Spatial Autoregressive Nonlinear Models in R with an Empirical Application in Labour Economics

Published Version:
Spatial Autoregressive Nonlinear Models in R with an Empirical Application in Labour Economics / Anna Gloria Billé. - STAMPA. - (2021), pp. 23-41. [10.4337/9781788976480.00008]

This version is available at: https://hdl.handle.net/11585/853802 since: 2022-02-07

Published:

DOI: http://doi.org/10.4337/9781788976480.00008

Spatial Autoregressive Nonlinear Models in R with an Empirical Application in Labour Economics

Anna Gloria Billé[1]

Department of Statistical Sciences, University of Padova, Padova, Italy

1. Introduction to Spatial Microeconometrics

Microeconomic data usually refer to individual-level data on the economic behaviour of individuals, firms or groups of them (e.g. households, industrial districts, etc.), and they are typically collected using cross-section or panel data surveys. The analysis of microeconomic data has a long history, both from a theoretical and an empirical point of view. There are several peculiarities that mark microeconomic data. First of all, data at a low disaggregated level are often discrete or censored/truncated by nature, leading to the use of **nonlinear models** such as probit/logit, Tobit, count data regressions, where the type of nonlinearity in the model specification refers to the *nonlinearity in parameters*, and for which iterative estimation procedure are necessary (McFadden 1984; Amemiya 1985; Maddala 1986). Secondly, disaggregation is often a source of *heterogeneity* that should be accounted for to obtain valid inferences in regression analyses. In addition, the growing availability of this type of data has brought out a number of additional problems, among which e.g. *sample selection*, *measurement*

---

[1] Anna Gloria Billé is Assistant Professor at Department of Statistical Sciences, University of Padova, Italy. Email: annagloria.bille@unipd.it

errors, *missing data*, see the book by Cameron and Trivedi (2005) for an overview of methods in microeconometrics. Differently from macroeconomic data, microeconomic data are more affected by these issues, but at the same time they provide a greater flexibility in specifying econometric models and a larger amount of information.

Adding spatial autoregressive processes to nonlinear models has increased the theoretical, methodological and computational problems in estimating the parameters and interpreting the results. Indeed, in a recent review, Billé and Arbia (2019) pointed out the main difficulties and the proposed solutions as well as the lack of a continuous development of this specific field in the econometric literature. Although spatial econometrics is now a consolidated discipline, most of the literature in spatial microeconometrics has only seen a growth of empirical applications with the use of micro data and spatial linear models, see e.g. Conley (2010), Arbia, Espa, and Giuliani (2018). In addition, to the best of my knowledge, methods e.g. to deal with *measurement error* problems have been developed only for spatial linear models (Le Gallo and Fingleton 2012). The argumentation and explanation of the different estimation procedures and their statistical properties according to different types of spatial limited dependent variable models is vast and somehow complicated. For a recent "*state of the art*" of the theoretical and methodological issues of spatial nonlinear models the reader is referred to Billé and Arbia (2019) and references therein.

This chapter is devoted to the analysis of spatial autoregressive nonlinear probit models, with a detailed illustration of how to use available packages in R through an empirical application in Labour Economics. The rest of the chapter is organised as follows. Section 2 lays the foundation for spatial econometric models by defining the weighting matrices. Section 3 specifies a spatial autoregressive nonlinear probit model and explains the marginal effects (subsection 3.1). Finally, Section 4 describes how to fit and interpret these type of models in R. An empirical illustration on the employments rates in Europe is also included (subsection 4.1).

2. Definition of Spatial Weighting Matrices

There are several different ways in which the spatial weights can be defined. Actually, their definition is somehow arbitrary, leading to potential problems of model misspecification due to wrongly assumed weighting matrices. This has led to an interesting debate in the literature. Many researchers began to evaluate the sensitivity of the estimated parameters on the use of different types of spatial weighting matrices through Monte Carlo simulations, see e.g. Billé (2013) in a spatial probit context, among many others. Some others critized the exogeneity assumption behind its construction, highlighting the importance of including information coming from the economic theory, see e.g. B. Fingleton and Arbia (2008) and Corrado and Fingleton (2012), or estimating the elements of this matrix through an appropriate procedure, see e.g. Pinkse, Slade, and Brett (2002), Bhattacharjee and Jensen-Butler (2013), Ahrens and Bhattacharjee (2015) and references below. Finally, J. P. LeSage (2014) argued that "*a Bayesian approach provides one way to introduce subjective prior information in choosing the weighting matrix*", while James LeSage and Pace (2014) stressed that if the partial derivatives defining the spatial marginal effects are correctly specified and the goal is to have approximately correct scalar summary measures (see subsection 3.1), then the weighting matrices have approximately no impact on estimates and inferences.

