-sample, the other 50% from the second sub-sample, and reassign them randomly to the first or to the second sub-sample.

(iii) Calculate the sequence of resampled statistics $\{S_1^g, \quad g = 1, 2, ..., G\}$, G being the number of resampled samples.

(iv) Compute the empirical p-value of the resampling experiment as:

$$pval(S_1) = \frac{1}{G}\sum_{g=1}^{G}\mathbb{I}(S_1^g > S_1) \tag{15}$$

where $\mathbb{I}(A)$ is the indicator function of event, A.

(v) Reject the null hypothesis of equal covariance matrices, (9), if: $pval(S_1) < \alpha$ for some chosen nominal significance level, α.

In case of rejecting the null, we may use the statistics S_2 and S_3 (defined in Garcia, 2012) to assess the characteristics of the break. S_3 evaluates the proportion of the break due to heteroskedasticity whereas S_2 measures the proportion attributable to changes in the covariances among the individuals.

4 Monte Carlo evidence

Our experimental DGP is a spatial panel data model, whose weights matrix remains stable under the null; it may be either a spatial error model, SEM, or a spatial lag model, SLM. In the first case:

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \mathbf{x}_t\boldsymbol{\beta} + (\mathbf{I} - \rho_u\mathbf{W})^{-1}\mathbf{u}_t; \quad t = 1, 2, 3, ...T \tag{16}$$

The model of the alternative hypothesis incorporates two SEM equations with different weighting matrices (similarly for the SLM case):

$$
\begin{aligned}
\mathbf{y}_t = \boldsymbol{\mu}_t + \mathbf{x}_t\boldsymbol{\beta} + (\mathbf{I} - \rho_u \mathbf{W}_1)^{-1}\mathbf{u}_t; &\quad t = 1, 2, ...T_b \\
\mathbf{y}_t = \boldsymbol{\mu}_t + \mathbf{x}_t\boldsymbol{\beta} + (\mathbf{I} - \rho_u \mathbf{W}_2)^{-1}\mathbf{u}_t; &\quad t = T_b + 1, ...T \\
\mathbf{W}_1 \neq \mathbf{W}_2 &
\end{aligned} \tag{17}
$$

The list of cases from which we extract issues for discussion comprises small to medium sample sizes in n and medium to large sample sizes in T, allowing for cases of high dimensionality. The tests are expected to work better with increasing T. Bai *et al.* (2009) report results with very large values of T, up to $12,800$ observations, which confirm the consistency of their T_N test; however this will seldom be realistic in applied work. Schott (2007), on the contrary, simulates very short time spans, $T = 4$, with serious dimensionality contraints ($n = 128$). The parameters for our Monte Carlo are the following:

- Time-dimension: sample sizes $T = \{10; 30; 60; 150; 250; 500\}$.
- Space-dimension: sample sizes $n = \{16; 36; 64; 144; 225; 400\}$.
- Regional shape: regular square lattice with side $n^{1/2}$.
- Weighting matrices. The weights matrix for the first period follows a rook contiguity pattern, $\mathbf{W}^R$. We distinguish two cases for the second period under the alternative: a queen contiguity pattern, $\mathbf{W}^Q$, and a pattern based on the inverse of the distances between the centroids of the cells, $\mathbf{W}^D$. The three matrices are row-standardized. The rook pattern is more akin to the queen case (this is a soft change in the weights in terms of Lee and Yu, 2012) than to the inverse of the distance; i.e., the Frobenius norm of the matrix $\mathbf{W}^R - \mathbf{W}^Q$, $n = 16$, is 1.50 but 1.75 in the case of $\mathbf{W}^R - \mathbf{W}^D$.
- Known or unknown breakpoints. The breakpoints, T_b, are located in the middle of the time sequence. Then we simulate two cases. In the first, it is supposed that we know where the breakpoint is located whereas this point is unknown in the second. In that case, the tests are obtained following a rolling process centred on T_b; the sequence includes 20% of the central observations according to the examples in Table 1 (T_a and T_c are the starting and ending points for the rolling sequence):

Table 1: Sample size, location of breakpoint,and testing intervals			
T	T_a	T_b	T_c
60	*24*	*30*	*36*
150	*61*	*75*	*90*
500	*201*	*250*	*300*

- Nature of the break, instantaneous or gradual. In the first case, the break takes place as in (17). For the second case, we define a '*breaking period*' as the time span along which change is happening; this period is equal to 10% of the sample, $T^{bp} = 0.1 \times T$. The weighting matrix that intervenes in the DGP for this transitional interval changes gradually in each period according to $\mathbf{W}^{bp}(t) = \frac{T_b - t}{T^{bp}} \times \mathbf{W}^R + (1 - \frac{T_b - t}{T^{bp}}) \times \mathbf{W}^i$ with $i \in \{Q, D\}$. That is, the transition from one matrix to the other begins in period $\left(T_b - T^{bp}\right)$ and it is not fully completed until T_b.

- Knowledge about the DGP. Generally, the DGP remains unkown to the user so the tests have been obtained for the covariance matrix of $\mathbf{y}$. Assuming that the spatial structure appears only in the errors of the equation (that is, is a SEM model) and that this is known, the user can benefit from a previous consistent estimation of the mean equation of the model, from which the corresponding (consistent) residuals can be obtained. In this framework, the tests should be applied to the covariance matrix of the residuals. Results are briefly discussed in Appendix A.

- Spatial parameter values $\rho \in \{0.1; 0.3; 0.5; 0.7; 0.9\}$

- Unobserved individual and idiosyncratic disturbances are drawn from standard normal distributions: $\mu_i \sim i.i.d.N(0, \sigma^2_\mu)$, $u_{it} \sim i.i.d.N(0, \sigma^2_u)$, $\sigma^2_\mu = 1$, $\sigma^2_u = 1$

- Exogenous regressor dimension, $k = 3$; $x_{ti1} = 1$; x_{ti2} and x_{ti3} drawn from $N(0; I_2)$ for $t = 1, 2, ..., T$; $i = 1, 2, ..., n$.

- Regression parameter values $\boldsymbol{\beta}' = (1, 1, \beta_3)$. The value of β_3 has been defined in order to assure a given signal-to-noise ratio in the pooled regression (with no spatial effects) according to the following Table:

Table 2: Expected R^2 and value of β_3			
	$R^2 \simeq 0.5$	$R^2 \simeq 0.6$	$R^2 \simeq 0.8$
$\beta_3 = \sqrt{\frac{\beta_2^2(1-R^2)\sigma_{x_2}^2 + R^2\left(\sigma_u^2+\sigma_\mu^2\right)}{(1-R^2)\sigma_{x_3}^2}}$	1.7	2	3

- For estimating size and power, we use 1,000 replications, which produce approximate standard errors for the 10%, 5% and 1% significance level of 0.00948, 0.00690 and 0.00314 respectively.

- Two of the tests, the LR of (10) and the T_N of (11) are based on the likelihood ratio approach, five other tests (Schott's t_{T_n}, the T_3 refined version of Srivastava *et al.*, Srivastava's T_2, Srivastava and Yanagihara's Q_2 and Li and Chen's $L_{T_1;T_2}$) are based on the squared Frobenius norm and, finally, there is Garcia's S_1 bootstrap test. For the likelihood ratio tests we need a concentration index smaller than 1, ($c_i = \frac{n}{T_i} < 1$) so there are cases where the LR based tests cannot be obtained (the same applies for Garcia's S_1); this is not the case for the Frobenius norm based tests.

4.1 Results of the Monte Carlo

Tables 3 and 4 and Figures 1 to 3 summarize the main results of the Monte Carlo[2]. The Tables correspond to the case of a breakpoint with known location whereas the Figures describe the case of a breakpoint with unknown location. The data that appear in the Tables are the percentage of rejections of the null hypothesis of (9), using the corresponding test. Figures represent the evolution of the test values over the testing interval in the two cases of instantaneous and progressive break in the weights matrix.

Tables 3a,b report estimated size, whereas Tables 4a,b,c,d show estimated power. The impact of the signal-to-noise-ratio is really small, both in size and power. Results tend to be slightly better for the case of a high signal-to-noise ratio but the difference is always below 0.05 points, so we focus on the case of a high signal-to-noise ratio.

[2]For the sake of brevity, we omit the details for the cases not shown in the Tables. The main tendencies, both in size as well in power, are maintained. Full details are available from the authors, upon request

Table 3a: Estimated size (significance level=0.05). Known breakpoint. SEM process
High signal-to-noise ratio. Matrix in the DGP: WR.

ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
LR	-	-	-	0.84	-	-	0.62	0.75	-	0.34	0.41	-
T_N	-	-	-	0.14	-	-	0.07	0.20	-	0.06	0.10	-
t_{Tn}	0.03	0.03	0.04	0.05	0.03	0.02	0.06	0.04	0.06	0.06	0.04	0.04
$L_{T1,T2}$	0.27	0.45	0.61	0.19	0.29	0.35	0.16	0.25	0.30	0.14	0.22	0.26
T_2	0.00	0.00	0.03	0.01	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00
T_3	0.01	0.02	0.09	0.05	0.00	0.00	0.06	0.00	0.00	0.06	0.02	0.01
Q_2	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S_1	-	-	-	0.62	-	-	0.46	0.61	-	0.41	0.55	-
ρ=0.5												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
LR	-	-	-	0.86	-	-	0.59	0.71	-	0.31	0.38	-
T_N	-	-	-	0.13	-	-	0.07	0.21	-	0.05	0.10	-
t_{Tn}	0.04	0.03	0.03	0.05	0.05	0.04	0.07	0.04	0.04	0.06	0.06	0.04
$L_{T1,T2}$	0.22	0.41	0.59	0.21	0.25	0.38	0.17	0.27	0.27	0.11	0.19	0.24
T_2	0.01	0.00	0.03	0.01	0.00	0.00	0.02	0.01	0.00	0.04	0.00	0.01
T_3	0.02	0.01	0.08	0.04	0.00	0.00	0.05	0.00	0.00	0.06	0.01	0.01
Q_2	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S_1	-	-	-	0.56	-	-	0.49	0.58	-	0.44	0.51	-
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
LR	-	-	-	0.81	-	-	0.42	0.55	-	0.18	0.22	-
T_N	-	-	-	0.10	-	-	0.07	0.22	-	0.06	0.09	-
t_{Tn}	0.04	0.04	0.02	0.04	0.04	0.04	0.06	0.05	0.05	0.07	0.05	0.05
$L_{T1,T2}$	0.20	0.42	0.52	0.20	0.20	0.32	0.15	0.22	0.20	0.13	0.18	0.23
T_2	0.01	0.01	0.01	0.02	0.00	0.00	0.02	0.00	0.00	0.04	0.01	0.00
T_3	0.01	0.00	0.08	0.04	0.00	0.00	0.06	0.00	0.00	0.07	0.02	0.01
Q_2	0.00	0.01	0.01	0.02	0.01	0.00	0.01	0.00	0.01	0.02	0.00	0.00
S_1	-	-	-	0.44	-	-	0.42	0.55	-	0.39	0.53	-

*: For the cases of the LR, T_N and S_1 tests, the percentages shown in this panel correspond to n=64.

Table 3b: Estimated size (significance level=0.05). Known breakpoint. SLM process High signal-to-noise ratio. Matrix in the DGP: WR.

ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
LR	-	-	-	0.88	-	-	0.66	0.76	-	0.41	0.48	-
T_N	-	-	-	0.11	-	-	0.08	0.21	-	0.05	0.07	-
t_{Tn}	0.05	0.02	0.03	0.06	0.04	0.03	0.07	0.05	0.05	0.06	0.05	0.05
$L_{T1,T2}$	0.35	0.52	0.73	0.29	0.32	0.38	0.24	0.29	0.32	0.19	0.20	0.28
T_2	0.01	0.00	0.02	0.02	0.00	0.00	0.03	0.04	0.03	0.05	0.04	0.03
T_3	0.03	0.00	0.10	0.05	0.00	0.00	0.06	0.01	0.00	0.05	0.02	0.01
Q_2	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00
S_1	-	-	-	0.66	-	-	0.48	0.65	-	0.43	0.54	-
ρ=0.5												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
LR	-	-	-	0.82	-	-	0.65	0.71	-	0.51	0.60	-
T_N	-	-	-	0.12	-	-	0.06	0.23	-	0.05	0.09	-
t_{Tn}	0.05	0.03	0.04	0.06	0.04	0.04	0.07	0.05	0.05	0.07	0.05	0.05
$L_{T1,T2}$	0.35	0.39	0.51	0.28	0.32	0.35	0.31	0.35	0.39	0.19	0.28	0.36
T_2	0.02	0.00	0.00	0.02	0.00	0.00	0.05	0.00	0.00	0.05	0.01	0.01
T_3	0.04	0.00	0.13	0.05	0.00	0.00	0.07	0.02	0.00	0.07	0.02	0.02
Q_2	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00
S_1	-	-	-	0.59	-	-	0.39	0.47	-	0.40	0.48	-
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
LR	-	-	-	0.79	-	-	0.36	0.46	-	0.27	0.31	-
T_N	-	-	-	0.13	-	-	0.07	0.24	-	0.05	0.08	-
t_{Tn}	0.06	0.04	0.03	0.07	0.04	0.04	0.08	0.05	0.04	0.08	0.05	0.05
$L_{T1,T2}$	0.28	0.31	0.48	0.29	0.37	0.39	0.24	0.29	0.33	0.19	0.23	0.27
T_2	0.02	0.01	0.00	0.03	0.01	0.00	0.05	0.01	0.01	0.03	0.03	0.02
T_3	0.08	0.00	0.00	0.07	0.01	0.00	0.09	0.03	0.01	0.09	0.04	0.03
Q_2	0.00	0.00	0.00	0.02	0.00	0.00	0.02	0.02	0.01	0.02	0.01	0.02
S_1	-	-	-	0.46	-	-	0.27	0.31	-	0.20	0.24	-

*: For the cases of the LR, T_N and S_1 tests, the percentages shown in this panel correspond to n=64.

Table 4a: Estimated power (significance level=0.05). Known breakpoint. SEM process
High signal-to-noise ratio. Matrices in the DGP: WR vs WD

ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.16	-	-	0.37	0.51	-	0.51	0.72	-
t_{Tn}	0.13	0.23	0.34	0.14	0.44	0.54	0.47	0.61	0.68	0.59	0.70	0.75
T_2	0.08	0.09	0.21	0.10	0.15	0.26	0.18	0.27	0.30	0.25	0.31	0.41
ρ=0.5												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.35	-	-	0.46	0.63	-	0.60	0.78	-
t_{Tn}	0.15	0.31	0.52	0.25	0.49	0.60	0.51	0.58	0.71	0.67	0.75	0.86
T_2	0.10	0.10	0.27	0.12	0.19	0.40	0.25	0.31	0.51	0.34	0.53	0.62
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.48	-	-	0.62	0.73	-	0.79	0.88	-
t_{Tn}	0.18	0.43	0.63	0.27	0.54	0.65	0.47	0.66	0.75	0.43	0.76	0.92
T_2	0.10	0.16	0.33	0.15	0.22	0.45	0.19	0.44	0.62	0.37	0.71	0.81

*: For the case of the T_N test, the percentages shown in this panel correspond to n=64.

Table 4b: Estimated power (significance level=0.05). Known breakpoint. SLM process
High signal-to-noise ratio. Matrices in the DGP: WR vs WD.

ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.17	-	-	0.26	0.46	-	0.54	0.85	-
t_{Tn}	0.07	0.16	0.26	0.11	0.27	0.44	0.31	0.56	0.58	0.62	0.92	0.92
T_2	0.04	0.14	0.31	0.10	0.27	0.48	0.31	0.96	0.98	0.50	1.00	1.00
ρ=0.5												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.35	-	-	0.83	0.87	-	0.99	1.00	-
t_{Tn}	0.17	0.33	0.53	0.32	0.83	0.87	0.83	1.00	1.00	0.99	1.00	1.00
T_2	0.11	0.46	0.58	0.26	1.00	1.00	0.53	1.00	1.00	0.74	1.00	1.00
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.75	-	-	1.00	1.00	-	1.00	1.00	-
t_{Tn}	0.19	0.93	0.98	0.34	1.00	1.00	0.90	1.00	1.00	1.00	1.00	1.00
T_2	0.10	0.86	0.95	0.15	1.00	1.00	0.27	1.00	1.00	0.35	1.00	1.00

*: For the case of the T_N test, the percentages shown in this panel correspond to n=64.

Table 4c: Estimated power (significance level=0.05). Known breakpoint. SEM process High signal-to-noise ratio. Matrices in the DGP: WR vs WQ												
ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.16	-	-	0.37	0.51	-	0.51	0.72	-
t_{Tn}	0.08	0.14	0.27	0.17	0.25	0.31	0.24	0.32	0.39	0.27	0.42	0.58
T_2	0.09	0.10	0.16	0.12	0.16	0.24	0.15	0.21	0.25	0.19	0.25	0.33
ρ=0.5												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.35	-	-	0.46	0.63	-	0.60	0.78	-
t_{Tn}	0.15	0.22	0.28	0.20	0.34	0.40	0.29	0.41	0.51	0.41	0.61	0.66
T_2	0.10	0.10	0.17	0.11	0.14	0.34	0.16	0.23	0.37	0.28	0.35	0.49
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.48	-	-	0.62	0.73	-	0.79	0.88	-
t_{Tn}	0.18	0.34	0.43	0.24	0.44	0.55	0.32	0.53	0.64	0.49	0.70	0.86
T_2	0.10	0.12	0.30	0.19	0.23	0.46	0.19	0.39	0.56	0.33	0.61	0.74
*: For the case of the T_N test, the percentages shown in this panel correspond to n=64.												

Table 4d: Estimated power (significance level=0.05). Known breakpoint. SLM process High signal-to-noise ratio. Matrices in the DGP: WR vs WQ												
ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.17	-	-	0.38	0.53	-	0.47	0.68	-
t_{Tn}	0.07	0.16	0.35	0.21	0.28	0.41	0.27	0.47	0.54	0.34	0.52	0.82
T_2	0.08	0.24	0.31	0.10	0.37	0.48	0.32	0.57	0.67	0.49	0.67	0.83
ρ=0.5												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.24	-	-	0.46	0.48	-	0.78	0.74	-
t_{Tn}	0.11	0.14	0.22	0.26	0.34	0.51	0.33	0.62	0.80	0.39	0.69	0.89
T_2	0.12	0.26	0.35	0.29	0.40	0.55	0.46	0.65	0.84	0.43	0.82	0.90
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.34	-	-	0.74	0.86	-	0.98	0.99	-
t_{Tn}	0.12	0.27	0.31	0.31	0.41	0.55	0.42	0.67	0.85	0.47	0.85	0.91
T_2	0.14	0.26	0.53	0.15	0.49	0.61	0.49	0.70	0.87	0.65	0.94	0.96
*: For the case of the T_N test, the percentages shown in this panel correspond to n=64.												

Size is a real problem for three cases: LR, $L_{T_1;T_2}$ of Li and Chen and S_1 of Garcia. These tests are strongly oversized, independently of (time, spatial) sample size or the degree of spatial interaction. Data on Bai *et al*'s T_N test corroborate the warning of Ledoit and Wolf (2002) with respect to LR type tests: these tests have a size problem when the concentration index is high. According to our results, for the T_N test to have reasonable empirical size the ratio of time observations to spatial units should be at least 10 to 1, so that $c_i \leq 0.10$. The other candidates work pretty well, especially Schott's t_{Tn} test whose estimated size is always around the theoretical 5% significance level. The tests using a lower bound to the Frobenius norm, T_2 and Q_2, are consistently under-sized, with size almost zero in most cases. However this is not, necessarily, a bad thing if accompanied by good power. Unfortunately, the estimated power function of the second is unacceptable (not shown in the tables) and we decided to exclude the Q_2 test from the analysis.

Another surprise comes from the T_3 test of Srivastava *et al.* (2014), a refinement of Schott's t_{Tn} test. The revision does not work in our framework where the T_3 test tends to be undersized (although for small spatial samples, its estimated size is moderately *above* the nominal significance level). However, like the Q_2 test, its estimated power function is unacceptable for almost all the cases (not shown in the paper), no matter sample size, spatial dependence or type of break. Consequently we have decided to exclude also the T_3 test from the analysis.

Some preliminary conclusions can be drawn from this part of the Monte Carlo:

- The LR, the Li and Chen $L_{T_1;T_2}$ and Garcia's S_1 tests are not appropriate tests for changes in the weights matrix of a spatial model. They are strongly affected by the high dimensionality problem which results in unacceptable size inflation.

- Neither can we recommend Srivastava *et al.*'s T_3 or Srivastava and Yanagihara's Q_2 test. These tests are consistently undersized and their estimated power function is unacceptable for the case we are studying. This lack of power is not corrected by increasing (time, spatial) sample size nor after introducing higher spatial dependence.

- Bai *et al*'s T_N test is affected by the high dimensionality problem, requiring a concentration index no greater than 0.1. Under these conditions, the empirical size of the test appears to be approximately correct (in spite of slight oversizing) and its power function increases monotonically with T, time span, and the value of the spatial autocorrelation coefficient.

- This monotonicity is maintained in the cases of the t_{Tn} of Schott and T_2 of Srivastava. The two tests attain good power especially in cases of high time spans. Also, their power increases quickly for cases of fixed T and large n especially if there is a medium to large coefficient of spatial dependence in the DGP.

- As expected, the three tests under consideration (T_N, t_{Tn} and T_2) work considerably better for SLM processes, which induce global patterns with stronger symptoms of dependence. Srivastava's T_2 appears slightly superior in the case of SLM processes whereas Schott's t_{T_N} seems preferable for dectecting breaks in local processes, like the SEM. Greater differences in the weight matrices imply higher power to detect the change as is clear by comparing results in Tables 4a,b with those of Table 4c,d.

Figures 1 to 4 show the behaviour of the three selected tests when the location of the breakpoint is not known. Each graph reproduces the average value of the corresponding statistic (vertical axis) for the hypothesis that there is a break in period t, in the horizontal axis. The breakpoint is located in period 75, but the user does not know that. The tests are evaluated sequentially in a rolling process initiated in period 60 and finished in period 90 (65 and 85 in Bai *et al.*'s T_N test).

Figure 1 describes the case of a change from a rook type pattern to a system of weights based on the inverse of distance, in a SLM model (top panel) and in SEM model (bottom panel). Figure 2 corresponds to a softer change, from a rook pattern to a queen pattern.

Overall, the T_N test works pretty well rejecting the null of equal covariance matrices in the whole interval inspected. The value of the test increases with the coefficient of spatial dependence but diminishes with the number of cross-sectional units (a reflection of the dimensionality problem). A U-pattern emerges in cases of a high concentration index ($n = 64$ and $c_i = 0.85$) combined with weak symptoms of spatial dependence. The maximum value of the sequence of T_N tests roughly coincides with the location of the breakpoint, in period 75. As expected, the procedure works significantly worse in the case of SEM processes and small changes in the weights matrix.

