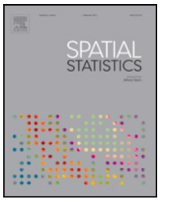
Spatial Statistics 70 (2025) 100939

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta



Dynamic spatial regimes for spatial panel data

Anna Gloria Billé [a], Roberto Benedetti [b], Paolo Postiglione [b,*]

[a] Department of Statistical Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy
[b] Department of Economic Studies, University "G. d'Annunzio" of Chieti–Pescara, Pescara, Italy

ARTICLE INFO

JEL classification:
C33
C38
C63

Keywords:
Spatio-temporal regimes
Spatial panel data model
Group-wise unobserved heterogeneity
Random effects

ABSTRACT

Spatial heterogeneity in terms of spatially-varying coefficients is often not properly considered in modeling economic data. This neglect might cause serious problems in the estimation of the parameters of a model specification when group-wise heterogeneity is at work. In this paper we propose a two-step algorithm for the identification of endogenous (data-driven) spatial regimes by using an iterative procedure that is based on weighting functions updated dynamically over time. In the first step, clusters of spatial units (i.e. spatial regimes) are defined using both space and time information. In the second step, a spatial panel data model with random effects is estimated with the spatial regimes identified in the previous step. The additional random effects assumption on the model specification ensures the possibility of controlling also for individual effects as well as group-wise slope coefficients. The proposed method is applied to two real data sets to illustrate our procedure.

1. Introduction

Unobserved heterogeneity is typically included in panel data models by considering individual and/or time specific effects that can be fixed or random, see e.g. Bai (2009). However, the way in which heterogeneity in both cross-sectional and time data can occur can be also of group-wise type. Moreover, heterogeneity can refer also to slope coefficients that vary across individuals or across time (Pesaran, 2006). Structural instability in the mean of the conditional distribution is then taking into account. In many empirical situations, researchers estimated different slope coefficients with cross-sectional subgroups of observations by using exogenous information like *classes* of observations. In some situations, no information is available to classify the observations into reasonable classes, in which cases latent class models or mixture models have been developed (Su et al., 2016).

In the spatial panel data econometrics literature, individual and/or time fixed/random effects have been also included, see e.g. Lee and Yu (2010) and Baltagi et al. (2013). One form of heterogeneity in space can be due to structural instability of the parameters (Anselin, 2010). The term *spatial regimes* (SRs) has been coined by Anselin (1988) in the context of group-wise or discrete spatial heterogeneity, referred to slope coefficients that changes among sub-groups of spatial units. Recently, to consider cases in which the group-wise heterogeneity is determined by different approaches in a data-driven context, the term *endogenous spatial regimes* (ESRs) has been also introduced (Postiglione et al., 2013; Billé et al., 2018; Anselin and Amaral, 2023). In this paper we develop an approach to identify endogenous spatial regimes into panel data models.

Within cross-sectional data, Billé et al. (2017) proposed a two-step procedure to identify endogenous spatial regimes in the first step and to also account for spatial dependence in the second step. The algorithm was developed in the spirit of the works by Cleveland and Devlin (1988) and Polzehl and Spokoiny (2000, 2006). In this paper we propose to extend their algorithm by allowing for *dynamic spatial regimes* (DSRs). The reason of considering DSRs is justified by the inherent characteristics of some

* Corresponding author.
E-mail address: paolo.postiglione@unich.it (P. Postiglione).

https://doi.org/10.1016/j.spasta.2025.100939
Received 7 March 2025; Received in revised form 12 August 2025; Accepted 22 October 2025
Available online 1 November 2025

phenomena to reveal pattern clustering that could vary also over time. The aim is to find a final partition of the spatial data through a smoothing iterative procedure that also takes into account for the past information over time. Therefore, the *dynamic* term is here used to emphasize that, during the iterative procedure, the weights used for local estimations depend also on the weights calculated at time $t-1$ by progressively using the information contained in the panel data. We do not allow, instead, for different spatial partitions of the data at each point in time (i.e. time-varying spatial regimes) , but the algorithm can be straightforwardly modified for this purpose. Notable contribution in this context is the paper by Sugasawa and Murakami (2021). We note that although their work is extremely significant in our context, it does not allow for dynamic regimes that change during the iterative procedure through information coming from the past weights and the Wald test statistics. Moreover, our algorithm not only estimates parameters during the iterative locally weighted regression (LWR) procedure, but also allows for the estimation of a spatial panel model with random effects and spatial regimes.

The rest of the paper is organized as follows. Section 2 describes the entire methodology for the identification of the DSRs and, after, the estimation of the parameters of a spatial panel data model specification with random effects. In particular, Section 2.1 explains in details the steps of the iterative algorithm to find DSRs, while Section 2.2 specifies a spatial panel data (SPD) model with DSRs and then proceeds with the estimation of the parameters of interest. Section 3 defines the short-term and long-term marginal impacts for the model in Section 2.2. Section 4 provides two empirical applications: (i) the first on cigarette demand with the dataset `Cigarette` in the R package `Ecdat`, (ii) the second one on Italian insurance data at provincial level with the dataset `Insurance` in the R package `splm` (Millo et al., 2012). In this Section we also provide some computational issues regarding the iterative procedure. Finally, Section 5 concludes.

2. Model specification and estimation strategy

In this Section we intend to explain a two-step estimation procedure able to both identify dynamic spatial regimes (DSRs) in the first step and consistently estimate structural model parameters in the second step. The following two subsections are referred to the above-mentioned two steps.

The first step in Section 2.1 explains in details an iterative statistical procedure able to find, at convergence, a partition in space of the spatio-temporal data. This algorithm is based on both locally weighted regressions and a smoothing procedure to continuously updating weights through the proper use of geographical distances among the spatial observations, Gaussian Kernel functions and Wald statistical tests. It is the natural extension for panel data of the algorithm proposed in Billé et al. (2017), in the spirit of Polzehl and Spokoiny (2000, 2006). It is important to emphasize that our method is fundamentally different from traditional LWR estimation (Brunsdon et al., 1998). Our algorithm defines an iterative procedure using LWRs at each stage and a smoothing rule that updates the weights by incorporating a dynamic temporal component. In each iteration, both the weighting structure and the coefficient estimates are refined based on statistical testing and evolving spatial dynamics. This design enables our method to go beyond the static, distance-based nature of LWR, allowing to capture for more complex, temporally adaptive spatial relationships.

Moreover, the procedure of identifying spatial regimes is also dynamic in a sense that the weights at time t depend also on the weights at time $t-1$. The *dynamic* term is here used to emphasize that, during the iterative procedure, the weights used for local estimations depend also on the weights calculated at time $t-1$ by progressively using the information contained in the panel data. Therefore, note that the term dynamic is referred to time-updated weight matrices, but then automatically this translates into dynamic spatial regimes. We do not allow, instead, for different spatial partitions of the data at each point in time (i.e. time-varying spatial regimes), but the algorithm can be straightforwardly modified for this purpose. Note that in the second step of our procedure we must estimate many slope coefficients which depend on the number of regimes each time. We will consider this option for future research.

The second step in Section 2.2 specifies a spatial panel data (SPD) model with DSRs, serial correlation and random effects, for estimating the parameters of interest. In particular, the model considers spatial dependence in the dependent variables, serial correlation and random effects in the error terms, while accounting for structural changes in space among the β coefficients. Note that these parameters are assumed to be time-invariant in the model specification, so that the estimation procedure provides global β coefficients by groups (regimes) of observations.

2.1. The algorithm: step 1

This Section describes the algorithm used to identify DSRs. This algorithm was developed in the spirit of the works by Polzehl and Spokoiny (2000, 2006), and it is the natural extension for panel data of the algorithm proposed in Billé et al. (2017). The procedure is also based on the idea of repeated local estimations (Cleveland and Devlin, 1988), i.e. estimation of the parameters with samples of observations for which different weights are assigned to different observations at each iteration. The idea is to estimate a set of parameters for a given location i in space and at time t by using linear regressions and least squares estimation.

We start by considering the following linear panel data model to be approximated

$$\mathbf{y}_t = \left(\boldsymbol{\beta}_t \odot \mathbf{X}_t\right) \mathbb{1} + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_t} \mathbf{I}\right) \quad \forall t \tag{1}$$

where $\mathbf{y}_t$ is an n-dimensional column vector at time t of continuous dependent variables, $\boldsymbol{\beta}_t$ is a n by k matrix of coefficients, i.e. we have one vector of k parameters for each observation i located in space at time t, $\odot$ is the Hadamard product operator, $\mathbf{X}_t$ is a n by k matrix of exogenous regressors at time t including the constant term, $\mathbb{1}$ is a k-dimensional column vector of ones, and $\boldsymbol{\varepsilon}_t$ is an n-dimensional column vector of innovations at time t. Obviously, the estimation of the entire matrix of coefficients $\boldsymbol{\beta}_t$ at each point

in time is not possible, since for each t we have only one observation and k parameters for each equation. To better figure out, the system of equations can be explicitly described as follows

$$\begin{aligned}
y_{1,t} &= \beta_{10,t} + \beta_{11,t}x_{11,t} + \cdots + \beta_{1k,t}x_{1k,t} + \varepsilon_{1,t} \\
y_{2,t} &= \beta_{20,t} + \beta_{21,t}x_{21,t} + \cdots + \beta_{2k,t}x_{2k,t} + \varepsilon_{2,t} \\
&\vdots \\
y_{i,t} &= \beta_{i0,t} + \beta_{i1,t}x_{i1,t} + \cdots + \beta_{ik,t}x_{ik,t} + \varepsilon_{i,t} \\
&\vdots \\
y_{n,t} &= \beta_{n0,t} + \beta_{n1,t}x_{n1,t} + \cdots + \beta_{nk,t}x_{nk,t} + \varepsilon_{n,t}
\end{aligned}$$

where the first subscript refers to the observation, the second to the parameter/variable, and finally t for time. The main point, as we will see later, is that we do not consider the entire time information to estimate the parameters, since the aim is to identify spatial regimes that progressively depends on time, i.e. the final spatial configuration obtained through our algorithm depends only on the past, i.e. $t-1$, and not the future.