Discussions on the definition on different spatial weighting matrices as well as model specifications can now be found in several spatial book references, see e.g. Anselin (1988), JP LeSage and Pace (2009), Elhorst (2014), Arbia (2014), Kelejian and Piras (2017). A first definition comes from the use of geographical information, as the spatial data are typically associated to georeferenced data, like e.g. firms, provinces, etc. Depending on the type of spatial units, i.e. points, lines or areas, distance-based definition or contiguity-based criteria to build the saptial weight matrix (*geometric* W) can be used, see e.g. Getis and Aldstadt (2004). Contiguity-based criteria typically define a *sparse* weighting matrix, see also e.g. Pace and Barry (1997). The starting point is to define a Boolean matrix as in the following. Let $W = \{w_{ij}\}$ be the spatial weighting matrix with elements equal to the weights among pairs of random variables y_i, y_j for $i, j = 1, \dots, n$, with n the sample size, then

$$\begin{cases} w_{ij} = 1 \quad iff \quad y_j \quad is\ close\ to \quad y_i, \quad j \neq i \\ w_{ij} = 0 \quad otherwise \end{cases}$$

where the term "*is close to*" is justified by the adopted criterion. Several contiguity-based criteria, like e.g. queen and rook, fall within the above definition, as well as a distance-based approach called k-nearest neighbours that will be explained in the section devoted to the empirical application. The peculiarity of contiguity-based criteria is the use of borders or angles to select the neighbourhood of each spatial unit, and therefore they are particularly suitable for areal/polygon data. Distance-based weights define, instead, *dense* weighting matrices. In this case, the geographical information comes from the numerical value of the distance (measured in *km*, *miles*, etc.) between pairs of spatial units. A general distance-based weight matrix can be define in the following way

$$w_{ij} = f(d_{ij}), \quad j \neq i$$

where $f(.)$ can be any continuous monotonically decreasing function that ensures decreasing weights as distances d_{ij} increase. Examples of such functions are, for instance, the negative exponential or the inverse-distance. Figure 1 graphically shows the difference between *sparse* and *dense* matrices, with the non-zero elements in black.