Figure 1: Unknown, instantaneous breakpoint. High signal-to-noise ratio. Matrices in the DPG: WR vs WD. T=150

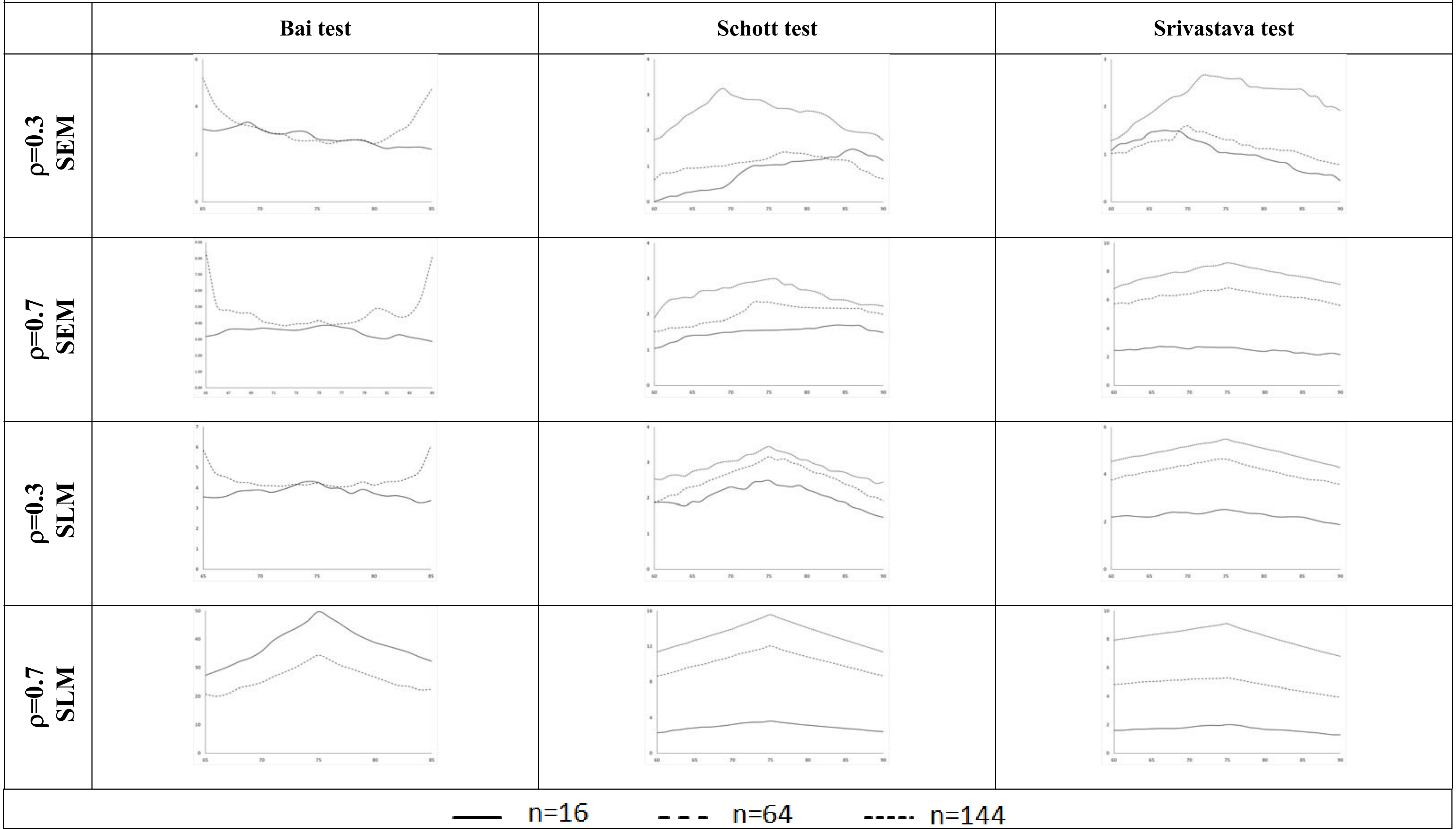
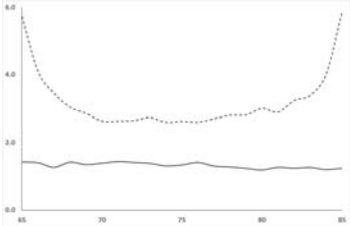
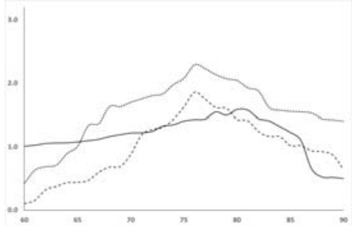
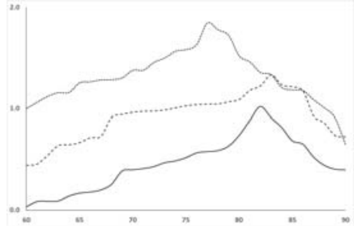
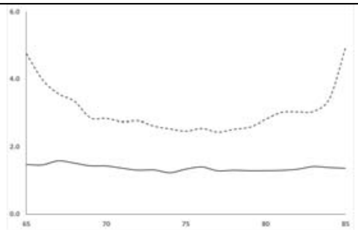
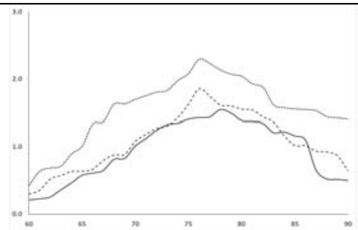
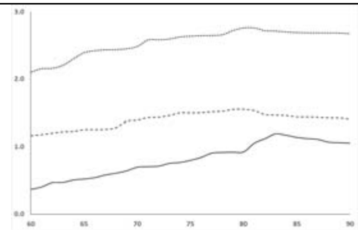
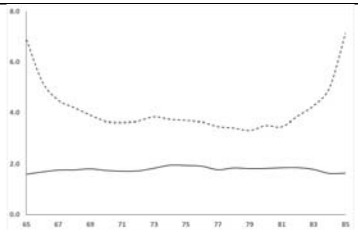
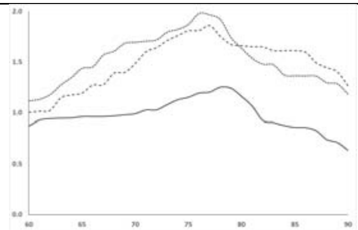
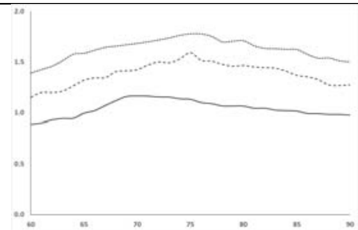
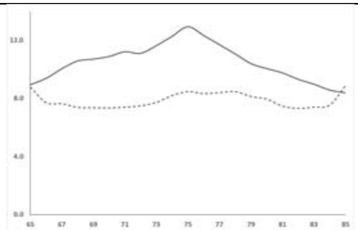
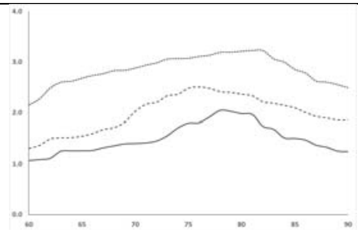
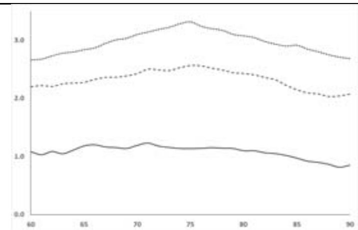
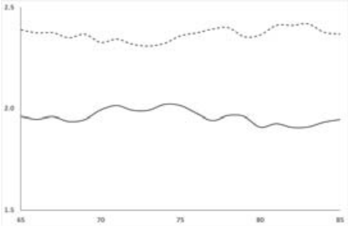
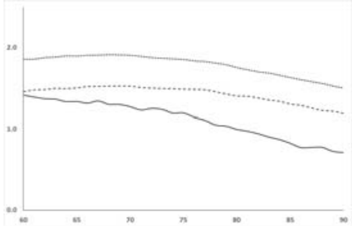
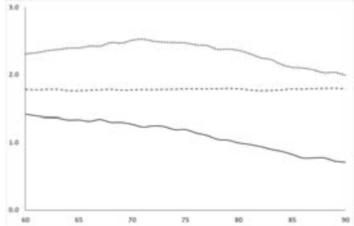
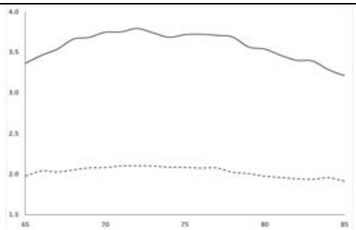
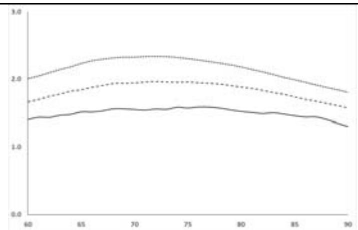
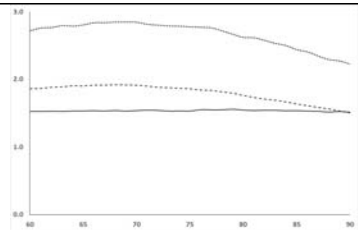
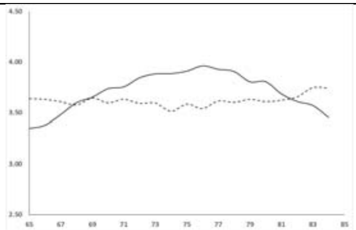
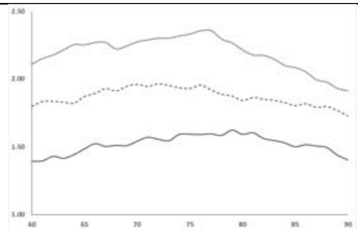
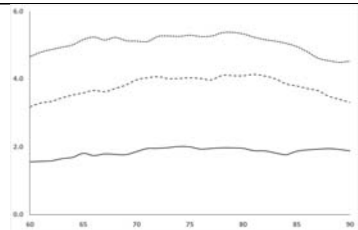
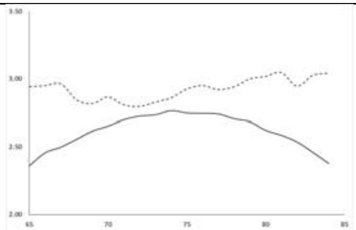
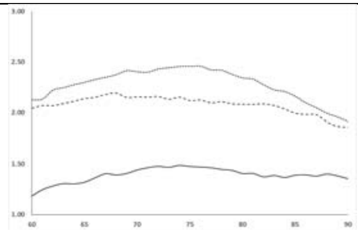
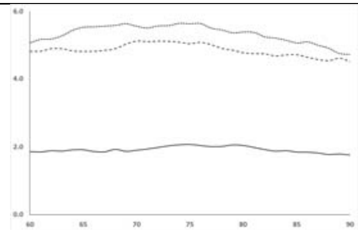


Figure 2: Unknown, instantaneous breakpoint. High signal-to-noise ratio. Matrices in the DPG: WR vs WQ. T=150

	Bai test	**Schott test**	**Srivastava test**
ρ=0.3 SEM			
ρ=0.7 SEM			
ρ=0.3 SLM			
ρ=0.7 SLM			

—— n=16 - - - n=64 ----- n=144

Figure 3: Unknown, gradual breakpoint. High signal-to-noise ratio. Matrices in the DPG: WR vs WD. T=150

	Bai test	Schott test	Srivastava test
$\rho=0.3$ SEM			
$\rho=0.7$ SEM			
$\rho=0.3$ SLM			
$\rho=0.7$ SLM			

——— n=16 - - - n=64 ----- n=144

The other two tests, the t_{Tn} of Schott and the T_2 of Srivastava, maintain a similar profile. Both work better with changes in the weights matrix of a SLM process and where the differences between the two matrices are substantial. The identification of the location of the breakpoint is effective in the SLM case and more uncertain under SEM processes. Smaller (spatial) sample sizes, lower values in the spatial autocorrelation coefficient, local processes of dependence or soft changes in the weights matrix are factors that worsen the functioning of these tests.

The discussion is completed by reviewing Figure 3 which illustrates the case of a gradual change in the weights matrix. Now the change from one matrix to the other is not instantaneous but gradual. The transition begins in period $t = 67$ and it is not fully completed until $t = 82$. As before, the tests are obtained in a rolling process from observations 60 to 90 (65 to 85 for the Bai *et al.* test). The T_N test works pretty well, in general, rejecting the null of equal covariance matrices in the whole testing period. The impact of the high-dimensionality problem affects the power curves of the test for the case of n=64, which now is very flat. The two other tests tend to work similarly well. The Schott test appears to be more sensitive to the gradual change of the weights matrix than the Srivastava test. As before, the change is more difficult to detect in the case of SEM processes and for low values of the spatial correlation coefficient. Finally, let us note the smoother curves obtained for the three tests. In all the cases, the estimated power functions are a bit more diffuse around a local peak than if the change is instantaneous; moreover, this local peak is next to the point where the transition process has completed half the total change ($T = 75$ once again).