To overcome the above over-parametrization problem, we provide a solution based on *local* estimation procedures. The term "local" is here used to consider only the estimation of a k by 1 vector of parameters $\boldsymbol{\beta}_{i,t}$ referred to one unit i in space (i.e. one equation) at time t exploiting a weighting matrix which is able to provide the necessary information coming from all the other observations in space.

The objective function for each local regression takes the following form

$$\mathbf{Q}\left(\boldsymbol{\beta}_{i,t}\right) = \left[\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta}_{i,t}\right]' \mathsf{W}_{i,t} \left[\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta}_{i,t}\right] \qquad \forall i,t \tag{2}$$

where $\mathsf{W}_{i,t}$ is an n-dimensional diagonal square matrix whose diagonal elements $\left(w_{i1,t}, w_{i2,t}, \ldots, w_{ij,t}, \ldots, w_{in,t}\right)$ denote the weights of each of the observed spatial data for the local regression at location i at time t. We will therefore have n diagonal spatial weighting matrices and n sets of local parameter estimates for each point in time. Then, a locally weighted least squares (LWLS) estimator is simply obtained by repeated weighted least squares

$$\hat{\boldsymbol{\beta}}_{i,t} = \left(\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{X}_t\right)^{-1}\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{y}_t \qquad \forall i,t \tag{3}$$

where the diagonal matrices $\mathsf{W}_{i,t}$ are updated throughout the smoothing procedure. Local estimates of $\sigma^2_{\varepsilon_{i,t}}$ are obtained as

$$\hat{\sigma}^2_{\varepsilon_{i,t}} = \frac{\left(\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}_{i,t}\right)' \mathsf{W}_{i,t} \left(\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}_{i,t}\right)}{n - 2v_1 + v_2} \qquad \forall i,t. \tag{4}$$

where $\hat{\boldsymbol{\beta}}_{i,t}$ is obtained via Eq. (3) and it is a k-dimensional column vector, $v_1 = tr(\mathbf{S}_t)$, $v_2 = tr(\mathbf{S}_t'\mathbf{S}_t)$ and $\mathbf{S}_t = \mathbf{X}_t\left(\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{X}_t\right)^{-1}\mathbf{X}_t'\mathsf{W}_{i,t}$.

In order to ensure the local estimation of the parameters the following assumption should be made.

Assumption 1. $\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{X}_t$ has a full column rank for all i and t.

Assumption 1 ensures the invertibility of the matrix $\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{X}_t$ in Eq. (3). It can be that for some iterations of the procedure the inverse does not exist. In order to overcome this problem, for those iterations in which the inverse does not exist, we calculate the Moore–Penrose generalized inverse $\mathbf{A}^+$ of $\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{X}_t$. Indeed, for some iterations it can be that some weights into the main diagonal of $\mathsf{W}_{i,t}$ are very close to zero. Therefore, computationally speaking, the algorithm is not able to ensure the invertibility of the matrix $\mathbf{X}_t'\mathsf{W}_{i,t}\mathbf{X}_t$. This translates in a lower number of observations, i.e. less information, taken from the other units in space.

We first need to initialize the updating procedure by defining initial weights in $\mathsf{W}_{i,t}$ as functions of distances among pairs of observations d_{ij} and a bandwidth value b. The neighborhood of observation i is defined by the value of b which determines the number of observations (i.e., a subsample) that receive a weight in building the estimate for i. The initial weights at iteration 0 are then

$$w_{ij}^0 = \mathcal{K}\left(d_{ij}; b^{opt}\right) \tag{5}$$

where $\mathcal{K}(.)$ is the Gaussian Kernel function which determine how rapidly these weights decline as distance increases, and b^{opt} is the optimal bandwidth value. Intuitively, when selecting a subsample of observations there is a trade-off between bias and variance (Brunsdon et al., 1996): because of the sample size reduction, the standard error of $\hat{\boldsymbol{\beta}}_{i,t}$ will increase, while the opposite will occur if instead we increase the subsample size. Due to this trade-off, since the LWLS estimate for point i, i.e. $\hat{\boldsymbol{\beta}}_{i,t}$, is only an approximation of the true value of the parameters in the same location $\boldsymbol{\beta}_{i,t}$, one can consider observations very closed to i providing them higher weights, with the justification that it is reasonable to assume a lower difference in the true magnitude of the $\boldsymbol{\beta}_{i,t}$-values among closed locations rather than among the ones far away. Exploiting the information coming from the locations in space, homogeneity among the $\boldsymbol{\beta}_{i,t}$-values is more pronounced between units close each other. For instance, it can be that enterprises behave similarly if they are all in a specific area and differently if they are not (Billé et al., 2018).

The Akaike information criterion (AIC), based on a trade-off between goodness of fit of the specified model and degrees of freedom, is a way to select b^{opt}, i.e.

$$b^{opt} = \min_b AIC = \min_b 2n\ln\left(\hat{\sigma}_{\varepsilon,t}\right) + n\ln(2\pi) + n\frac{n + \operatorname{tr}\left(\mathbf{S}_t\right)}{n - 2 - \operatorname{tr}\left(\mathbf{S}_t\right)} \tag{6}$$

where $\mathbf{S}_t$ is defined above and $t = 1$. Note that, in a panel data setting we can ideally define an optimal bandwidth value for each point in time. However, our procedure is based on defining the initial weights and the optimal bandwidth value only at time $t = 1$ since our smoothing procedure starts at $t = 1$ and then concludes when $t = T$. Finally, the adaptive bandwidth rather than a fixed bandwidth is used. The reason is that the adaptive bandwidth guarantees the same number of observations (i.e., the same amount of information) for each local estimation. By minimizing Eq. (6) we then find the b-smallest distances for each initial local estimation. Each element of the geographical distances is then normalized through the maximum distance. In this way, through an appropriate Kernel function, the weights that contribute to each local estimation do not depend on the effective position of the points in space.

Once the initial weights are defined, we calculate the initial estimates $\left(\hat{\boldsymbol{\beta}}^0_{i,t}, \sigma^0_{\varepsilon_{i,t}}\right)$ through Eqs. (3) and (4). From the second step and until the condition $|w^{l-1}_{ij,t} - w^l_{ij,t}| < \omega \;\; \forall i,j \;\; i \neq j$ holds for each time t, with ω a fixed small value, we compute updated weights for each iteration, i.e. $w^l_{ij,t} \;\; \forall l,t$. At the same iteration, say l, and the same time t, we compare the local estimates obtained from different spatial units, i.e. $\hat{\boldsymbol{\beta}}^l_{i,t}$ with $\hat{\boldsymbol{\beta}}^l_{j,t} \;\; \forall i,j \;\; i \neq j$ and $\forall l,t$, by using the following Wald test statistics

$$\chi^l_{ij,t} = \left(\hat{\boldsymbol{\beta}}^l_{i,t} - \hat{\boldsymbol{\beta}}^l_{j,t}\right)' \left(\boldsymbol{\Sigma}^l_t\right)^{-1} \left(\hat{\boldsymbol{\beta}}^l_{i,t} - \hat{\boldsymbol{\beta}}^l_{j,t}\right) \qquad \forall i,t \tag{7}$$

where

$$\boldsymbol{\Sigma}^l_t = \frac{\left[tr\left(\mathbf{W}^l_{i,t}\right)\boldsymbol{\Sigma}^l_{i,t} + tr\left(\mathbf{W}^l_{j,t}\right)\boldsymbol{\Sigma}^l_{j,t}\right]}{tr\left(\mathbf{W}^l_{i,t}\right) + tr\left(\mathbf{W}^l_{j,t}\right)}$$

is the weighted average of the two covariance matrices of $\hat{\boldsymbol{\beta}}^l_{i,t}$ and $\hat{\boldsymbol{\beta}}^l_{j,t}$, i.e.

$$\boldsymbol{\Sigma}^l_{i,t} = \left[\left(\mathbf{X}_t\mathbf{W}^l_{i,t}\mathbf{X}_t\right)^{-1}\mathbf{X}'_t\mathbf{W}^l_{i,t}\right]\left[\left(\mathbf{X}_t\mathbf{W}^l_{i,t}\mathbf{X}_t\right)^{-1}\mathbf{X}'_t\mathbf{W}^l_{i,t}\right]'\left(\hat{\sigma}^2_{\varepsilon_{i,t}}\right)^l$$

and

$$\boldsymbol{\Sigma}^l_{j,t} = \left[\left(\mathbf{X}_t\mathbf{W}^l_{j,t}\mathbf{X}_t\right)^{-1}\mathbf{X}'_t\mathbf{W}^l_{j,t}\right]\left[\left(\mathbf{X}_t\mathbf{W}^l_{j,t}\mathbf{X}_t\right)^{-1}\mathbf{X}'_t\mathbf{W}^l_{j,t}\right]'\left(\hat{\sigma}^2_{\varepsilon_{j,t}}\right)^l,$$

respectively. Each statistic $\chi^l_{ij,t}$ follows a χ^2_k distribution with k degrees of freedom. This result is ensured by assuming that each $\hat{\boldsymbol{\beta}}^l_{i,t} \sim \mathcal{N}\left(\boldsymbol{\beta}^l_{i,t}, \boldsymbol{\Sigma}^l_{i,t}\right)$, which is a standard result if the number of observations is reasonably large in all defined subsamples (Ibragimov and Müller, 2010; Bester et al., 2011). We also assume that $\hat{\boldsymbol{\beta}}^l_{i,t}$ is approximately independent of $\hat{\boldsymbol{\beta}}^l_{j,t} \;\; \forall i,j$ at iteration l and time t, since the variables $y_{i,t} \;\; \forall i$ are assumed to be iid Normally-distributed.