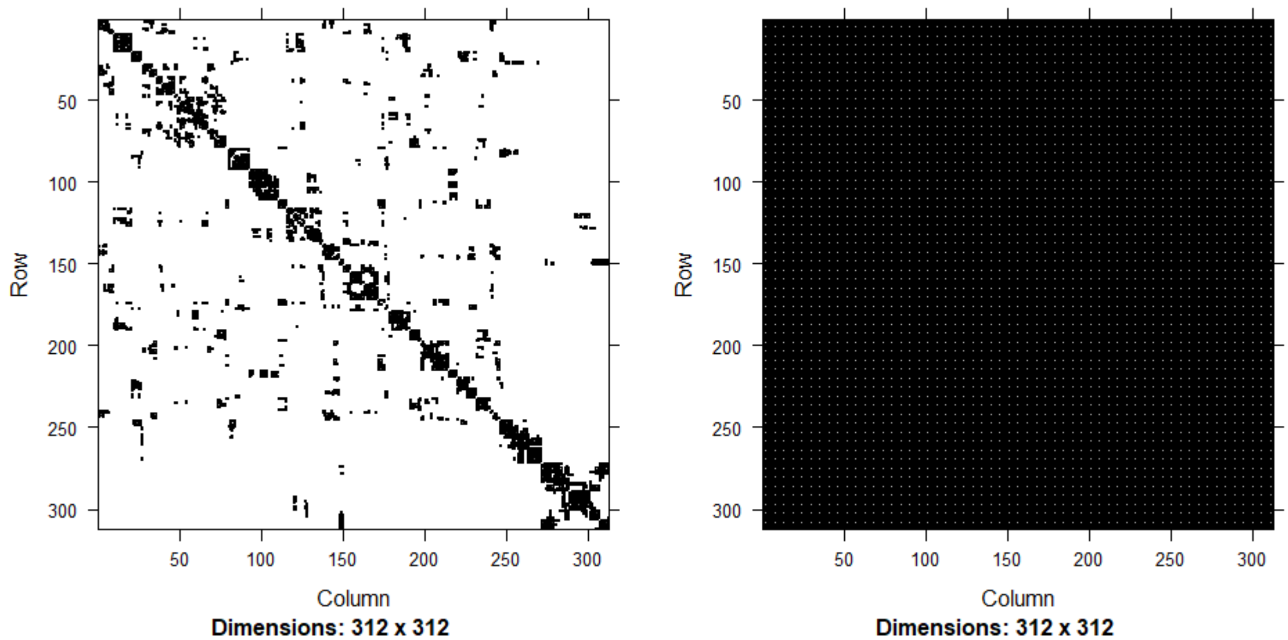

Figure 1 Sparse (left) and Dense (right) Weighting Matrices

The cut-off approach is a particular type of distance-based definition that leads to *sparse* weighting matrices. The aim is to put at zero all the weights referred to distances greater than a pre-specified distance (*radius*, r),

$$\begin{cases} w_{ij} = f\left(d_{ij}\right) \quad iff \quad d_{ij} < r, \quad j \neq i \\ w_{ij} = 0 \quad otherwise. \end{cases}$$

Regardless the way of defining the spatial weights, all the spatial weighting matrices must be such that: (i) all the diagonal elements are zero, $w_{ii} = 0 \ \ \forall i$, (ii) their row and column sums are uniformly bounded in absolute value. Assumption (i) is a simple normalization rule and means that each spatial unit is not viewed as its own neighbor, while (ii) limits the spatial correlation to a manageable degree and ensures that the spatial process is not explosive. Moreover, nonnegative weighting matrices are typically used, $w_{ij} \geq 0 \ \ \forall \ i \neq j$, as follows from the definition of *distance* (metric). For theoretical details see e.g. Kelejian and Prucha (1998), L.-F. Lee (2004), Kelejian and Prucha (2010).

W is typically row-normalized such that $\sum_j w_{ij} = 1, \forall i$. This ensures that the autoregressive term of the model lies in the interval $(-1,1)$. Indeed, the parameter space of the autoregressive term depends in general on the eigenvalues of W. Although the row-normalization rule provides an easy interpretation of the spatial model, i.e. each geo-located dependent variable depends on a weighted average of neighbouring dependent variables, it does not ensure the equivalence of the model specifications before and after normalizations of the weights, with the exception of the use of the k-nearest neighbour approach. To ensure this equivalence, an alternative normalization rule based on the spectral radius of W has been proposed by Kelejian and Prucha (2010). Finally, although the definition of *distance* requires the symmetric property, the weighthing matrix can be in some cases an asymmetric matrix, whose definition can be particularly useful in the following cases.

Spatial econometric methods can be also used to modeling, for instance, economic agents' behaviours or financial data, see Catania and Billé (2017), since the aim is to easily capture cross-sectional dependence through an appropriate parametrization. However, individual data of this type are not georeferenced, making difficult the connection between them. Economic definitions of distance can be useful in this case, see e.g. Case, Rosen, and Hines Jr (1993) and Anselin and Bera (1998). Moreover, the inclusion of spatially lagged dependent variables Wy, both in the linear and in the nonlinear model specification, adds an endogenity problem due to the simultaneity of n equations. Therefore, the exogeneity assumption behind the geographical definition of the spatial weights has been criticized e.g. by Pinkse and Slade (2010), and proper estimation methods have been developed, see Kelejian and Piras (2014) and Qu and Lee (2015). For instance, Qu and Lee (2015) proposed the use of additional exogenous regressors, say X_2, to control for potential endogeneity of the spatial weight matrix in a cross-sectional setting. The weights are defined as

$$w_{ij} = h\left(Z_{ij}, d_{ij}\right), \quad j \neq i$$

where $h(.)$ is a bounded function, Z is a matrix of variables, and d_{ij} is the distance between two units. Within a two-stage IV estimation, they first regress the endogenous matrix Z on X_2 through the OLS estimator applied to the following equation

$$Z = X_2 \mathit{\Gamma} + \varepsilon$$

where $\mathit{\Gamma}$ is the matrix of coefficients and ε is a column vector of innovations. Then, they used $(Z - X_2 \mathit{\Gamma})$ as control variables in the linear spatial model to control for the potential endogeneity of W in the second stage. Furthermore, the authors argued that the spatial weight matrix W can be exogenous/predetermined, but the term Wy remains endogenous due to the potential correlation among the error terms of the two equations.