5 Case studies

5.1 Case study I: Unemployment in Spanish regions

A distinctive feature of the Spanish economy is the high unemployment, whose spatial dimension is characterized by the persistent and remarkable differences between regions (e.g., Lopez-Bazo *et al*, 2005).

We can provide different explanations for such disparities. Some rely on the assumption of equilibrium, others on the notion of disequilibrium. According to Marston (1985), unemployment is a steady-state function of factor endowment. As long as they are stable through time but different between regions, we cannot expect rapid changes in the short run. From this perspective, unemployment is basically a local question that reflects imbalances in each regional labour market. The neoclassical paradigm posits the existence of a competitive equilibrium among the regions, where the unemployment rate will level off (Blanchard and Katz, 1992). In the short run, regional disparities result from market rigidities or mobility restrictions that should disappear in the long-run. The adjustment process will be faster or slower depending on regional circumstances, that may persist for a long time. Partridge and Rickman (1997) merge both approaches by combining disequilibrium factors (e.g., economic growth, technology) and equilibrium mechanisms, (e.g., industrial composition, wages) with demographic trends, locational amenities and institutions.

Furthermore, there is an extensive literature (e.g., Zeilstra and Elhorst, 2014) showing that unemployment rates are not only driven by intra-regional factors; extra-regional factors have also a role. Beyer and Smets (2015), from a slightly different perspective, highlight the importance of common factors in the European case (related, for example, to country effects) which strengthens the interdependence of regional unemployment figures.

The interest of the Spanish case is reinforced by the institutional framework. Franco's dictatorship ended in the 70s (Franco died in 1975). Then, the country embarked on a transition from a highly centralized system to a quasi-federal structure with 17 regions, Autonomous Communities (CCAA from now on). The set up of the so-called "*Autonomic State*" was finished in 1983. The decentralisation process included own-source revenues for sub-central governments and tax sharing agreements between the regions and the central government.

Actually, CCAA have substantial control over tax rates and on different legislative areas such as education (transferred between 1995 and 1999), public health (transferred in 2002), labour markets, consumers and commercial activities, etc. Key dates for the new system are the agreements of 1986, 1992, 1996, 2001 and 2009, clearly visible in Figure 4 which depicts the percentage of public expenditure under control of the regions. Currently, regional governments are responsible for 45%, approximately, of total public expenditure and almost 60% of government employment.

Spanish regional unemployment combines two important features for us: it is expected to exhibit a strong cross-sectional dependence and, probably, this structure has suffered changes because of the decentralization process. We are going to study this case for the period 1980:1 to 2013:4, using quarterly data.

Figure 4. Percentage of public expenditure under control of the CCAA

source: Gil-Serrate *et al.* (2011)

We follow the mixed approach to labour markets of Partridge and Rickman (1997) but the selection of variables has been highly constrained by data availability. A key element is unemployment by CCAA, taken from the 'Encuesta de Población Activa', Working Population Survey (Instituto Nacional de Estadística, INE, several years). Wages and economic activity in the region are the main factors driving unemployment. Quarterly Labour Cost Survey (INE, several years) is the source of the former and Spanish Regional Accounts (INE, several years) that of the second. INE produces only annual estimates of regional GDP that have been disaggregated into quarterly data according to DiFonzo (1990), using regional employment and regional industrial production indices as high-frequency indicators.

The three panel series are I(1), as indicated in Table 6, where lp stands for the log of the regional gross domestic product, lu is the log of regional unemployment and lw the log of regional wages. $PMSB$ is the panel modified Sargan-Barghava test of Bai and Ng (2010) for testing non-stationarity in the idyosincratic component of the panel series, MQ_c is the Bai and Ng (2004) test to determine the number of common stochastic trends in the common factors of the panel series (whose final number is indicated as r in the table); $CIPS^*$ is the Pesaran (2007) test for panel unit roots; t_a, and t_b

are Moon and Perron (2004) tests for panel unit roots and $n.c.f.$ indicates the number of common factors in the Moon-Perron tests as determined by the Akaike Information Criteria. All the test include individual effects and time trend. $PMSB$, $CIPS^{*}$, t_a, and t_b are asymptotically distributed as standard normals (p-value in brackets) whereas the critical values for the MQ_c statistic appear in Table 1 of Bai and Ng (2004). The Moon-Perron tests cast some doubt on the nonstationarity of lu and lw which, however, are overwhelmingly offset by the other two tests (the Choi tests, not in the table, corroborate this conclusion).

Table 6. Panel unit root tests

$H_0 : I(1)\ vs\ H_A : I(0)$						
	lp		lu		lw	
$PMSB$	-0.4683	(0.3200)	-0.4831	(0.3147)	-0.9794	(0.1638)
MQ_c	-27.5893	$r = 4$	-30.8682	$r = 3$	-3.0300	$r = 1$
$CIPS^{*}$	-2.4220	(0.6750)	-2.4931	(0.2750)	-2.1339	(0.8150)
t_a	0.5300	(0.7019)	-1.7276	(0.0420)	-1.6730	(0.0472)
t_b	0.4895	(0.6878)	-1.7380	(0.0411)	-0.6189	(0.2680)
$n.c.f.$	4		3		4	
$H_0 : I(2)\ vs\ H_A : I(1)$						
	lp		lu		lw	
$PMSB$	-2.3647	(0.0092)	-2.4552	(0.0063)	-3.2125	(0.0007)
MQ_c	-77.7600	$r = 0$	-82.4687	$r = 0$	-134.6321	$r = 0$
$CIPS^{*}$	-4.0609	(0.0100)	-6.4003	(0.0100)	-5.1347	(0.0100)
t_a	-371.0119	(0.0000)	-326.4307	(0.0000)	-344.9117	(0.0000)
t_b	-35.6467	(0.0000)	-53.0366	(0.0000)	-61.6577	(0.0000)
$n.c.f.$	4		3		4	

Moreover, the three variables appear to be cointegrated according to Westerlund and Pedroni tests shown inTable 7. The panel statistics of Pedroni (1999) are based on pooling different estimates, from the estimated residuals, across members while the group statistics simply average these estimates. ρ and t statistics can be seen as variations of the analogous ρ and t tests of Phillips and Perron; the first is a non-parametric variance ratio statistic. The four tests are asymptotically distributed as standard normal. The variance ratio test is right-sided, while the other tests are left-sided. Similarly, the P statistics of Westerlund (2007) pool information over all the cross-sectional units whereas the G statistics are obtained as weighted averages of individual estimates. They are asymptotically normally distributed; Z-value denotes the standardized value of the statistic. The pvalue is robust to cross-sectional dependence and has been obtained after 400 bootstraps. All the tests include individual effects and time trend. Maximum truncation lags are set to 4 and determined using data dependent criteria.

Table 7. Panel cointegration tests		
PEDRONI tests		
	Panel statistics	Group statistics
Variance ratio	1.938^*	
ρ statistic	-8.884^*	-8.406^*
t statistic	-8.315^*	8.137^*
ADF statistic	-0.305	0.015
WESTERLUND tests		
	Z-value	pvalue
G_t test	-3.241^*	0.008
G_a test	0.869	0.605
P_t test	-3.711^*	0.006
P_a test	-1.815	0.097
*: denotes rejection of the null of no cointegration at a 5% significance level.		

Eight of the eleven cointegration tests reject the null of no cointegration whereas the two ADF tests in the case of Pedroni and the two tests based on the long run variance estimators of Westerlund detect problems. Overall, it seems the evidence supporting the assumption of cointegration is stronger.

This period has been crucial for the evolution of the contemporary Spanish economy and includes many important events such as inclusion in the European Common Market in 1986, the crisis of 1993, the full decentralization of the Public Administration or the economic crash in the second semester of 2007 and subsequent downturn. We look for structural breaks using the Banerjee and Carrión-i-Silvestre test (2015), which reveals a significant breakpoint in the second quarter of 2000 (weaker symptoms of a second break were detected also at the beginning of the 1990s and/or in 2008, depending on the region)[3]. Let us note that the procedure detects the presence of, at least, 4 non-stationary common factors, which would mean

[3] We do not report the results of their t_e^τ cointegration test because there is strong evidence of cross-sectional dependence in the idiosyncratic errors, distorting its asymptotic distribution.

that the observed variables do not cointegrate by themselves alone (common factors are needed to obtain a significant long-run relationship). We will not go deeper into this question, which may be connected with the problems of cointegration raised in Table 7. The model we are using, for reasons of lack of information, is an oversimplified version of the functioning of a labour market and, probably, it is somewhat misspecified. Our impression is that those nonstationary common factors are related to elements of the national or international cycle not included in the model.

Literature on labour markets agrees on the negative impact of economic activity on unemployment. The relation is more controversial in relation to wages: a positive impact is predicted from a neoclassical perspective although there are doubts from a social-democratic view. Using previous results, we approach this discussion through an Autoregressive Distributed Lag Model, $ARDL(p, q_1, q_2)$; see Pesaran *et al*, 2001. The lengths p, q_1 and q_2 of the $ARDL$ have been fixed paying attention to the assumption of serially uncorrelated disturbances and to the Akaike Information Criterion. Results indicate that the most adequate lag lengths, to eliminate residual serial correlation, are $p = 4$, $q_1 = 2$ and $q_2 = 2$. The ARDL has been parameterized into an error correction equation such as the following:

$$\left.\begin{array}{c} \Delta lu_{it} = \gamma_i \left(lu_{it-1} - \beta_{1i} lp_{it} - \beta_{2i} lw_{it} - \beta_{3i} d_{it} \times lw_{it} \right) + \alpha_i + \pi_i d_{it} \\ \sum\limits_{j=1}^{4} \theta_{1j,i} \Delta lu_{it-j} + \sum\limits_{j=1}^{2} \theta_{2j,i} \Delta lp_{it-j} + \sum\limits_{j=1}^{2} \theta_{3j,i} \Delta lw_{it-j} + u_{it}; \\ i = 1, ..., 17; t = 1980:1, ..., 2013:4; d_{it} = 1 \ if \ t \geqslant 2000:2 \end{array}\right\} \quad (18)$$

Another problem to consider is the homogeneity of the parameters of the cointegration equation among the different CCAA. A way of solving this question is by comparing the MG and PMG estimates. The first (Pesaran and Smith, 1995) relies on estimating n time-series regressions, one for each CCAA, and averaging the coefficients, whereas the PMG estimator (Pesaran *et al.*, 1999) is a mixture of pooling and averaging of coefficients. The comparison can be made through a simple Hausman test where the MG estimates are consistent under the null (homogeneous parameters) and the alternative (heterogeneous parameters) hypothesis, whereas the PMG estimates are efficient under the null but biased under the alternative. In this case, the Hausman statistic takes a value of 2.44 (3 *d.o.f.*) and a p-value of 0.4861 which points in favor of the mixed PMG estimates. Table 8 reports the results corresponding to the common panel cointegration equation (top panel) and the average of the short-run parameters (including a measure of dispersion of these estimates). Misspecification tests for temporal and cross-sectional dependence complete the information. The details of the short-run estimates of the $ARDL(4, 2, 2)$ corresponding to each Autonomous Community appear in Table B1 in Appendix B.

A striking result in this Table is the increased impact of regional wages after the structural break in 2000:2. The relation with unemployment was inelastic before this date (5% confidence interval is $0.53 - 1.04$) but becomes highly elastic afterwards (the confidence interval is $2.78 - 3.66$). This result highlights the increasing importance of purely regional factors in the determination of unemployment, as a consequence of the decentralization process. Economic activity, as expected, appears as an important factor reducing regional imbalances, with a more than proportional, stable impact.