Larger values of $\chi^l_{ij,t}$ means higher distances in values between $\hat{\boldsymbol{\beta}}^l_{i,t}$ and $\hat{\boldsymbol{\beta}}^l_{j,t}$, which implies lower weights at each next iteration. In this case, the probability that units i and j at time t belong to different groups/regimes increases. The initial assigned weights are then updated in a sense that they decrease if the values of the test statistics are large, and increase otherwise. During the iterative procedure, we also allow the weights at time t to depend on the weights at time $t-1$. Formally, the updated weights are

$$w^l_{ij,t} = \mathcal{K}\left(d^l_{ij}; b^{opt}\right)\mathcal{K}_t\left(\chi^l_{ij,t}; \tau\right)w_{ij,t-1} \tag{8}$$

where $\mathcal{K}\left(d^l_{ij}; b^{opt}\right)$ are the initial weights with $d^l_{ij} = d_{ij}/(lt)$ and $l = 1, \ldots, L$, $\mathcal{K}_t\left(\chi^l_{ij,t}; \tau\right)$ is the Gaussian Kernel function of the $\chi^l_{ij,t}$ statistic, and $w_{ij,t-1}$ are the weights at time $t-1$. Note that when $t = 1$ then $w_{ij,t-1} = 1$, so the algorithm starts with the calculation of the weights at different iterations considering only the information coming from the first cross–sectional data of the entire panel. Then, from $t = 2$, the weights used for local estimations depend also on the weights calculated at time $t-1$, which are specific of locations i and j, by progressively using the information contained in the panel data, and for which we define the spatial regimes as *dynamic*. The term $d^l_{ij} = d_{ij}/(lt)$ is divided by both the iteration l and the time t to progressively reduce the weight associated to pure geographical distances, while τ is a parameter that scales the test statistic. Since that each $\chi^l_{ij,t}$ asymptotically follows a χ^2_k distribution, it is reasonable to choose τ as a function of the α–quantile of the χ^2_k distribution with k degrees of freedom (Polzehl and Spokoiny, 2006). In our case, we set this parameter equal to $\tau = \{0.0005, 0.001, 0.0015\}$. The weights are finally re-updated by averaging them with those obtained in the previous iteration

$$\tilde{w}^l_{ij,t} = (1-\eta)\tilde{w}^{l-1}_{ij,t} + \eta w^l_{ij,t}. \tag{9}$$

Eq. (9) can be seen as an *exponential smoothing* process, similar to the one in time series, whose stability is reached faster or lower depending on the value assumed by the smoothing parameter $0 < \eta < 1$. In this paper we consider also different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$.

The iterative procedure can be summarized in the following steps:

(1) Define the initial weights w^0_{ij} by using Eqs. (5) and (6) and compute the initial local estimates $\left(\hat{\boldsymbol{\beta}}^0_{i,t}, \hat{\sigma}^0_{\varepsilon_{i,t}}\right)$ through Eqs. (3) and (4). From iteration $l = 1$ the initial weights are defined as $\mathcal{K}\left(d^l_{ij}; b^{opt}\right)$ with $d^l_{ij} = d_{ij}/(lt)$ and $l = 1, \ldots, L$.

(2) At each iteration, say l, simultaneously compare the local estimates $\left(\hat{\boldsymbol{\beta}}^l_{i,t}, \hat{\boldsymbol{\beta}}^l_{j,t}\right)$, $\forall i,j = 1, \ldots, n, i \neq j$ and $\forall t$, by using Wald test statistics of Eq. (7).

(3) Calculate new weights as the product of the initial weights, a Kernel function of the Wald test statistics and the weights at time $t-1$ as in Eq. (8).
(4) To stabilize the convergence procedure, re-update the weights as in Eq. (9).
(5) Repeat steps (1)–(4) till the condition $\max|w_{ij,t}^{l-1} - w_{ij,t}^{l}| < \omega \;\; \forall i,j,i \neq j$ holds for each time t, where ω is a fixed small number.

Repeated local estimations are allowed by using locally weighted least square estimators, which evolve weights that need to be updated. The smoothing updating procedure of the weights finally converge to a vector of 0,1 values, where the weights that correspond to 1 are assigned to the units that form a cluster (or regime), and 0 is used for all the other units.

2.2. Model with DSRs: step 2

In this Section we specify a spatial panel data (SPD) model with DSRs. Among the panel model alternatives without spatial regimes proposed in the literature, we consider a SPD model with serial correlation and individual random effects. We then extend this model specification to account for DSRs. In this way we are able to identify two distinct heterogeneous effects: the former induced by the clustering process over space and time in order to highlight group-wise unobserved heterogeneity among the β coefficients, while the latter determined by the inclusion of the individual random effects in order to highlight unit-specific unobserved heterogeneity.

Let $\mathbf{y}_t$ be an n-dimensional vector at time t of continuous dependent variables. The spatial panel data (SPD) model with DSRs and individual random effects is specified as follows

$$\begin{aligned} &\mathbf{y}_t = \rho \mathsf{M}\mathbf{y}_t + \tilde{\mathbf{X}}_t\tilde{\boldsymbol{\beta}} + \mathbf{u}_t, \\ &\mathbf{u}_t = \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(0, \sigma_e^2\mathbf{I}\right) \end{aligned} \tag{10}$$

where $\mathsf{M}\mathbf{y}_t$ is the autoregressive process at time t in the dependent variables with spatial coefficient ρ, $\boldsymbol{\alpha} \sim \mathcal{N}\left(\underline{0}, \sigma_\alpha^2\mathbf{I}\right)$ are the individual random effects,

$$\tilde{\mathbf{X}}_t = \begin{bmatrix} \mathbf{X}_{1,t} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \mathbf{X}_{c,t} \end{bmatrix}$$

is a n by k block diagonal matrix of exogenous regressors at time t with the subscript c that identifies the total number of clusters/regimes, and

$$\tilde{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_c \end{bmatrix}$$

are their coefficient values. Looking at the structures of $\tilde{\mathbf{X}}_t = \{\mathbf{X}_{i,t}\}$ and $\tilde{\boldsymbol{\beta}} = \{\boldsymbol{\beta}_i\}$, the generic matrix element $\mathbf{X}_{i,t}$ is then the matrix of regressors referred to cluster/regime i at time t for which the number of observations taken is unknown before the first step of the estimation procedure. In the same way, the generic element $\boldsymbol{\beta}_i$ is a k-dimensional column vector referred to cluster/regime i. Both $\tilde{\mathbf{X}}_t$ and $\tilde{\boldsymbol{\beta}}$ are then partitioned matrices.

In order to obtain the reduced form model and properly estimate the parameters in Eq. (10), the following regularity conditions are assumed.

Lemma 2.1. *(Kelejian and Prucha, 2010) Let $\overline{\tau}$ denotes the spectral radius of the square n-dimensional M matrix, i.e.: $\overline{\tau}_{\mathsf{M}} = max\{|\omega_1|, \dots, |\omega_n|\}$, where $\omega_1, \dots, \omega_N$ are the eigenvalues of M, respectively. Then, $\mathbf{A}_\rho := (\mathbf{I} - \rho\mathsf{M})$ is nonsingular for all values of ρ in the interval $\left(-1/\overline{\tau}, 1/\overline{\tau}\right)$.*

Assumption 2 (a)**.** The spatial weights matrix M is a row-stochastic time-invariant square matrix with zero diagonal elements. (b) $\rho \in \left(-1/\overline{\tau}, 1/\overline{\tau}\right)$.

Assumption 3. Matrices M and $\mathbf{A}_\rho^{-1}$ are uniformly bounded in both row and column sum norms for all $\rho \in \left(-1/\overline{\tau}, 1/\overline{\tau}\right)$ (see e.g. Lee (2004), Assumption 5, page 1902).

Assumption 4. $\tilde{\mathbf{X}}_t$ has a full column rank for all t.

The reduced form model can be written as

$$\begin{aligned} &\mathbf{y}_t = \mathbf{A}_\rho^{-1}\tilde{\mathbf{X}}_t\tilde{\boldsymbol{\beta}} + \mathbf{A}_\rho^{-1}\mathbf{u}_t, \\ &\mathbf{u}_t = \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(0, \sigma_e^2\mathbf{I}\right) \end{aligned} \tag{11}$$

The estimation procedure is based on the maximum likelihood estimator (Millo, 2014), by using the function *spml* in the R package `splm` (Millo et al., 2012). After parameter estimation we select the best model for different values of τ and η by using AIC and BIC criteria for spatial panel data model with random effects, and the spatial Chow test as a form of Likelihood Ratio (LR) test (Anselin, 1988).

3. Marginal effects

After parameter estimation, when dealing with panel data it is of particular interest to consider the marginal effects with respect to a continuous regressor, say $\mathbf{x}_{h,t}$. In this case, since our panel model is not dynamic, we only consider the definition of the short-term effects, but definition of the long-term effects for dynamic models is straightforward. If in addition the panel data model accounts for spatial dependence in the dependent variables, those effects can be split into direct effects and indirect effects (Debarsy et al., 2012).

The short-term marginal effects with respect to $\mathbf{x}_{h,t}$ are

Definition 3.1 (*Regime-Specific Short-Term Marginal Effects*)**.**

$$\frac{\partial E\left(\mathbf{y}_t|\mathbf{X}_{i,t}\right)}{\partial \mathbf{x}_{ih,t}} = \mathbf{A}_\rho^{-1}\left[\beta_{ih}\right] \quad \forall i$$

where β_{ih} is the coefficient referred to regime i and variable h. This definition is coherent with the one proposed by LeSage and Chih (2016, Eq. (5)) for heterogeneous SAR (HSAR) panel data models, with the difference that in our case we estimate regime-specific beta coefficients and the spatial autoregressive coefficient is global. Other definitions of the marginal impacts related e.g. to spatial Durbin panel data models with regimes are defined in LeSage and Sheng (2014), and applied for instance by Benedetti et al. (2020). In the spatial Durbin specification complications arise in determining the block or sub-matrix of the spatial weighting matrix to be premultiplied to the regressors. Statistical significance of these impacts can be provided by different methods like the delta method, bootstrapping, or simulations. In this paper we use the function *impacts* in the R package `spatialreg` (Bivand and Piras, 2015), which also provides simulated p-values for each average marginal effect.

4. Empirical application

In this Section we describe the two datasets used to highlight the performances of our algorithm in identifying DSRs and we specify their relative model specifications. Both the two datasets are free available in R. This first dataset is called `Cigarette` in the package `Ecdat`, while the second dataset is called `Insurance` in the package `splm`.

4.1. Data and model specifications

The dataset `Cigarette` is a panel dataset about cigarette consumption with 48 cross-sectional units (United States) and 11 years of observations, from 1985 to 1995. Notable empirical applications on cigarettes demand are e.g. Kelejian and Piras (2014), Debarsy et al. (2012), Baltagi and Li (2004), Baltagi and Griffin (2001).