The speed of adjustment (γ_i) is significant and negative in the 17 CCAA with a minimum value in the case of Navarre, -0.021, and a maximum in La Rioja, -0.233. The cross-equation restriction of homogeneity in this parameter is strongly rejected with a likelihood ratio test of 50.40 (pvalue of 0.0000). The short-run impact of regional economic activity on unemployment (sum of parameters θ_{21} and θ_{22}) is negative and more than proportional; in fact, the null hypothesis of a zero short-run impact, against the alternative of a negative impact, is rejected in 14 out of the 17 regions. The situation is not so clear in the case of wages where this hypothesis cannot be rejected in almost half of the CCAA.

The CD test indicates the existence of strong spatial interaction in the residual of the $ARDL(4, 2, 2)$ for the Spanish unemployment series. A natural reaction would be to fix the misspecification by including a SEM structure. Obvious questions are:

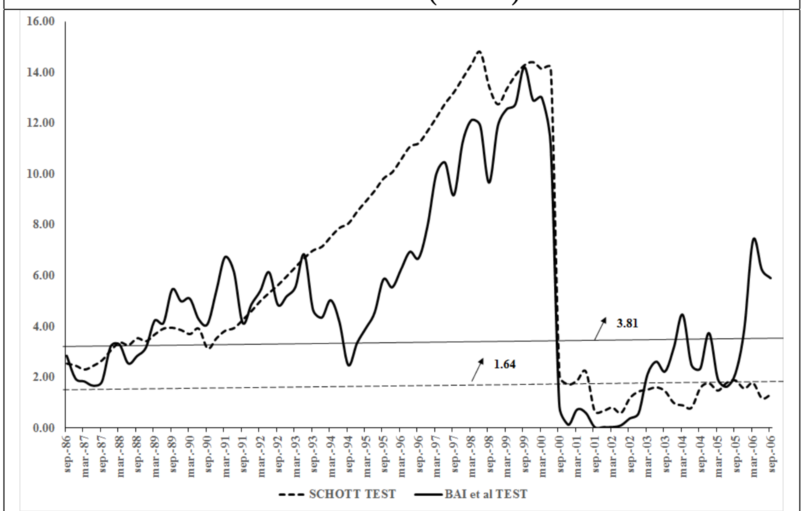
- what weighting matrix?
- is it in fact the same for the three decades?

Table 8. Estimation of the ARDL(4,2,2) for the unemployment case (18)		
Long run coefficients		
	Estimate	p-value
β_1 (gdp)	1.6324	0.0000
β_2 (wages)	0.7896	0.0000
β_3 (wages×time dummy)	2.4287	0.0000
Short run coefficients		
	Mean estimate	stand. dev.
γ_i	−0.0891	0.0334
α_i	1.5823	0.0226
π_i	−0.5105	0.0386
θ_{11}	0.1656	0.1901
θ_{12}	−0.2963	0.1842
θ_{13}	0.3348	0.1419
θ_{14}	−0.0669	0.3400
θ_{21}	−1.9912	0.0774
θ_{22}	0.1592	0.4133
θ_{31}	−0.1144	0.3057
θ_{32}	−0.0331	0.3640
Serial correlation	Estimate	p-value
r_1	−0.1114	0.1738
r_2	0.0206	0.3878
r_3	0.0166	0.3916
r_4	−0.0400	0.3584
r_5	0.0011	0.3989
Cross-sect correl.	Estimate	p-value
CD test	37.2447	0.0000

Note: r_j denotes the mean correlation coefficient between the cross-sections residuals separated by j periods. Given that T is larger than N, it is compared with a N(0,1). CD: denotes Pesaran's test of cross-sectional dependence in panel data. See Arellano and Bond (1991) and Pesaran (2004), respectively.

Different hypotheses can be formulated in relation to the first question, none of them definitive. The nonparametric procedure of Bhattacharjee and Jensen-Butler (2013) produces the weighting matrix listed in Table B2 in Appendix B. The answer to the second question is negative: this matrix cannot be taken as constant for the three decades. Figure 5 shows the rolling estimates of the T_N test of Bai *et al* (2009) and of the t_{Tn} test of Schott (2007). The search span contains 60% of the observations (that is, 80 quarters) and it is centered in the middle of the sample (1998:1). Both tests peak in the second semester of 2000, where the Banerjee and Carrión-i-Silvestre procedure identifies a similar break in the cointegration equation. Notice that this date roughly coincides with the completion of the transference of the second large set of competencies to the regions. We do not have a clear cut explanation for the sharp fall produced in both tests after the peak.

Figure 5. Results of the Schott (2007) and Bai *et al.* (2009) tests



Under the alternative hypothesis there are, at least, two weighting matrices working in the sample, as they appear in Tables B3 and B4 in Appendix B. Once more, the matrices have been estimated using the nonparametric algorithm of Bhattacharjee and Jensen-Butler (2013).

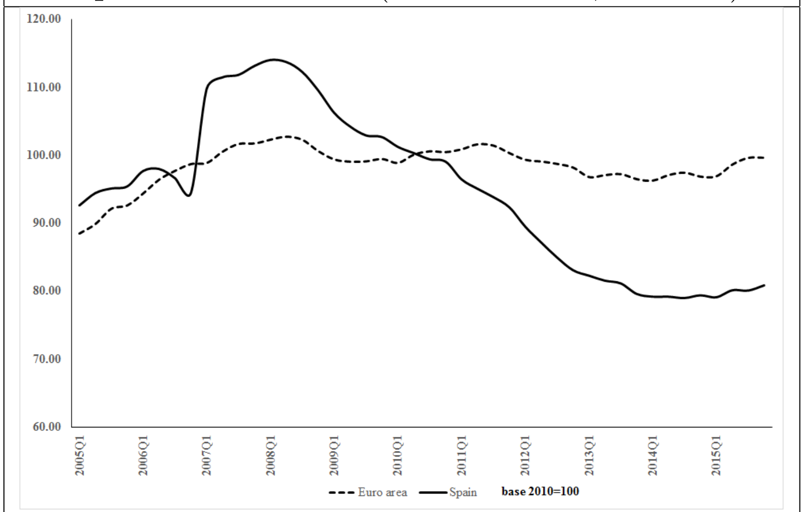
Some comments are in order here. First, the **W** matrix obtained for the first period is sparser than that for the second. This is consistent with the decentralization hypothesis: as the regions gain control over their own decisions, the interrelation between them increases. Following this change, there has been a realignment among the regions in terms of 'good' and 'bad' neighbours. During the first period, the ranking of regions with a positive impact on the others is headed by geographically central regions, such as Castille-Leon, Castille-La Mancha or Madrid; in the second period, the group of regions with positive influence has moved towards the Mediterranean axis (both Castilles, Catalonia, Valencian Community are there). On the contrary, peripheral regions appear at the top of the ranking of regions with negative impulses in the first period (Balearic Islands, Cantabria, Murcia), which is made by Northern and peripheral regions in the second (Asturias, Cantabria, Canary Islands, Extremadura). This picture is also consistent with the gradual but firm displacement of the centre of gravity of the Spanish economy from the North-West to the South-East of the peninsula. Another interesting result is the 'lack of geography' in these matrices whose relation with the concept of contiguity is very feeble, if not entirely absent. This also accords with the distinction between micro and macro data in the sense that aggregating spatial data leads to the weakening of purely geographical relations, favouring other types of comovements such as national common factors.

5.2 Case Study II: Housing prices in Spanish municipalities

The second example refers to housing prices in Spanish municipalities. National housing prices rose strongly in the expansion cycle that preceded the global downturn of 2007; this increment was one of the largest in Europe. However the crisis led to severe corrections in the sector, stronger than the European mean. The boom period started, approximately, in 1995 and the housing stock increased by 50 percent from 1995 to 2007, with an average of more than half a million houses built per year. Meanwhile the population increased by 6.5 million people (18%) and the GDP per capita by more than 4% on average. The market collapsed after the financial crash of 2007 with cuts in prices by more than 30% on average (see Figure 6), GDP per capita decreased around 2.3% per year in the period 2008 to 2014 and there was an out-migration of almost half a million people.

Some factors are specific to the Spanish case, such as the massive affluence of baby-boomers in the early 2000s increasing the group of young working adults, prone to demand housing services. The loosening of financial conditions in the years before the crisis facilitated the access to credit, which spurred the demand for housing in a market with a strong preference for ownership. Moreover, fiscal policy favoured ownership, with subsidies, deductions in personal income taxes and several incentives to reinvestment that were partially removed from the Spanish tax system in 2012.

Figure 6. House price indices, 2005:1-2015:4. Spain vs euro area (base 2010=100, Eurostat)



There is a rich literature devoted to real estate markets. A common starting point is the notion of real housing user cost of capital (Cameron *et al.*, 2006, or Holly *et al.*, 2010), which confronts the consumer with a utility maximization problem in an inter-temporal model of consumption, with two different goods: a composite consumption good and housing services. Real housing user cost of capital results from the first order conditions of the optimization problem in which interest rate, housing stock, income and demographic factors are key variables driving the formation of prices. In addition, real estate markets tend to be segmented into several sub-markets (owner-occupied houses, second residences, tourist activities, investors' property developments, etc.), each one of them with a strong spatial dynamic (e.g., Can, 1990 and 1992; Brady, 2011; Martinez and Maza, 2003, for the Spanish case).

Our case study refers to housing prices for the 284 Spanish municipalities with more than 25,000 residents (Figure 7 shows their location). The data emanate from the Spanish Department of Housing (Dirección General de Arquitectura, Vivienda y Suelo, Ministerio de Fomento), and comprise quarterly data from 2005:1 to 2015:4. However, we are not going into the direction of modelling because there is an acute problem of data unavailability at a municipal level in Spain. We merely test for structural breaks in the cross-sectional connections among the municipalities.

The cross-sectional dimension of this dataset, 284 municipalities, is 6 times greater than its time dimension, 44 quarters in the sample. There is a strong dimensionality problem that prevents the use of LR based tests. Moreover, given the short time dimension, it is not possible to estimate the underlying weighting matrix. As previously remarked, at most we can speculate about its stability.

Figure 7. Location of the municipalities included in the sample

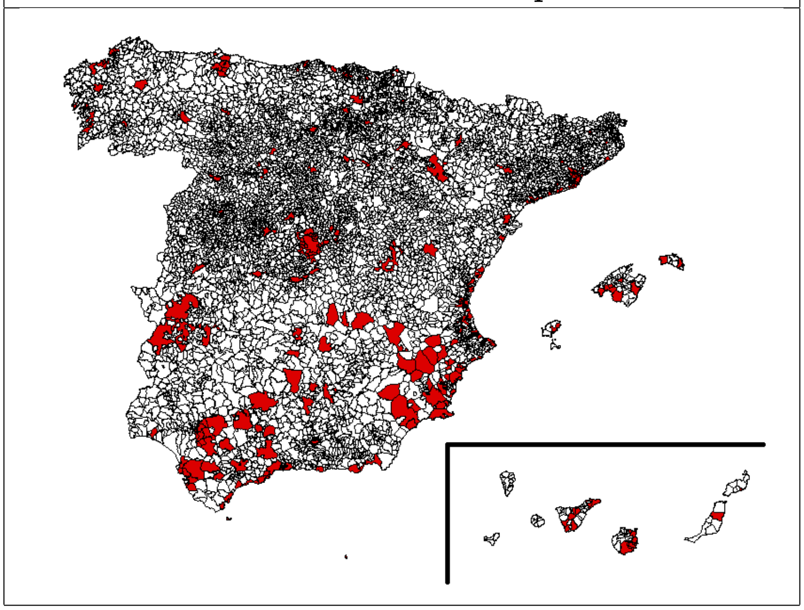
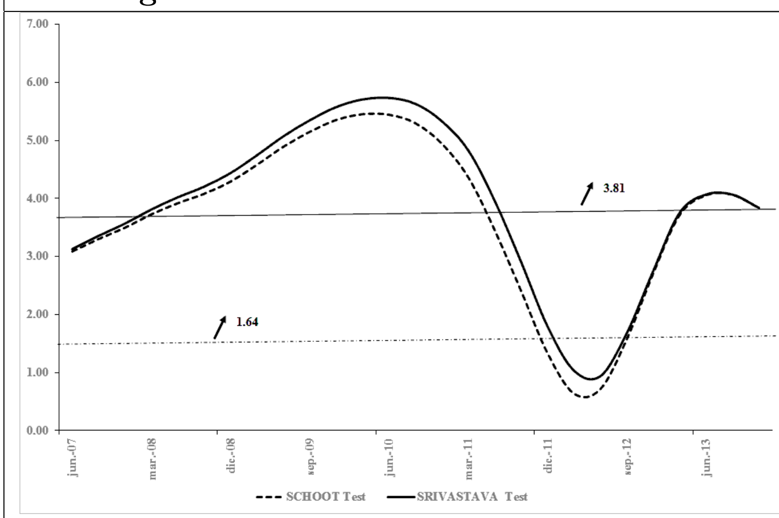


Figure 8 shows the rolling estimates of the t_{Tn} test of Schott and of the T_2 test of Srivastava. As before, the search span contains 60% of the observations (that is, 26 quarters) and it is centered in the middle of the sample (2010:2). The two tests offer a very similar view of this case: there is a wide time interval where they detect a break in the covariance matrix: from the beginning of 2008 to the second semester of 2011, with a peak in the first quarter of 2010. This period coincides with the outbreak of the crash and the bursting of the Spanish housing bubble.