In the first step estimation, we specify the linear panel model as in Eq. (1) to be approximated by using the iterative procedure as

$$ln(\texttt{packpc})_t = \beta_1 ln(\frac{\texttt{avgprs}}{\texttt{cpi}})_t + \beta_2 ln(\frac{\texttt{income} \times \texttt{cpi}}{\texttt{pop}})_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_t} I\right) \tag{12}$$

where the dependent variable is the logarithm of the number of cigarette packs per capita for different states at time t and can be considered as a measure of real per capita cigarette sales. The covariates are the logarithm of the average price during fiscal year, including sales taxes, over the consumer price index at time t and the logarithm of the income per capita multiplied by the consumer price index at time t. Therefore, they are expressed in real terms and the estimated coefficients represent elasticities.

For the second step estimation we refer to the model in Eq. (10). The time-invariant weighting matrix $\mathsf{M} = \{m_{ij}\}$ is built using a queen contiguity criterion, i.e. two states are close to each other if they share a boundary. The idea of applying a spatial model to the cigarettes demands derives from the fact that prices vary across states mainly due to variations in state taxes on cigarettes. Thus, consumers living near state borders may decide to buy cigarettes in neighboring states if the price is lower. This consumer behavior represents a clear justification for considering spatial dependence in the statistical model.

The second empirical application is related to the dataset `Insurance`. The dataset consists of insurance consumption across 103 Italian provinces from 1998 to 2002. For this dataset we follow the specification used in the paper (Millo and Carmeci, 2011), that we report in the following

$$\begin{aligned} ln(\texttt{ppcd})_t = {} & \beta_1 ln(\texttt{rgdp})_t + \beta_2 ln(\texttt{bank})_t + \beta_3 ln(\texttt{den})_t + \beta_4 \texttt{rirs}_t + \beta_5 ln(\texttt{agen})_t + \beta_6 \texttt{school}_t + \beta_7 \texttt{vaagr}_t + \\ & + \beta_8 ln(\texttt{fam})_t + \beta_9 ln(\texttt{inef})_t + \beta_{10} ln(\texttt{trust})_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_t} I\right) \end{aligned} \tag{13}$$

where the dependent variable is the logarithm of the real per capita premiums in 2000 euros, non-life insurance excluding mandatory motor third-party liability at time t, while the covariates at time t are: the real per-capita GDP (in logarithms), real per-capita bank deposits (in logarithms), population density per square Km population (in logarithms), real interest rate on lending to families and small enterprises, density of insurance agencies per 1000 inhabitants (in logarithms), share of people with second grade schooling or more, share of value added in agriculture, average number of family members (in logarithms), judicial inefficiency index (average years to settle first degree of civil case, in logarithms), survey result to the question "do you trust others?" (in logarithms).

The time-invariant weighting matrix $\mathsf{M} = \{m_{ij}\}$ is here built using a k-nearest neighbor approach with $k = 8$, and then normalized such that $\sum_j m_{ij} = 1$. We supposed $k = 8$ as a reasonable number of neighboring provinces to choose considering the total number of Italian provinces, since the global mean of Italian provinces in each region is 5.24 (Billé et al., 2023).

Table 1
AIC, AICadj, BIC values for different values of τ and η with the Cigarette data. The lowest values are highlighted in blue.

	Clusters with $\tau = 0.0005$											
	No Clusters	$\eta = 0.25$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	$\eta = 0.45$	$\eta = 0.50$	$\eta = 0.55$	$\eta = 0.60$	$\eta = 0.65$	$\eta = 0.70$	$\eta = 0.75$
AIC	−1417.605	−1424.260	−1425.841	−1432.608	−1432.317	−1427.967	−1443.864	−1439.536	−1439.653	−1456.844	−1461.080	−1475.220
AICadj	−1417.559	−1423.654	−1425.236	−1432.002	−1431.711	−1427.030	−1442.927	−1437.710	−1437.826	−1455.501	−1459.254	−1473.394
BIC	−1396.260	−1364.493	−1366.074	−1372.841	−1372.549	−1355.393	−1371.290	−1341.347	−1341.463	−1371.462	−1362.891	−1377.031
	Clusters with $\tau = 0.0010$											
	No Clusters	$\eta = 0.25$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	$\eta = 0.45$	$\eta = 0.50$	$\eta = 0.55$	$\eta = 0.60$	$\eta = 0.65$	$\eta = 0.70$	$\eta = 0.75$
AIC	−1417.605	−1424.260	−1425.841	−1432.608	−1432.317	−1427.967	−1443.864	−1439.536	−1444.017	−1461.080	−1461.080	−1463.535
AICadj	−1417.559	−1423.654	−1425.236	−1432.002	−1431.711	−1427.030	−1442.927	−1437.710	−1442.191	−1459.254	−1459.254	−1461.149
BIC	−1396.260	−1364.493	−1366.074	−1372.841	−1372.549	−1355.393	−1371.290	−1341.347	−1345.828	−1362.891	−1362.891	−1352.538
	Clusters with $\tau = 0.0015$											
	No Clusters	$\eta = 0.25$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	$\eta = 0.45$	$\eta = 0.50$	$\eta = 0.55$	$\eta = 0.60$	$\eta = 0.65$	$\eta = 0.70$	$\eta = 0.75$
AIC	−1417.605	−1424.260	−1425.841	−1432.608	−1432.317	−1427.967	−1443.864	−1439.536	−1450.484	−1461.080	−1454.239	−1463.098
AICadj	−1417.559	−1423.654	−1425.236	−1432.002	−1431.711	−1427.030	−1442.927	−1437.710	−1448.098	−1459.254	−1451.853	−1460.074
BIC	−1396.260	−1364.493	−1366.074	−1372.841	−1372.549	−1355.393	−1371.290	−1341.347	−1339.488	−1362.891	−1343.243	−1339.294

Table 2
AIC, AICadj, BIC values for different values of τ and η with the Insurance data. The lowest values are highlighted in blue.

	Clusters with $\tau = 0.0005$											
	No Clusters	$\eta = 0.25$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	$\eta = 0.45$	$\eta = 0.50$	$\eta = 0.55$	$\eta = 0.60$	$\eta = 0.65$	$\eta = 0.70$	$\eta = 0.75$
AIC	−1240.096	−1260.506	−1249.822	−1248.724	−1233.618	−1220.475	−1220.475	−1223.182	−1222.978	−1209.882	−1209.882	−1209.882
AICadj	−1239.571	−1252.081	−1236.402	−1235.303	−1220.198	−1200.734	−1200.734	−1203.441	−1203.237	−1190.141	−1190.141	−1190.141
BIC	−1184.921	−1065.275	−1007.905	−1006.806	-991.701	-931.872	-931.872	-934.578	-934.375	-921.278	-921.278	-921.278
	Clusters with $\tau = 0.0010$											
	No Clusters	$\eta = 0.25$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	$\eta = 0.45$	$\eta = 0.50$	$\eta = 0.55$	$\eta = 0.60$	$\eta = 0.65$	$\eta = 0.70$	$\eta = 0.75$
AIC	−1240.096	−1260.506	−1249.822	−1248.724	−1233.618	−1220.475	−1234.297	−1223.182	−1222.978	−1222.978	−1222.978	−1222.978
AICadj	−1239.571	−1252.081	−1236.402	−1235.303	−1220.198	−1200.734	−1214.556	−1203.441	−1203.237	−1203.237	−1203.237	−1203.237
BIC	−1184.921	−1065.275	−1007.905	−1006.806	−991.701	−931.872	−945.694	−934.578	−934.375	−934.375	−934.375	−934.375
	Clusters with $\tau = 0.0015$											
	No Clusters	$\eta = 0.25$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	$\eta = 0.45$	$\eta = 0.50$	$\eta = 0.55$	$\eta = 0.60$	$\eta = 0.65$	$\eta = 0.70$	$\eta = 0.75$
AIC	−1240.096	−1260.506	−1249.822	−1248.724	−1233.618	−1220.475	−1234.297	−1223.182	−1222.978	−1222.978	−1222.978	−1222.978
AICadj	−1239.571	−1252.081	−1236.402	−1235.303	−1220.198	−1200.734	−1214.556	−1203.441	−1203.237	−1203.237	−1203.237	−1203.237
BIC	−1184.921	−1065.275	−1007.905	−1006.806	−991.701	−931.872	−945.694	−934.578	−934.375	−934.375	−934.375	−934.375

Table 3
Estimates of spatial panel data model with random effects.

Cigarette						Insurance					
Variable	Estimate	Std. Error	t-value	Pr(> |t|)		Variable	Estimate	Std. Error	t-value	Pr(> |t|)	
(Intercept)	4.94637	0.12866	38.4453	<2e−16	***	(Intercept)	−3.65126	0.853329	−4.2788	1.88e−05	***
lnrprice	−0.53919	0.027543	−19.5762	<2e−16	***	lnrgdp	0.46989	0.070025	6.7103	1.94e−11	***
lnrincome	0.013631	0.059153	0.2304	0.8177		lnbank	0.115591	0.03337	3.4639	0.000532	***
ρ	0.47599	0.034752	13.697	<2.2e−16	***	lnden	0.096934	0.020928	4.6317	3.63e−06	***
σ^2_α	9.1183	1.9832	4.5977	4.27e−06	***	rirs	−0.01031	0.004608	−2.2375	0.025254	*
						lnagen	0.117689	0.037664	3.1248	0.00178	**
						school	0.005315	0.002099	2.532	0.011341	*
						vaagr	−0.00816	0.003686	−2.2131	0.026892	*
						lnfam	−0.3403	0.122936	−2.7681	0.005639	**
						lninef	−0.25521	0.053933	−4.7319	2.22e−06	***
						lntrust	0.988044	0.468309	2.1098	0.034875	*
						ρ	0.410191	0.046076	8.9024	<2.2e−16	***
						σ^2_α	10.4596	1.9261	5.4303	5.63e−08	***

4.2. Results

In this Section we report the main results based on the analysis of the two datasets `Cigarette` and `Insurance`. According to the first step described in Section 2.1, we first need to obtain the initial bandwidth as in Eq. (6) to define the initial weights in Eq. (5) for the iterative procedure. By using the coordinates information and the function *bw.gwr* in the R package `GWmodel` (Gollini et al., 2015), we found the initial optimal bandwidth b^{opt} equals to 21 and 56 for `Cigarette` and `Insurance`, respectively.

Tables 1 and 2 show the AIC, adjusted AIC and BIC values for different values of $\tau = \{0.0005, 0.001, 0.0015\}$ and $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$ for `Cigarette` and `Insurance`, respectively. As expected, the BIC criterion chooses the most parsimonious model specification in both cases, i.e. the model estimated without spatial regimes. AIC and adjusted AIC values are coherent each other, choosing a more flexible model with spatial regimes. Differences can be found between the two datasets. In `Cigarette`, the combination of $\tau = 0.0005$ and $\eta = 0.75$ is to be preferred, which corresponds to the map k in Fig. 1. While, Figs. 2 and 3 show the maps for $\tau = 0.001$ and $\tau = 0.0015$, respectively. In `Insurance`, the preferred models are relative to $\eta = 0.25$, regardless of the value of τ. The relative maps are labels with a in Figs. 4–6. As we can observe, the spatial regimes in these three maps are exactly the same, leading to the same model specification in the second step.

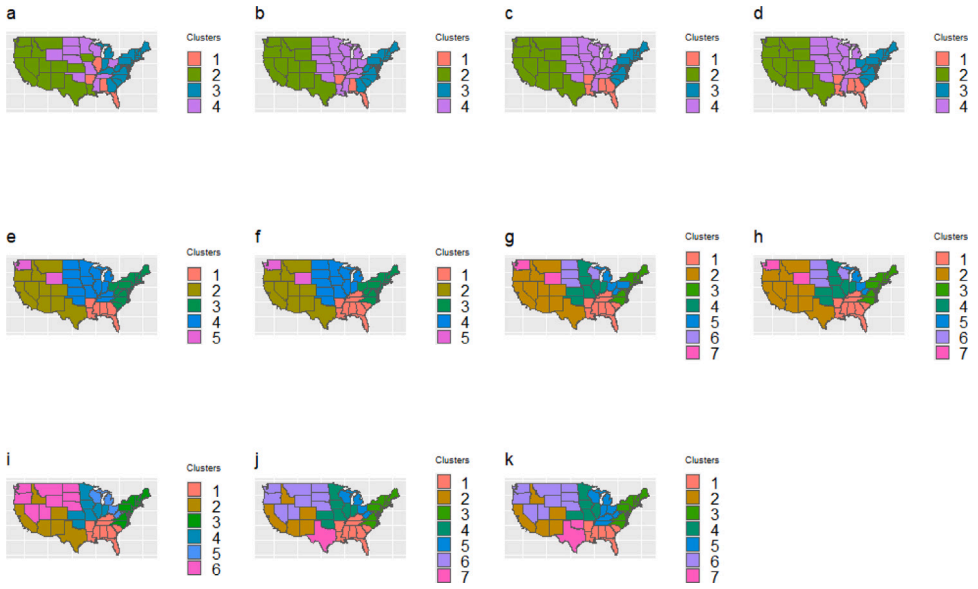


Fig. 1. Dynamic spatial regimes in US with $\tau = 0.0005$ and for different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$. The labels from (a) to (k) refer to the different values of η from 0.25 to 0.75.

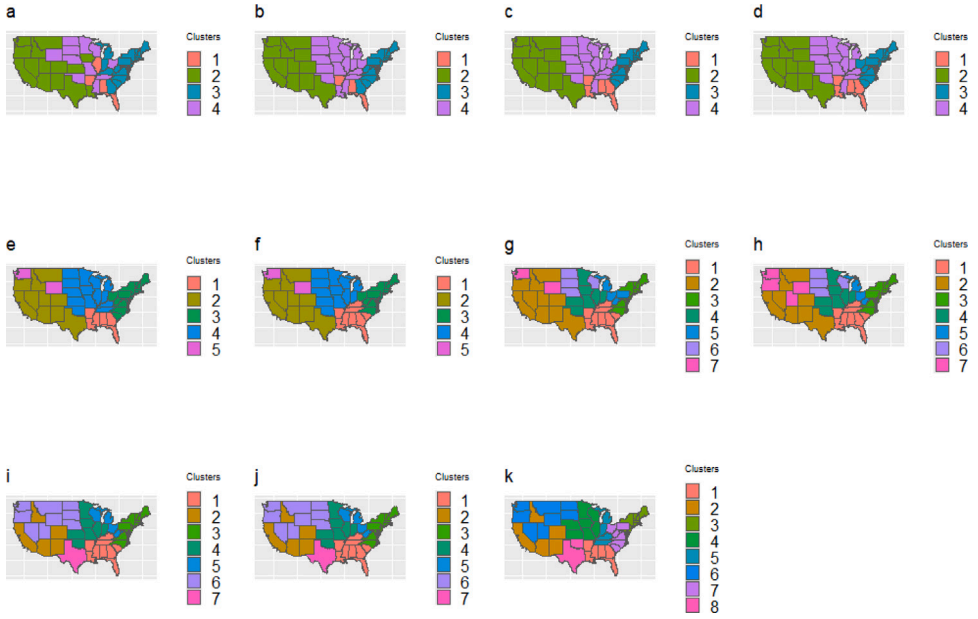


Fig. 2. Dynamic spatial regimes in Italy with $\tau = 0.001$ and for different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$. The labels from (a) to (k) refer to the different values of η from 0.25 to 0.75.

Alternatively, one can consider the distribution of the spatial Chow test statistic over different data partitions, according again to the different values of τ and η. The distributions of the spatial Chow test statistic are shown in Fig. 7 for `Cigarette` on the left and `Insurance` on the right. These distributions are similar to the F statistic distribution (Chow test) to identify the unknown potential structural break in a time series. In `Cigarette`, the maximum value reached by the spatial Chow test statistic distribution selects coherently the model at point 11, i.e. the combination of $\tau = 0.0005$ and $\eta = 0.75$. In `Insurance`, there are two maximum values that correspond to maps labeled with f in Figs. 5 and 6, which are actually the same. This can be due for instance to the fact that for finite samples, different statistics of the spatial Chow test can lead to different conclusions. Moreover, we can note that the distribution of the spatial Chow test statistic is quite the same fixing τ for `Cigarette`, while for `Insurance`, the statistic distribution with $\tau = 0.0005$ is different from the ones with $\tau = \{0.001, 0.0015\}$.

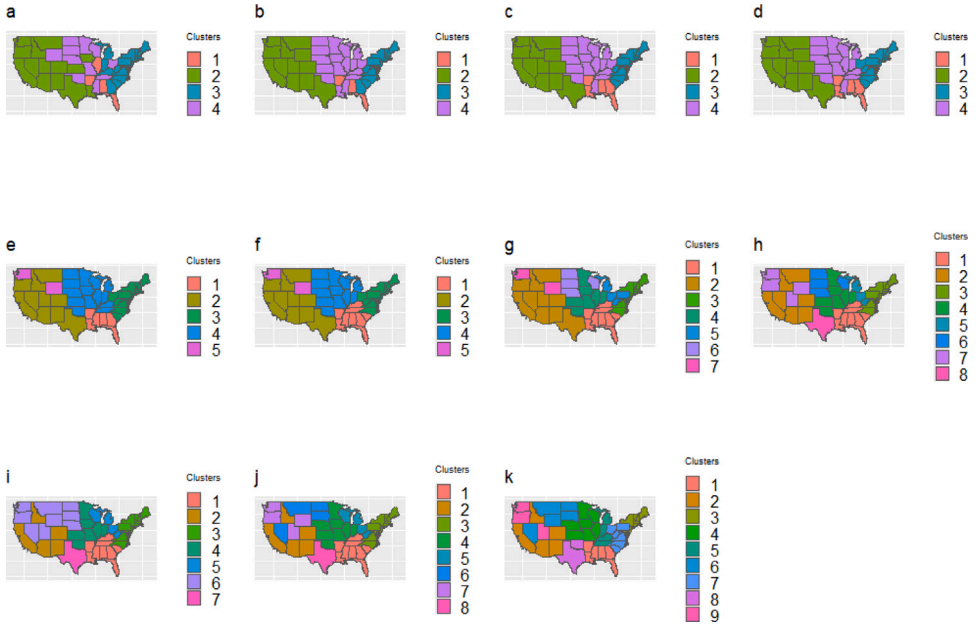


Fig. 3. Dynamic spatial regimes in US with $\tau = 0.0015$ and for different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$. The labels from (a) to (k) refer to the different values of η from 0.25 to 0.75.

In Table 3 we report the estimated coefficients of the spatial panel data models with random effects, but without spatial regimes, for the two empirical applications. While, according to the lowest AIC and AICadj values, we report the estimation of parameters of the model in Eq. (10) in Table 4. From the comparison of the two Tables, sensible differences can be found. First of all, in `Cigarette` we found 7 regimes while in `Insurance` 4 regimes. As mentioned before, the related maps are labeled with k in Fig. 1 and a in Figs. 4 for `Cigarette` and `Insurance`, respectively. When considering the spatial regimes, the magnitude and the significance of the group-wise coefficients with respect to a specific variable of the model specification can be quite different from each other, both in `Cigarette` and `Insurance` data. Moreover, in both cases, we found a coherence between the sign and the significance of the coefficients of the global model specification and the one with spatial regimes, with the following exceptions.

Specifically, looking at `Cigarette`, in Table 3 only the variable referred to income is not significant. When considering the model with spatial regimes in Table 4, both group-wise intercepts and the coefficients of the price variable are significant in all the regimes, with a coherent sign, whereas the variable income becomes significant in a subset of the total identified regimes. It is of particular interest that for some groups of states the estimated coefficients of the variable income are significant and negative, for some others they are not significant and for one specific regime the coefficient is significant and positive. Specifically, the coefficient of this variable is positive and significant in the Eastern Mid West, while it is negative and significant in the South, the Western Mid West and Northern West. The estimated spatial autocorrelation coefficient as well as the variance of the random effects are significant in both the global model and the model with spatial regimes. In particular, the estimated spatial autocorrelation coefficient ρ has a positive sign and its magnitude is similar in the two model specifications. This suggests a similar behavior in tobacco consumption among US Countries, with a propagation effects. The variance of the random effects is significant which include in our model specification an important source of individual heterogeneity.