Of course, this break may be due to other causes different from the hypothesized change in the weighting matrix (for example, changes of agents' expectations, in the perception of risk, etc.), but the instability in the cross-sectional structure of the markets is a potential disturbing factor that cannot be neglected. This is consistent with the realignment of the Spanish real estate sector after the crash of 2007. Actually, the sector is working at a very moderate level of activity and it is highly dependent on the demand for residential housing, in very specific and well chosen locations. Real estate markets evolve at a lower scale than before and the activity is a bit more diffuse on the territory; rehabilitation and restoration of properties occupy a more prominent role to the detriment of purely touristic sub-markets. In sum, the sector appears to have acquired a more balanced spatial distribution which may have caused a break in the connections captured by the weights matrix connecting local markets.

Figure 8. Schott and Srivastava tests.



6 Conclusion

In recent years there has been a growing interest in different questions related to $\mathbf{W}$. There appears to be a general consensus in the sense that it is not enough to build the weights matrix exogenously according to some unspecified prior knowledge. On the contrary, the author should make explicit his/her knowledge in order that the results of the estimated models may be interpreted. Usually, one has several candidate matrices, which amounts to a decision problem that should be solved using some of the different algorithms that exist in the literature. In other cases, assuming that we have panel data, the **W** matrix can be estimated from the data.

Our contribution focuses on the (implicit) assumption of time constancy of **W** in a panel data framework. If the time span is large or there have been external shocks, it seems reasonable to expect changes also in the cross-sectional connections. The approach in this paper consists in testing the existence of possible breaks in **W** through the corresponding covariance matrix. The literature devoted to the topic of comparing covariance matrices is large and heterogeneous, from which we have chosen a set of potentially useful tests.

The Monte Carlo experiment reported in the paper leads us to a clear conclusion: the T_N test of Bai *et al* (2009), the t_{T_n} test of Schott (2007) and the T_2 test of Srivastava (2007) are candidates that adapt well to our case. The first test tends to be slightly oversized, especially for small time sample sizes. The T_2 test is consistently undersized whereas the t_{T_n} test of Schott is more balanced in relation to size. The estimated power function of the three tests improves very quickly, both with the time span and the spatial coefficient. The tests work quite well in the case where the location of the breakpoint is unknown and also when the change in the weight matrix is gradual. On the negative side, because the T_N test is a corrected version of a standard Likelihood Ratio, it needs a concentration c_i index in each subsample lower than 1. This limitation does not apply for the other two tests, built around the notion of the Frobenius norm. Their behavior is worse for processes of local dependence, such as the SEM, where the symptoms of spatial dependence are weaker. Moreover, the distance between the weight matrices involved in the change is a crucial factor to guarantee a proper functioning of the tests.

As said in the Introduction, this paper is part of ongoing research whose objective is to endogenize fully the building of the weighting matrix in a spatial model. Tests for its stability are a small step forward, and need to be combined with more flexible estimation algorithms. The potential usefulness of these techniques, combined with other tools recently developed in relation to the weights matrix has been illustrated by means of two case studies coming from the contemporary Spanish economy. Overall, the results of our approach corroborate the dominant position with respect to this economy.

7 Appendix A: The SEM case in the errors of a linear model

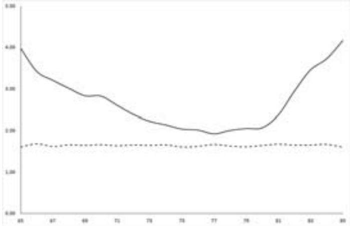
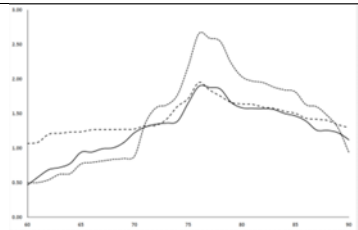
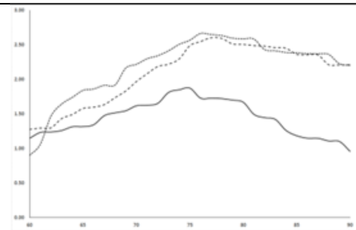
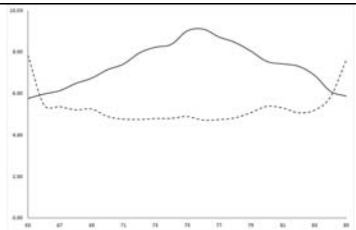
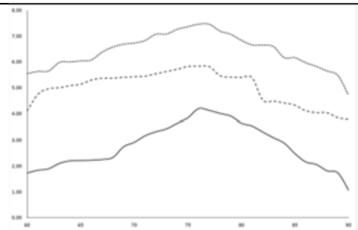
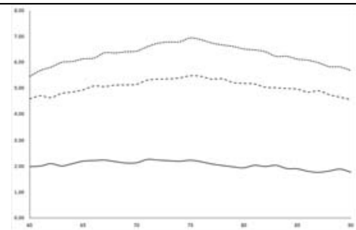
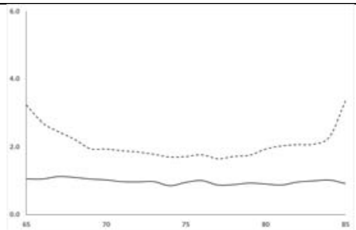
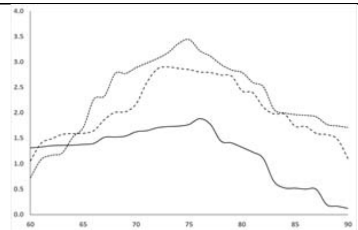
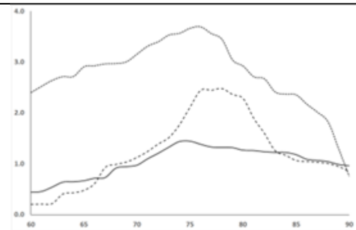
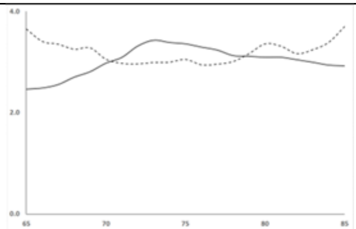
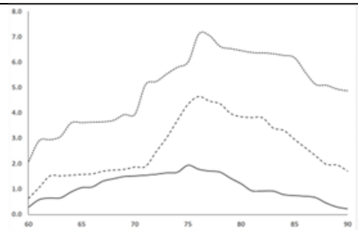
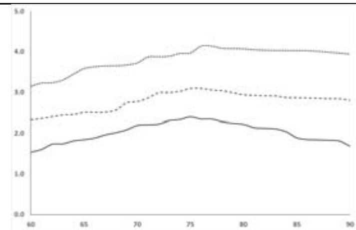
This Appendix describes the results obtained for an extension of the Monte Carlo discussed in Section 4. Here we treat the case where the user knows that the spatial structure appears only in the error terms of the equation. The DGP simulated corresponds to expressions (16), null hypothesis, and (17), alternative hypothesis. This (partial) knowledge of the DGP allows the user to (i) estimate the equation of the mean using a consistent estimator (LS in our Monte Carlo), (ii) obtain the corresponding residuals and (iii) solve the test for breaks using the covariance matrix of the residuals from step (ii). We expect that this additional knowledge in the experiment improves the behaviour of the tests. Main results, for the T_N test of Bai *et al* (2009), the t_{T_n} test of Schott (2007) and the T_2 test of Srivastava (2007) appear in Tables A1 and A2 below, for the case of a known breakpoint, and in Figure A1, for the case of an unknown instantaneous breakpoint (results for the case of a gradual change in the weight matrix are much in line with those shown below; details are available from the authors upon request).

This partial knowledge is highly beneficial for the three tests, both in power and in size. The T_N test appears to be, still, slightly oversized especially for high values of c_i, the concentration index. The T_2 of Srivastava continues to be undersized although the intensity of the spatial dependency tends to correct this deficit. The size of the t_{T_n} test of Schott is right for all the cases in the experiment. The impact on the estimated power of the three tests is clear. Overall, data on power have increased around 30%-40% in relation to the situation described in Tables 4a-4d. The power of the tests improves with the sample size (in the cross-section and/or in the time dimension), with the strength of the spatial dependence and with the distance between the matrices in the break. Figure A1 confirms this amelioration for the case of an unkown instantaneous break. The U-shaped pattern is still present in the estimated power function of the T_N test for low values of the spatial dependence coefficient. The behaviour of the other two tests is more efficient now in the sense that the peak of the sequence of values is better defined in the proximity of the (unknown) breakpoint.

Table A1: Estimated size (significance level=0.05). Known breakpoint. SEM in the errors of a linear model. High signal-to-noise ratio. Matrix in the DGP: WR.												
ρ=0.3												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.09	-	-	0.06	0.11	-	0.05	0.12	-
t_{Tn}	0.04	0.04	0.05	0.04	0.04	0.06	0.05	0.05	0.04	0.06	0.06	0.07
T_2	0.02	0.02	0.03	0.01	0.01	0.01	0.03	0.01	0.01	0.02	0.01	0.02
ρ=0.7												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.14	-	-	0.16	0.15	-	0.10	0.06	-
t_{Tn}	0.04	0.04	0.05	0.05	0.06	0.05	0.05	0.03	0.05-	0.06	0.06	0.05
T_2	0.04	0.03	0.03	0.04	0.04	0.04	0.06	0.05	0.05	0.07	0.05	0.05
*: For the T_N tests the percentages shown in this panel correspond to n=64.												

Table A2: Estimated power (significance level=0.05). Known breakpoint. SEM in the error of a linear model. High signal-to-noise ratio.												
ρ=0.3 Matrices in the DGP: WR vs WD												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.46	-	-	0.51	0.72	-	0.64	0.88	-
t_{Tn}	0.32	0.39	0.43	0.41	0.64	0.72	0.55	0.80	0.89	0.73	0.82	0.93
T_2	0.28	0.29	0.44	0.45	0.49	0.76	0.58	0.67	0.80	0.63	0.75	0.86
ρ=0.7 Matrices in the DGP: WR vs WD												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.78	-	-	0.83	0.82	-	0.91	0.89	-
t_{Tn}	0.38	0.45	0.64	0.47	0.74	0.85	0.69	0.87	0.95	0.87	1.00	1.00
T_2	0.40	0.46	0.53	0.52	0.68	0.78	0.66	0.78	0.89	0.79	1.00	1.00
ρ=0.3 Matrices in the DGP: WR vs WQ												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.28	-	-	0.41	0.52	-	0.54	0.58	-
t_{Tn}	0.19	0.26	0.38	0.32	0.47	0.61	0.41	0.60	0.66	0.63	0.82	0.93
T_2	0.17	0.22	0.42	0.35	0.39	0.63	0.45	0.57	0.70	0.58	0.65	0.79
ρ=0.7 Matrices in the DGP: WR vs WQ												
T	**30**			**60**			**150**			**250**		
n	**16**	**144**	**225**	**16**	**144**	**225**	**16**	**144***	**225**	**16**	**144***	**225**
T_N	-	-	-	0.42	-	-	0.53	0.59	-	0.82	0.71	-
t_{Tn}	0.31	0.35	0.51	0.37	0.61	0.75	0.51	0.66	0.78	0.73	0.81	1.00
T_2	0.37	0.42	0.55	0.42	0.51	0.68	0.56	0.69	0.80	0.75	0.98	1.00
*: For the T_N tests the percentages shown in this panel correspond to n=64.												