Looking at `Insurance`, instead, in Table 3 all the variables are highly significant, with a positive and significant ρ and a significant impact of the individual random effects. When considering the model with spatial regimes in Table 4, differently from `Cigarette`, we can observe highly differences in terms of the statistical significance of the coefficients among different regimes for all the covariates. For instance, the GDP variable is not significant only in the third regime, which is referred to the southern area of Italy. This means that the southern GDP of Italy has effectively no impact on insurance consumption in the same area, and insurance consumption is probably due to are determinants. Real per-capita bank deposits is instead effectively positive and significant mostly in the central part of Italy. In the fourth regime, the share of people with second grade schooling or more, and the share of value added in agriculture have two significant coefficients with opposite sign with respect to the global model and with respect to other regimes, i.e. they are not coherent from each other. This regime includes the southern part of Sardinia, a province of Calabria and a province of Molise. For these areas, the first variable negatively affects insurance consumption, while the second one positively affects insurance consumption. The coefficients of the judicial inefficiency index and the trust variables are, instead, coherent with the global model, but they are only significant mainly in the Northern and Central Italy and only in the Northern Italy, respectively. Finally, the estimated spatial autocorrelation coefficient as well as the variance of the random effects are again significant in both the model with no spatial regimes and the model with spatial regimes. This suggests both similar behaviors in insurance consumption and a significant presence of individual heterogeneity.

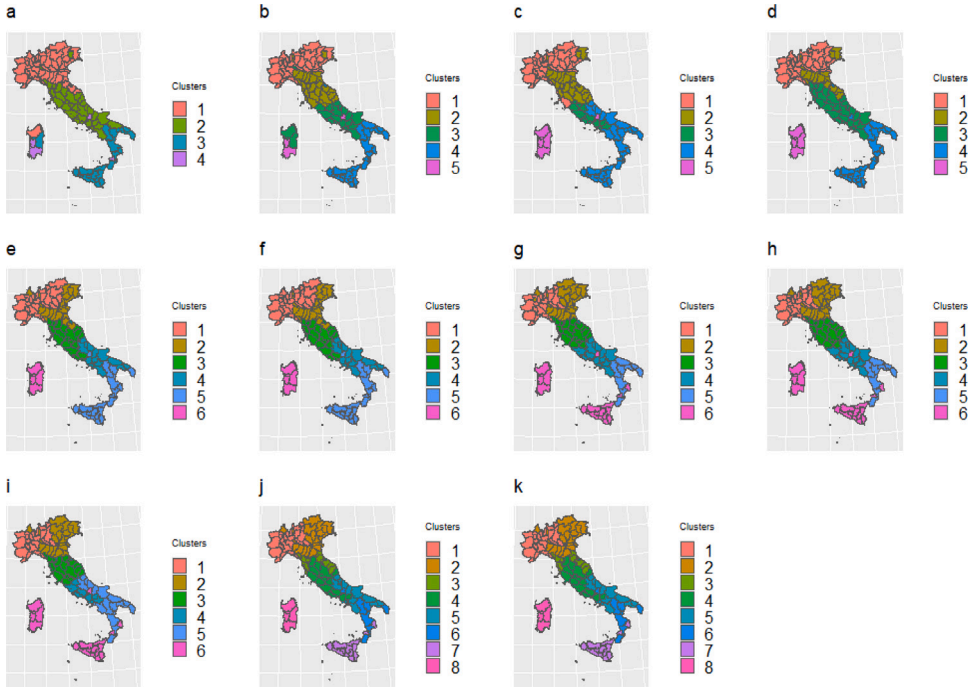


Fig. 4. Dynamic spatial regimes in Italy with $\tau = 0.0005$ and for different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$. The labels from (a) to (k) refer to the different values of η from 0.25 to 0.75.

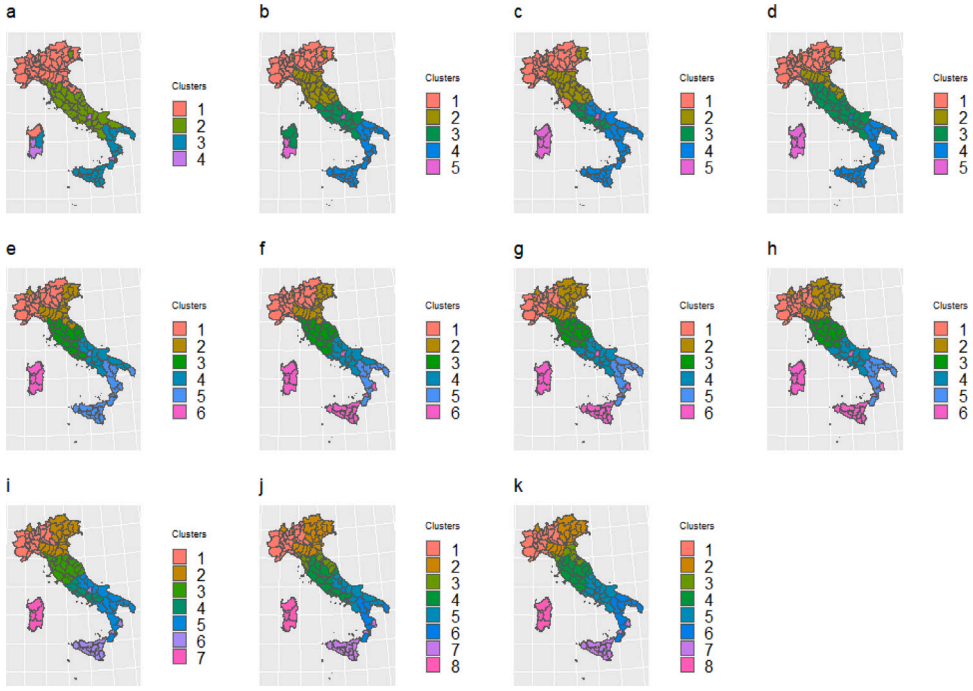


Fig. 5. Dynamic spatial regimes in US with $\tau = 0.0010$ and for different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$. The labels from (a) to (k) refer to the different values of η from 0.25 to 0.75.

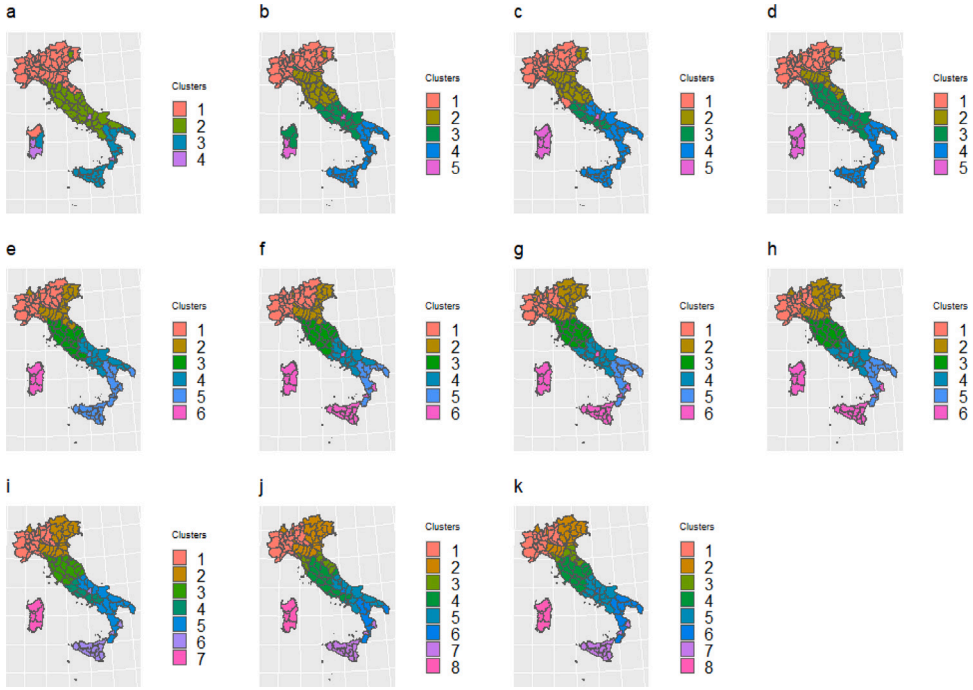


Fig. 6. Dynamic spatial regimes in Italy with $\tau = 0.0015$ and for different values of $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$. The labels from (a) to (k) refer to the different values of η from 0.25 to 0.75.

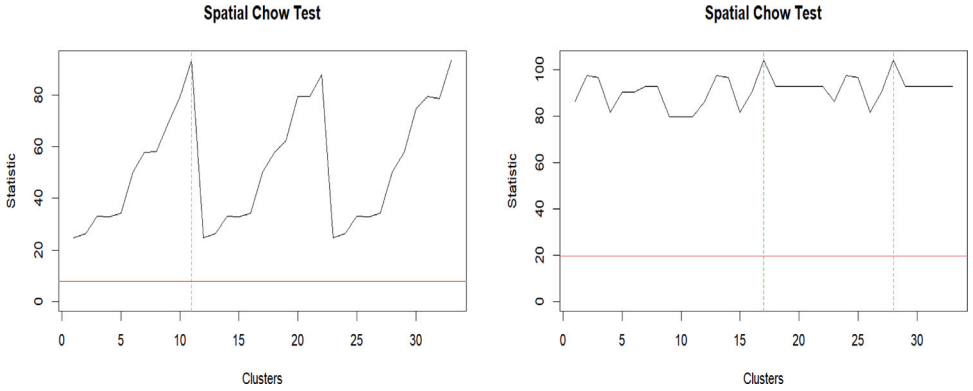


Fig. 7. Spatial Chow test statistic distributions over different data partitions. Distribution for Cigarette data on the left and distribution for Insurance data on the right. The α-quantile of a χ^2_k with k equals the number of parameters in the unrestricted model and $\alpha = 0.05$ is the red line, while the maximum value/s for the statistics are highlighted by vertical green dotted lines. From 1 to 11 $\tau = 0.0005$, from 12 to 22 $\tau = 0.0010$, from 23 to 33 $\tau = 0.0015$. For each of these values of τ, $\eta = \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$.

Finally, Table 5 summarizes the average regime-specific marginal effects in terms of the direct, indirect and total impacts, with their relative significance provided by the simulated p-values. When considering spatial panel data models, synthetic measures of this type can suggest the right interpretation of the marginal impact on the conditional expected value of y_t with respect to a small variation of a specific regressor. The significance of the marginal impacts are coherent with respect to the significance of the coefficients in the model specifications.

To what pertain the computation time, the procedure is not time-demanding. It depends more on the number of regressors in the model specification rather than on the cross-sectional and time dimension. Moreover, the values of τ and η do not report significant changes in the computation time. For the two datasets we obtain a computation time of 69.79 and 114.29 s, respectively.

Table 4
Estimates of spatial panel data model with random effects and DSRs.

Cigarette						Insurance					
Variable	Estimate	Std. Error	t-value	Pr(> $\|t\|$)		Variable	Estimate	Std. Error	t-value	Pr(> $\|t\|$)	
kk1	4.426454	0.311461	14.2119	<2.2e−16	***	kk1	−1.81963	1.447836	−1.2568	0.20883	
kk2	6.352495	0.428619	14.8208	<2.2e−16	***	kk2	−7.03084	1.491001	−4.7155	2.41e−06	***
kk3	5.625204	0.228852	24.5801	<2.2e−16	***	kk3	−1.14837	3.113021	−0.3689	0.712208	
kk4	6.611258	0.477903	13.8339	<2.2e−16	***	kk4	−4.99327	3.757579	−1.3289	0.183896	
kk5	4.456528	0.305076	14.6079	<2.2e−16	***	kk1:lnrgdp	0.300236	0.124445	2.4126	0.01584	*
kk6	5.269456	0.268218	19.6461	<2.2e−16	***	kk2:lnrgdp	0.726534	0.132645	5.4773	4.32e−08	***
kk7	9.09089	0.950404	9.5653	<2.2e−16	***	kk3:lnrgdp	0.154509	0.154406	1.0007	0.316989	
kk1:lnrprice	−0.17416	0.074846	−2.327	0.019967	*	kk4:lnrgdp	1.179921	0.320607	3.6803	0.000233	***
kk2:lnrprice	−0.71711	0.062205	−11.5282	<2.2e−16	***	kk1:lnbank	0.066331	0.043267	1.5331	0.125261	
kk3:lnrprice	−0.63167	0.046526	−13.5766	<2.2e−16	***	kk2:lnbank	0.306215	0.067454	4.5396	5.64e−06	***
kk4:lnrprice	−0.43043	0.096735	−4.4496	8.61e−06	***	kk3:lnbank	−0.05497	0.09119	−0.6028	0.546617	
kk5:lnrprice	−0.51743	0.081007	−6.3874	1.69e−10	***	kk4:lnbank	−0.17368	0.175647	−0.9888	0.32276	
kk6:lnrprice	−0.41682	0.068627	−6.0737	1.25e−09	***	kk1:lnden	0.068885	0.028452	2.4211	0.015472	*
kk7:lnrprice	−0.89121	0.108503	−8.2137	<2.2e−16	***	kk2:lnden	0.10181	0.038138	2.6695	0.007596	**
kk1:lnrincome	−0.35215	0.121963	−2.8874	0.003885	**	kk3:lnden	0.129006	0.069427	1.8581	0.063148	.
kk2:lnrincome	−0.16447	0.165616	−0.9931	0.320679		kk4:lnden	0.219292	0.102643	2.1365	0.032642	*
kk3:lnrincome	0.03696	0.111346	0.3319	0.739938		kk1:rirs	−0.01444	0.008045	−1.7949	0.07267	.
kk4:lnrincome	−0.7135	0.267931	−2.663	0.007745	**	kk2:rirs	−0.00393	0.008729	−0.4506	0.65231	
kk5:lnrincome	0.298865	0.1466	2.0386	0.041486	*	kk3:rirs	−0.00429	0.007327	−0.5855	0.558223	
kk6:lnrincome	−0.26221	0.140295	−1.869	0.061624	.	kk4:rirs	−0.00495	0.018015	−0.275	0.783302	
kk7:lnrincome	−0.86675	0.444617	−1.9494	0.051244	.	kk1:lnagen	0.01448	0.076921	0.1882	0.850687	
ρ	0.420933	0.035414	11.886	<2.2e−16	***	kk2:lnagen	0.015892	0.059425	0.2674	0.789138	
σ^2_α	9.315	2.0929	4.4507	8.56e−06	***	kk3:lnagen	0.321853	0.085553	3.762	0.000169	***
						kk4:lnagen	0.321475	0.181032	1.7758	0.075768	.
						kk1:school	0.004452	0.003073	1.4489	0.147356	
						kk2:school	0.006505	0.00375	1.7348	0.082773	.
						kk3:school	0.022291	0.005935	3.7559	0.000173	***
						kk4:school	−0.02456	0.011191	−2.1948	0.028179	*
						kk1:vaagr	−0.0118	0.007938	−1.4859	0.137316	
						kk2:vaagr	0.002761	0.008173	0.3378	0.735483	
						kk3:vaagr	−0.01554	0.005655	−2.7483	0.005991	**
						kk4:vaagr	0.039017	0.015723	2.4816	0.013081	*
						kk1:lnfam	−0.5107	0.183649	−2.7809	0.005422	**
						kk2:lnfam	−0.4239	0.223164	−1.8995	0.057498	.
						kk3:lnfam	−0.15091	0.253769	−0.5947	0.552065	
						kk4:lnfam	−1.57899	0.656438	−2.4054	0.016156	*
						kk1:lninef	−0.29134	0.09686	−3.0079	0.002631	**
						kk2:lninef	−0.22679	0.086415	−2.6244	0.008681	**
						kk3:lninef	−0.04564	0.122677	−0.3721	0.70985	
						kk4:lninef	−0.2539	0.372236	−0.6821	0.495187	
						kk1:lntrust	2.209979	0.772492	2.8608	0.004225	**
						kk2:lntrust	0.723709	0.668381	1.0828	0.278906	
						kk3:lntrust	2.257212	2.108332	1.0706	0.284343	
						kk4:lntrust	0.570154	1.65393	0.3447	0.7303	
						ρ	0.26652	0.055694	4.7854	1.71e−06	***
						σ^2_α	9.7978	1.9231	5.0949	3.49e−07	***

5. Conclusions

Spatial units can be different in size and in other characteristics, suggesting the presence of spatial heterogeneity in the data. While individual (spatial) heterogeneity is often introduced through fixed or random effects, spatial heterogeneity in terms of spatially-varying slope coefficients is one of the possible forms of heterogeneity and is often neglected in the empirical analysis of geographical data. The definition of these homogeneous groups represents, instead, a crucial issue in many regional economic studies. This negligence can lead to serious problems of misspecification of the model and inaccurate estimates. This paper aims at highlighting the role played by this form of spatial heterogeneity in economic data when the available information is based on panel data.

To this purpose, we propose an extension of an algorithm previously introduced in the literature for identifying endogenous spatial regimes for spatial panel data models. The basic idea starts by considering a transformation of the standard linear regression model in order to account for heterogeneous slope coefficients. Then, an additional flexibility on the model specification is introduced while accounting for other commonly defined effects. In the final model specification, we consider two forms of heterogeneity: the first one induced by the group-wise slope coefficients and the second one implied by the inclusion of individual random effects. This dual choice guarantees an greater control of spatial heterogeneity in the econometric model.

Our proposed algorithm is based on a two-step procedure. The first step consists in an iterative statistical method where our geographical sample is partitioned into clusters of spatial units. The algorithm is based on a continuous updating process of the

Table 5
Average regime-specific marginal impacts.