Figure A1: Unknown, instantaneous breakpoint. High signal-to-noise ratio. SEM in the errors of a linear model.

	Bai test	**Schott test**	**Srivastava test**
ρ=0.3 WR vs WQ. T=150			
ρ=0.7 WR vs WQ. T=150			
ρ=0.3 WR vs WD T=150			
ρ=0.7 WR vs WD. T=150			

APPENDIX B: Additional results for the case of Spanish Regional Unemployment.

Table B1: Detailed results for the ARDL(4,2,2) short-run estimates

	ANDA		ARAG		ASTU		BALE		CANA		CANT		CLEO		CMAN		CATA	
	Estima.	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**
γ_i	-0.054	0.001	-0.095	0.002	-0.197	0.000	-0.142	0.001	-0.087	0.001	-0.096	0.002	-0.091	0.000	-0.114	0.000	-0.056	0.005
α_i	1.184	0.002	1.597	0.004	3.238	0.000	2.299	0.002	1.573	0.002	1.413	0.006	1.707	0.000	1.987	0.000	1.228	0.009
π_i	-0.305	0.001	-0.564	0.003	-1.256	0.000	-0.742	0.001	-0.443	0.001	-0.567	0.002	-0.550	0.000	-0.599	0.000	-0.348	0.006
θ_{11}	0.316	0.023	0.316	0.060	0.097	0.323	-0.008	0.399	0.143	0.249	0.026	0.394	0.336	0.020	0.420	0.002	0.426	0.001
θ_{12}	-0.404	0.116	-0.729	0.019	-0.175	0.331	-0.425	0.171	-0.246	0.271	-0.352	0.206	-0.749	0.006	-0.764	0.004	-0.478	0.054
θ_{13}	0.477	0.051	0.756	0.003	-0.020	0.398	0.679	0.011	0.219	0.264	0.469	0.071	0.798	0.001	0.597	0.009	0.345	0.118
θ_{14}	-0.173	0.040	-0.225	0.006	0.051	0.321	-0.225	0.005	-0.042	0.347	-0.164	0.048	-0.247	0.002	-0.144	0.052	-0.114	0.128
θ_{21}	-1.418	0.018	-2.906	0.007	-1.334	0.163	-4.870	0.000	-2.165	0.008	-2.066	0.048	-1.061	0.139	-1.372	0.013	-2.864	0.001
θ_{22}	-0.448	0.247	0.718	0.249	1.020	0.099	1.846	0.075	-0.073	0.396	0.445	0.333	-0.144	0.379	-0.462	0.171	0.242	0.367
θ_{31}	0.136	0.109	-0.245	0.069	-0.748	0.000	0.121	0.365	-0.104	0.292	-0.151	0.165	-0.090	0.197	-0.081	0.350	-0.026	0.380
θ_{31}	-0.075	0.115	0.026	0.374	0.174	0.113	-0.207	0.154	-0.017	0.389	-0.012	0.393	-0.055	0.194	0.017	0.390	-0.127	0.017

	CVAL		EXTR		GALI		MADR		MURC		NAVA		PAVA		RIOJ	
	Estima.	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**	**Estima.**	**pvalue**
γ_i	-0.105	0.000	-0.160	0.000	-0.084	0.000	-0.089	0.001	-0.083	0.001	-0.021	0.337	-0.040	0.012	-0.233	0.000
α_i	2.132	0.001	2.607	0.000	1.588	0.001	1.873	0.002	1.391	0.002	0.318	0.333	0.765	0.016	3.015	0.000
π_i	-0.572	0.000	-0.891	0.000	-0.479	0.000	-0.543	0.002	-0.433	0.001	-0.100	0.364	-0.286	0.010	-1.336	0.000
θ_{11}	0.193	0.137	-0.088	0.348	0.217	0.133	0.075	0.359	0.104	0.306	-0.275	0.180	0.167	0.154	-0.096	0.350
θ_{12}	-0.454	0.094	0.121	0.370	-0.535	0.071	-0.260	0.271	-0.077	0.383	-0.067	0.392	0.168	0.303	0.068	0.390
θ_{13}	0.328	0.152	-0.047	0.392	0.554	0.029	0.220	0.273	0.109	0.356	0.213	0.303	-0.393	0.068	-0.066	0.387
θ_{14}	-0.054	0.312	0.012	0.394	-0.160	0.053	-0.064	0.290	-0.055	0.299	-0.114	0.167	0.184	0.014	0.008	0.397
θ_{21}	-3.167	0.000	1.181	0.141	-0.839	0.141	-3.117	0.009	-2.581	0.000	-2.303	0.010	-2.967	0.000	-5.633	0.000
θ_{22}	0.744	0.148	-0.531	0.254	0.563	0.133	0.709	0.278	-0.885	0.165	-0.183	0.387	-0.861	0.185	3.309	0.004
θ_{31}	-0.064	0.303	-0.378	0.005	-0.261	0.003	-0.017	0.398	0.211	0.042	-0.269	0.150	0.022	0.365	-0.798	0.000
θ_{32}	-0.051	0.242	0.061	0.287	0.011	0.391	-0.129	0.237	-0.162	0.009	-0.035	0.378	-0.074	0.033	0.193	0.091

NOTE: ANDA: Andalucia; ARA: Aragon; ASTU: Asturias; BALE: Balearic Islands; CANA: Canary Islands; CANT: Cantabria; CLEO: Castille-Leon; CMAN: Castille-La Mancha; CATA: Catalonia; CVAL: Valencian Community; EXTR: Extremadura; GALI: Galicia; MADR: Community of Madrid; NAVA: Navarre; PAVA: Basque Country; RIOJ: La Rioja

Table B2 . Estimated weighting matrix for the Spanish UNEMPLOYMENT case, under the NULL hypothesis (no structural break)

	AND	ARA	AST	BAL	CAN	CAB	CLE	CMA	CAT	CVA	EXT	GAL	MAD	MUR	NAV	PAV	RIO	|SUM|	"+"	"-"
AND	**0.00**	-	-0.07	0.10	-	0.09	0.09	0.08	0.07	0.10	0.08	-	-	0.07	-0.06	-	-	0.75	0.57	-0.18
ARA		**0.00**	0.12	-0.06	0.07	0.07	0.09	0.15	-	0.04	-0.05	-	0.04	-	-	-	0.06	0.85	0.61	-0.24
AST			**0.00**	-	0.02	0.10	0.13	0.07	-0.09	0.11	0.15	0.09	-	-0.07	-	0.05	0.06	1.18	0.70	-0.47
BAL				**0.00**	-0.08	-0.10	-	-	0.06	-	0.07	0.10	0.08	-	0.08	-	-	0.83	0.25	-0.58
CAN					**0.00**	-	0.09	-	0.08	0.08	-	-0.11	0.09	0.09	0.07	-	0.04	0.89	0.49	-0.40
CAB						**0.00**	0.07	0.11	0.08	-0.05	-	0.09	0.08	-0.07	-	0.08	-	1.06	0.63	-0.43
CLE							**0.00**	0.09	0.12	0.09	0.08	0.04	0.06	0.10	0.07	-	-	1.16	1.12	-0.03
CMA								**0.00**	-	0.14	0.07	-	-	-	0.08	-	-	0.89	0.79	-0.11
CAT									**0.00**	0.04	0.04	0.19	-0.05	0.04	-0.04	0.11	-	0.86	0.65	-0.21
CVA										**0.00**	-0.06	0.05	0.06	0.06	-	0.06	-	0.97	0.69	-0.28
EXT											**0.00**	0.05	0.13	0.04	-	0.04	0.05	0.95	0.68	-0.28
GAL												**0.00**	-	0.08	0.09	-	0.06	1.08	0.75	-0.33
MAD													**0.00**	0.06	-	-	0.09	0.84	0.64	-0.20
MUR														**0.00**	-	-	-	0.63	0.45	-0.18
NAV															**0.00**	0.06	0.09	0.74	0.41	-0.33
PAV																**0.00**	-	0.62	0.53	-0.08
RIO																	**0.00**	0.64	0.53	-0.11

Number of Contacts different from zero (1% significance level)	178
Percentage of non zero cells	65.44%

NOTE: The weighting matrix has been estimated using the procedure described in Bhattacharjee and Jensen-Butler (2013), that assumes symmetry. The ARDL(4,2,2) of Table 8 and 9 have has been bootstrapped under the null of stability of the weighting matrix in order to estimate the standard deviation of the estimated weights of the weighting matrix.
Weights not significant at a 1% significance level are denoted with a dash,-, in the table above.
|SUM|: denotes the sum of the absolute weights in corresponding row; "+" ("-"): denotes de sum of the positive (negative) weights in the corresponding row.

Table B3. Estimated weighting matrix for the Spanish UNEMPLOYMENT case, under the ALTERNATIVE hypothesis (structural break in 2000:3) PERIOD 1980:1-2000:2

	AND	**ARA**	**AST**	**BAL**	**CAN**	**CAB**	**CLE**	**CMA**	**CAT**	**CVA**	**EXT**	**GAL**	**MAD**	**MUR**	**NAV**	**PAV**	**RIO**	**|SUM|**	**“+”**	**“-“**
AND	**0.00**	-	-	0.15	-	0.11	0.15	0.05	-	-	0.06	-0.08	-	0.10	-0.12	-	-	1.08	0.68	0.40
ARA		**0.00**	0.10	-0.11	-	0.09	0.11	0.18	-	-	-0.09	-	-	-	0.11	-	0.07	1.01	0.59	0.42
AST			**0.00**	-	-	-	0.14	0.07	-	0.07	0.15	-	0.17	-	-	0.12	-	1.05	0.86	0.18
BAL				**0.00**	-0.11	-	-	-	-	-	-	0.13	-	-	0.11	-	0.04	1.04	0.27	0.77
CAN					**0.00**	-	0.11	-	-	0.09	-	-	0.09	0.08	0.06	0.08	-	0.88	0.57	0.31
CAB						**0.00**	-	-	-	-	0.08	0.10	0.10	-0.13	-0.09	0.09	0.08	1.14	0.58	0.57
CLE							**0.00**	-	-	0.06	-	-	-	0.08	0.08	-	0.08	1.09	1.05	0.04
CMA								**0.00**	-	0.19	0.13	-	-	-	-	-	-	0.89	0.70	0.19
CAT									**0.00**	0.07	-	0.09	-	-	-	0.11	-	0.65	0.51	0.14
CVA										**0.00**	-	0.08	-	0.12	-	-	-	0.96	0.77	0.19
EXT											**0.00**	-	0.15	-	-	-	0.12	1.00	0.64	0.36
GAL												**0.00**	-	0.08	0.10	-	0.07	0.93	0.61	0.32
MAD													**0.00**	0.06	-	-	-	0.93	0.79	0.13
MUR														**0.00**	-	-	-	0.77	0.38	0.39
NAV															**0.00**	0.09	-	0.86	0.36	0.50
PAV																**0.00**	-	0.82	0.61	0.21
RIO																	**0.00**	0.70	0.52	0.18

Number of Contacts different from zero (1% significance level) 114

Percentage of non zero cells 41.91%

NOTE: The weighting matrix has been estimated using the procedure described in Bhattacharjee and Jensen-Butler (2013), that assumes symmetry. The ARDL(4,2,2) of Table 8 and 9 have has been bootstrapped under the null of stability of the weighting matrix in order to estimate the standard deviation of the estimated weights of the weighting matrix.
Weights not significant at a 1% significance level are denoted with a dash,-, in the table above.
|SUM|: denotes the sum of the absolute weights in corresponding row; “+” (“-“): denotes de sum of the positive (negative) weights in the corresponding row.

Table B4. Estimated weighting matrix for the Spanish UNEMPLOYMENT case, under the ALTERNATIVE hypothesis (structural break in 2000:3) PERIOD 2000:3-2013:4

	AND	**ARA**	**AST**	**BAL**	**CAN**	**CAB**	**CLE**	**CMA**	**CAT**	**CVA**	**EXT**	**GAL**	**MAD**	**MUR**	**NAV**	**PAV**	**RIO**	**|SUM|**	**"+"**	**"-"**
AND	**0.00**	-	-0.16	-	-0.04	0.06	-	0.10	-	0.22	0.08	0.09	-0.11	0.07	-	-	-	0.80	0.44	0.37
ARA		**0.00**	0.15	-	0.12	0.11	-	0.17	-	0.06	-	-	-	-	-0.09	-	0.07	0.92	0.72	0.19
AST			**0.00**	-	0.04	0.16	0.15	0.06	-0.20	0.17	0.20	0.19	-0.13	-0.16	-	-	0.15	2.01	0.57	1.44
BAL				**0.00**	-0.04	-0.10	-	-	-0.01	-	0.16	0.08	-	-	-	-	0.07	0.70	0.28	0.43
CAN					**0.00**	-	0.05	-	0.17	0.05	-0.05	-0.17	0.09	0.10	0.15	-	-	1.18	0.40	0.78
CAB						**0.00**	0.09	0.09	0.15	-0.10	-0.11	0.05	0.09	-	0.12	0.09	-0.05	1.41	0.64	0.77
CLE							**0.00**	0.15	0.19	0.15	0.10	-	0.09	0.12	0.08	-	-0.11	1.35	1.14	0.22
CMA								**0.00**	-	0.06	-	-	-	-	0.16	-0.05	-	1.07	0.80	0.27
CAT									**0.00**	0.08	0.15	0.28	-0.13	0.07	-0.13	0.09	-	1.37	0.82	0.55
CVA										**0.00**	-0.12	-	0.13	-	-0.07	0.04	-	1.35	0.73	0.62
EXT											**0.00**	0.08	0.13	0.09	-	-	-0.05	1.43	0.73	0.70
GAL												**0.00**	-	0.08	-	0.10	0.06	1.30	0.93	0.37
MAD													**0.00**	0.06	-	-	0.13	1.01	0.51	0.50
MUR														**0.00**	-	-	0.10	0.67	0.50	0.17
NAV															**0.00**	0.05	0.21	1.32	0.65	0.67
PAV																**0.00**	-	0.57	0.41	0.16
RIO																	**0.00**	1.11	0.59	0.52

Number of Contacts different from zero (1% significance level) 88

Percentage of non zero cells 61.76%

NOTE: The weighting matrix has been estimated using the procedure described in Bhattacharjee and Jensen-Butler (2013), that assumes symmetry. The ARDL(4,2,2) of Table 8 and 9 have has been bootstrapped under the null of stability of the weighting matrix in order to estimate the standard deviation of the estimated weights of the weighting matrix.
Weights not significant at a 1% significance level are denoted with a dash,-, in the table above.
|SUM|: denotes the sum of the absolute weights in corresponding row; "+" ("-"): denotes de sum of the positive (negative) weights in the corresponding row.

References

Aldstadt, J. and A. Getis, 2006, Using AMOEBA to create a spatial weight matrix and identify spatial clusters. *Geographical Analysis*, **38** 327–343.

Angulo, A., P. Burridge and J. Mur, 2016, Testing for a structural break in the weight matrix of the Spatial Error or Spatial Lag model. Working Paper Department of Economic Analysis; University of Zaragoza

Anderson, T., 2003, *An Introduction to Multivariate Analysis*, 3rd ed. New York: John Wiley & Sons.

Arellano, M., and S. Bond, 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies,* **58** 277–29.

Ahrens, A. and A. Bhattacharjee, 2015, Two-Step Lasso Estimation of the Spatial Weights Matrix. *Econometrics*, **3** 128-155.

Bai, Z., Jiang, D., Yao, J. and S. Zheng, 2009, Corrections to LRT on large-dimensional covariance matrix by RMT. *The Annals of Statistics*, **37** 3822-3840.

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

Banerjee, A. and J. Carrión-i-Silvestre, 2015, Cointegration in panel data with structural breaks and cross-sectional dependence. *Journal of Applied Econometrics,* **30** 1–23.

Battacharjee, A. and C. Jensen-Butler, 2013, Estimation of the Spatial Weights Matrix under Structural Constraints. *Regional Science and Urban Economics*, **43** 617-634.

Beenstock, M. and D. Felsenstein, 2012, Nonparametric estimation of the spatial connectivity matrix using spatial panel data. *Geographical Analysis*, **44** 386-397.

Benjanuvatra, S.,2015, QML estimation of the spatial weight matrix in the MR-SAR model. York, DERS University of York Working Paper

Beyer, R. and F. Smets, 2015), Labour market adjustments in Europe and the US: How different? European Central Bank. Working Paper Series 1767.

Blanchard, O., and Katz, L., 1992, Regional evolutions. *Brookings Papers on Economic Activity*, **1** 1-75.

Brady, R., 2011, Measuring the diffusion of housing prices across space and over time. *Journal of Applied Econometrics*, **26** 213–231.

Brun, J., C. Carrère, P. Gillaumont and J. de Melo, 2005, Has distance died? Evidence from a panel gravity model. *The World Bank Economic Review* **19**, 99-120.

Cameron, G., J. Muellbauer, and A. Murphy, 2006, Was there a British house price bubble? Evidence from a regional panel. Economics Series Working Papers 276, University of Oxford

Can, A., 1990, The Measurement of Neighborhood Dynamics in Urban House Prices. *Economic Geography*, **66** 254-272

Can, A., 1992, Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, **22** 453-474,

Case, A., 1991, Spatial Patterns in Household Demand. *Econometrica*, **59** 953-965.

Choi, I., 2002, Combination unit root tests for cross-sectionally correlated panels, in Econometric Theory and Practice: *Frontiers of Analysis and Applied Research: Essays in Honor of Peter C. B. Phillips.* Cambridge University Press, Cambridge

Corrado, L. and B. Fingleton, 2012, Where is the economics in spatial econometrics? *Journal of Regional Science*, **52** 210-239

DiFonzo, T., 1990, The estimation of M disaggregate time series when contemporaneous and temporal aggregates are known. *The Review of Economics and Statistics*, **72** 178-82.

Druska, V. and W. Horrace, 2004, Generalized Moments Estimation for Spatial Panel Data: Indonesian Rice Farming. *American. Journal of. Agricultural. Economics*, **86** 185-198

Fernández E, M. Mayor and J. Rodríguez, 2009, Estimating spatial autoregressive models by GME–GCE techniques. *International Regional Science Review*, **32** 148-172.

García, C, 2012, A simple procedure for the comparison of covariance matrices. BMC *Evolutionary Biology* **12**.

Getis A. and J. Aldstadt, 2004, Constructing the spatial weight matrix using a local statistic. *Geographical Analysis*, **36** 90–104

Gil-Serrate, R., J. López-Laborda and J. Mur, 2011, Revenue Autonomy and Regional Growth: An Analysis of the 25-Year Process of Fiscal Decentralisation in Spain. *Environment and Planning A*, **43** 2626-2648.

Griffith, D., 1996, Some Guidelines for Specifying the Geographic Weights Matrix Contained in Spatial Statistical Models. In *Practical Handbook of Spatial Statistics*, edited by S. Arlinghaus. Boca Raton: CRC.

Harris, R., J. Moffat and V. Kravtsova, 2011, In search of 'W'. *Spatial Economic Analysis*, **6** 249-270.

Holly, S., M. Pesaran and T. Yamagata, 2010, A spatiotemporal model of house prices in the US. *Journal of Econometrics,* **158** 160-173.

Kelejian, H. and I. Prucha, 2007, HAC estimation in a spatial framework. *Journal of Econometrics*, **140** 131-154.

Kelejian, H. and G. Piras, 2014, Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics*, **46** 140–149.

Kooijman, S., 1976, Some Remarks on the Statistical Analysis of Grids Especially with Respect to Ecology. *Annals of Systems Research* **5**.

Ledoit, O. and M. Wolf, 2002, Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics*, **30** 1081-1102.

Lee, L-f. and J. Yu, 2012, QML Estimation of Spatial Dynamic Panel Data Models with Time Varying Spatial Weights Matrices, *Spatial Economic Analysis*, **7** 31-74.

Li J. and S. Chen, 2012, Two Sample test for high-dimensional covariance matrices. *The Annals of Statistics*, **40** 908–940.

López-Bazo, E. T. del Barrio. and M. Artis, 2005, The geographical distribution of unemployment in Spain. *Regional Studies,* **39** 305-318.

Marston, S., 1985, Two Views of the Geographic Distribution of Unemployment. *Quarterly Journal of Economics* **100**: 57–69.

Martínez, J. and L. Maza, 2003, Analysis of house prices in Spain. Bank of Spain Working Paper Series 0307.

Meen, G., 1996, Spatial aggregation, spatial dependence and predictability in the UK housing market. *Housing Studies*, **11** 345-372.

Moon, H. and B. Perron, 2004, Testing for a unit root in panels with dynamic factors. *Journal of Econometrics,* **122** 81-126,

Mur, J. and J. Paelinck, 2010, Deriving the W-matrix via p-median complete correlation analysis of residuals. *Annals of Regional Science*, **47** 253-267

Oud, J. and H. Folmer, 2008, A Structural Equation Approach to Models with Spatial Dependence. *Geographical Analysis,* **40** 152–166.

Paci, R. and S. Usai, 2009, Knowledge flows across European regions. *The Annals of Regional Science*, **43** 669-690.

Patridge M. and D. Rickman, 1997, The Dispersion of US State Unemployment Rates: The Role of Market and Non-market Equilibrium Factors. *Regional Studies* **31** 503–606.

Pedroni, P., 1999, Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics,* **61**, 653-670.

Pedroni, P., 2001, Purchasing Power Parity Tests in Cointegrated Panels. *Review of Economics and Statistics,* **83** 727-731.

Pesaran, M. H., 2004, General diagnostic tests for cross section dependence in panels. University of Cambridge, Faculty of Economics, Cambridge Working Papers in Economics No. 0435.

Pesaran, M., 2007, A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics* **22**, 265-312.

Pesaran, M. and R. P. Smith, 1995, Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics,* **68** 79-113.

Pesaran, M., Y. Shin, and R. P. Smith, 1999, Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association,* **94** 621-634.

Pesaran, M., Y. Shin and R. Smith, 2001, Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* **16**, 289-326.

Qu, X. and L-f. Lee, 2015, Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics,* **184** 209–232.

Schott, J., 2007, A test for the equality of covariance matrices when the dimension is large relative to the sample size. *Computational Statistics & Data Analysis* **51** 6535-6542.

Srivastava, M., 2007, Multivariate theory for analyzing high-dimensional data. *Journal of Japan Statistical Society* **37** 53-86.

Srivastava, M. and H. Yanagihara, 2010, Testing the equality of several covariance matrices with fewer observations than the dimension. *Journal of Multivariate Analysis*, **101** 1319-1329.

Srivastava, M., H. Yanagihara and T. Kubokawa, 2014, Tests for covariance matrices in high dimension with less sample size. *Journal of Multivariate Analysis*, **130** 289-309.

Westerlund, J., 2007, Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics.* **69** 709–748.

Zeilstra, J. and P. Elhorst, 2014, Integrated Analysis of Regional and National Unemployment Differentials in the European Union. *Regional Studies,* **48** 1739-1755.