Cigarette							Insurance						
Variable	DIRECT	p-value	INDIRECT	p-value	TOTAL	p-value	Variable	DIRECT	p-value	INDIRECT	p-value	TOTAL	p-value
kk1:lnrprice	−0.18316	0.026768	−0.11761	0.036352	−0.30077	0.028207	kk1:lnrgdp	0.302868	0.018169	0.106462	0.045013	0.40933	0.019744
kk2:lnrprice	−0.75413	<2.22e−16	−0.48426	1.65e−08	−1.23839	<2.22e−16	kk2:lnrgdp	0.732904	3.74e−08	0.257626	0.001928	0.99053	3.44e−07
kk3:lnrprice	−0.66427	<2.22e−16	−0.42656	9.39e−09	−1.09083	<2.22e−16	kk3:lnrgdp	0.155863	0.298189	0.054788	0.333743	0.210651	0.301367
kk4:lnrprice	−0.45265	3.60e−05	−0.29067	0.000495	−0.74332	6.67e−05	kk4:lnrgdp	1.190266	0.000299	0.418394	0.010263	1.60866	0.000474
kk5:lnrprice	−0.54414	1.26e−09	−0.34941	4.52e−06	−0.89355	1.17e−08	kk1:lnbank	0.066912	0.124884	0.023521	0.173211	0.090433	0.129046
kk6:lnrprice	−0.43834	5.15e−10	−0.28147	1.33e−05	−0.71981	2.30e−08	kk2:lnbank	0.308899	6.82e−06	0.108582	0.002091	0.417482	9.70e−06
kk7:lnrprice	−0.93722	<2.22e−16	−0.60183	6.77e−08	−1.53905	1.47e−14	kk3:lnbank	−0.05545	0.47481	−0.01949	0.481146	−0.07495	0.472653
kk1:lnrincome	−0.37033	0.004136	−0.23781	0.010609	−0.60814	0.005313	kk4:lnbank	−0.1752	0.30985	−0.06159	0.338464	−0.23679	0.311057
kk2:lnrincome	−0.17296	0.356389	−0.11106	0.359592	−0.28402	0.355797	kk1:lnden	0.069489	0.020711	0.024426	0.058594	0.093916	0.02428
kk3:lnrincome	0.038868	0.680153	0.024958	0.690891	0.063826	0.683685	kk2:lnden	0.102703	0.006695	0.036101	0.034942	0.138804	0.008501
kk4:lnrincome	−0.75033	0.013153	−0.48182	0.02106	−1.23215	0.014704	kk3:lnden	0.130137	0.070105	0.045745	0.117381	0.175882	0.074449
kk5:lnrincome	0.314293	0.038086	0.201821	0.047448	0.516114	0.039472	kk4:lnden	0.221215	0.038976	0.07776	0.066028	0.298975	0.038973
kk6:lnrincome	−0.27575	0.05157	−0.17707	0.050556	−0.45281	0.048578	kk1:rirs	−0.01457	0.079194	−0.00512	0.117512	−0.01969	0.081711
kk7:lnrincome	−0.91149	0.041771	−0.58531	0.052327	−1.4968	0.043653	kk2:rirs	−0.00397	0.518569	−0.00139	0.522244	−0.00536	0.515619
							kk3:rirs	−0.00433	0.564987	−0.00152	0.586666	−0.00585	0.568437
							kk4:rirs	−0.005	0.774487	−0.00176	0.761586	−0.00675	0.769716
							kk1:lnagen	0.014607	0.783155	0.005134	0.782248	0.019741	0.781337
							kk2:lnagen	0.016031	0.785421	0.005635	0.779654	0.021666	0.782322
							kk3:lnagen	0.324674	3.62E−05	0.114127	0.010315	0.438802	0.000125
							kk4:lnagen	0.324293	0.051305	0.113993	0.094151	0.438287	0.055962
							kk1:school	0.004491	0.113883	0.001579	0.145409	0.00607	0.11332
							kk2:school	0.006562	0.116259	0.002307	0.159324	0.008868	0.119389
							kk3:school	0.022487	0.000165	0.007904	0.005129	0.030391	0.000154
							kk4:school	−0.02478	0.038262	−0.00871	0.07521	−0.03349	0.040353
							kk1:vaagr	−0.0119	0.113989	−0.00418	0.140755	−0.01608	0.111408
							kk2:vaagr	0.002786	0.859365	0.000979	0.874304	0.003765	0.862452
							kk3:vaagr	−0.01568	0.006265	−0.00551	0.03464	−0.02119	0.008265
							kk4:vaagr	0.039359	0.007366	0.013835	0.029525	0.053195	0.008447
							kk1:lnfam	−0.51518	0.008099	−0.18109	0.03805	−0.69627	0.009868
							kk2:lnfam	−0.42762	0.068586	−0.15031	0.119266	−0.57793	0.074481
							kk3:lnfam	−0.15223	0.618003	−0.05351	0.645406	−0.20574	0.623033
							kk4:lnfam	−1.59283	0.014476	−0.5599	0.039536	−2.15273	0.015388
							kk1:lninef	−0.2939	0.00223	−0.10331	0.020321	−0.39721	0.002916
							kk2:lninef	−0.22877	0.006338	−0.08042	0.025782	−0.30919	0.00666
							kk3:lninef	−0.04604	0.649408	−0.01618	0.678959	−0.06223	0.655176
							kk4:lninef	−0.25612	0.38321	−0.09003	0.393521	−0.34615	0.379941
							kk1:lntrust	2.229355	0.005375	0.783648	0.028193	3.013003	0.006282
							kk2:lntrust	0.730054	0.262695	0.256624	0.300488	0.986678	0.266831
							kk3:lntrust	2.277002	0.31516	0.800397	0.358307	3.077399	0.320178
							kk4:lntrust	0.575153	0.765759	0.202174	0.768843	0.777327	0.765217

weights over space and time, using geographical distances among the spatial observations and the weights obtained at time $t-1$. The final output of the algorithm is then a single spatial configuration that has also considered temporal information during the iterative procedure. The second step estimates a spatial panel data model with random effects by using the spatial regimes identified in the previous step. Given the flexibility of the algorithm, its implementation also provides a flexible tool for the use of any spatial panel data model in the analysis of group-wise unobserved heterogeneity.

Our two-step algorithm is implemented through the use of two datasets that have been used, for illustrative purposes, in several spatial econometric studies. Penalized criteria such as AIC, AICadj confirm that the more flexible spatial panel data models with dynamic spatial regimes were to be preferred in both cases, coherently with the results of the spatial Chow test.

The empirical evidence clearly demonstrates that coefficient values, and notably their signs, can differ significantly across different spatial regimes. This phenomenon is a clear manifestation of the presence of spatial non-stationarity in economic behaviors, indicating that the relationships between explanatory variables and outcomes vary depending on the territorial context. In many cases, such variation cannot be adequately captured by global models that assume homogeneity across space, as these models tend to overlook critical structural differences among the regions under study. This evidence strongly highlights the need for place-specific policy approaches. Policies defined in response to the heterogeneous economic behaviors observed across regimes cannot rely on one-size-fits-all solutions. Instead, they must be tailored to the unique characteristics and dynamics of each territory. Recognizing and addressing these localized differences is essential to formulate effective interventions that respond to the specific challenges and opportunities inherent in each spatial regime.

It is worth noting that our technique can also be viewed as a method for the identification of spatial clusters, without necessarily estimating a spatial panel data model. Moreover, even when we do not have geographical information, the algorithm can be used to identified clusters of data by properly defining the weights with other metrics. Finally, our algorithm can be easily extended to the case of identifying time-varying spatial regimes, i.e. spatial regimes that vary over time or, in other words, clusters of spatial units for each wave (time) of our panel. In this way, for example, a policy maker might evaluate the persistence of a region belonging to a certain group or its movement to another group. Future research will be dedicated to this methodological improvement.

Data availability

The data that support the findings of this study are openly available in R.

References

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Springer Science & Business Media.
Anselin, L., 2010. Thirty years of spatial econometrics. Pap. Reg. Sci. 89 (1), 3–26.
Anselin, L., Amaral, P., 2023. Endogenous spatial regimes. J. Geogr. Syst. 1–26.
Bai, J., 2009. Panel data models with interactive fixed effects. Econometrica 77 (4), 1229–1279.
Baltagi, B.H., Egger, P., Pfaffermayr, M., 2013. A generalized spatial panel data model with random effects. Econometric Rev. 32 (5–6), 650–685.
Baltagi, B.H., Griffin, J.M., 2001. The econometrics of rational addiction: the case of cigarettes. J. Bus. Econom. Statist. 19 (4), 449–454.
Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer, pp. 283–295.
Benedetti, R., Palma, D., Postiglione, P., 2020. Modeling the impact of technological innovation on environmental efficiency: a spatial panel data approach. Geogr. Anal. 52 (2), 231–253.
Bester, C.A., Conley, T.G., Hansen, C.B., 2011. Inference with dependent data using cluster covariance estimators. J. Econometrics 165 (2), 137–151.
Billé, A.G., Benedetti, R., Postiglione, P., 2017. A two-step approach to account for unobserved spatial heterogeneity. Spat. Econ. Anal. 12 (4), 452–471.
Billé, A.G., Salvioni, C., Benedetti, R., 2018. Modelling spatial regimes in farms technologies. J. Prod. Anal. 49, 173–185.
Billé, A.G., Tomelleri, A., Ravazzolo, F., 2023. Forecasting regional GDPs: a comparison with spatial dynamic panel data models. Spat. Econ. Anal. 1–22.
Bivand, R., Piras, G., 2015. Comparing implementations of estimation methods for spatial econometrics. J. Stat. Softw. 63, 1–36.
Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. 28 (4), 281–298.
Brunsdon, C., Fotheringham, S., Charlton, M., 1998. Geographically weighted regression: modelling spatial non-stationarity. J. R. Stat. Soc.: Ser. D (The Statistician) 47 (3), 431–443.
Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Amer. Statist. Assoc. 83 (403), 596–610.
Debarsy, N., Ertur, C., LeSage, J.P., 2012. Interpreting dynamic space–time panel data models. Stat. Methodol. 9 (1–2), 158–171.
Gollini, I., Lu, B., Charlton, M., Brunsdon, C., Harris, P., 2015. GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. J. Stat. Softw. 63 (17), 1–50.
Ibragimov, R., Müller, U.K., 2010. t-Statistic based correlation and heterogeneity robust inference. J. Bus. Econom. Statist. 28 (4), 453–468.
Kelejian, H.H., Piras, G., 2014. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. Reg. Sci. Urban Econ. 46, 140–149.
Kelejian, H.H., Prucha, I.R., 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. J. Econometrics 157 (1), 53–67.
Lee, L.-F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica 72 (6), 1899–1925.
Lee, L.-F., Yu, J., 2010. Estimation of spatial autoregressive panel data models with fixed effects. J. Econometrics 154 (2), 165–185.
LeSage, J.P., Chih, Y.-Y., 2016. Interpreting heterogeneous coefficient spatial autoregressive panel models. Econom. Lett. 142, 1–5.
LeSage, J.P., Sheng, Y., 2014. A spatial econometric panel data examination of endogenous versus exogenous interaction in Chinese province-level patenting. J. Geogr. Syst. 16, 233–262.
Millo, G., 2014. Maximum likelihood estimation of spatially and serially correlated panels with random effects. Comput. Statist. Data Anal. 71, 914–933.
Millo, G., Carmeci, G., 2011. Non-life insurance consumption in Italy: a sub-regional panel data analysis. J. Geogr. Syst. 13, 273–298.
Millo, G., Piras, G., et al., 2012. splm: Spatial panel data models in R. J. Stat. Softw. 47 (1), 1–38.
Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74 (4), 967–1012.
Polzehl, J., Spokoiny, V.G., 2000. Adaptive weights smoothing with applications to image restoration. J. R. Stat. Soc. Ser. B Stat. Methodol. 62 (2), 335–354.
Polzehl, J., Spokoiny, V.G., 2006. Propagation-separation approach for local likelihood estimation. Probab. Theory Related Fields 135, 335–362.
Postiglione, P., Andreano, M.S., Benedetti, R., 2013. Using constrained optimization for the identification of convergence clubs. Comput. Econ. 42, 151–174.
Su, L., Shi, Z., Phillips, P.C., 2016. Identifying latent structures in panel data. Econometrica 84 (6), 2215–2264.
Sugasawa, S., Murakami, D., 2021. Spatially clustered regression. Spat. Stat. 44, 100525.