Still within the IV approach, Kelejian and Piras (2014) directly proposed estimating the elements of W through a linear approximation with a finite set of parameters in a panel data setting. A recent alternative way to estimate the elements of the weighting matrix is based on exploiting the time information in a spatio-temporal model, see Billé and Catania (2018). This spatio-temporal model specification, combined with the generalized autoregressive score (GAS) procedure, is able to estimate via MLE the elements of the spatial weighting matrix $W = \{w_{ij}\}$ through a proper parametrization of the weights over time. In this way, the estimation procedure is able to identify the *radius* within which the spatial effects have their highest impacts. Finally, Otto and Steinert (2018) proposed the use of the least absolute shrinkage and selection operator (LASSO) approach to estimate the weighting matrix (with also structural breaks) in a spatio-temporal model. However, none of the above approaches for endogenous W matrices have been yet developed for spatial nonlinear model specifications.

3. Model Specification and Interpretation

In this section we define one of the possible general specifications of the spatial probit model, and its nested-model specifications. For details on different spatial binary probit specifications see Billé and Arbia (2019). Then, we provide the model interpretation and in subsection 3.1 the definition of the marginal effects.

Let y be a n-dimensional vector of binary dependent variables. A spatial (first-order) autoregressive probit model with (first-order) autoregressive disturbances (SARAR(1,1)-probit) (Billé and Leorato 2019; Martinetti and Geniaux 2017) can be defined in the following way

$$\begin{cases} y^* = \rho W_1 y^* + X\beta + u \quad u = \lambda W_2 u + \varepsilon \quad \varepsilon \sim \mathcal{N}(0, I) \\ y = \mathbb{1}(y^* > 0) \end{cases}$$

where y^* is a n-dimensional vector of continuous latent dependent variables, typically associated to individual unobserved utility functions, W_1 and W_2 are (possibly) two different n-dimensional spatial weighting matrices to govern different spatial processes, X is the n by k matrix of regressors, u is the n-dimensional vector of autoregressive error terms, ε is a vector of homoskedastic innovations of the same dimension with $\sigma_\varepsilon^2 = 1$ for identification, and $\mathbb{1}(.)$ is the indicator function such that $y_i = 1 \Leftrightarrow y^* > 0 \quad \forall i$. The terms $\rho W_1 y^*$ and $\lambda W_2 u$ capture the global *spatial spillover effects* in the latent dependent variables and among the shocks, respectively. The term "*spillover effects*" refers to the indirect effects due to the neighboring random variables. This concept will be clarified below. Two nested-model specifications can be easily obtained by setting $\lambda = 0$ or $\rho = 0$. The two models are called spatial (first-order) autoregressive probit (SAR(1)-probit) model and spatial (first-order) autoregressive error probit (SAE(1)-probit) model, respectively.

When we consider nonlinear binary models, we are generally interested in evaluating the changes in the probability of being equal to 1 by the binary dependent variables given the set of regressors of the model specification, i.e. $E(y|X) = P(y = 1|X)$. If we assume that the error terms are normally distributed, then a probit model can be considered and the above probabilities can be evaluated by using the normal cumulative density function, i.e. $P(y = 1|X) = \Phi(.)$, where the cdf $\Phi(.)$ is a function of the unknown parametrs $\theta = (\rho, \lambda, \beta')'$ and the exogenous regressors X. Therefore, the estimated β coefficients cannot be interpreted as the marginal effects anymore, due to both the nonlinearity in parameters and the presence of spatial dependence. In section 3.1 we explain in details how to calculate proper marginal effects for these types of models.

Due to the simultaneity of the model specification, reduced forms of the spatial models are typically derived. Under some regularity conditions, see e.g. Billé and Leorato (2019), the above model can be written in reduced form as

$$\begin{cases} y^* = A_\rho^{-1} X\beta + \nu \quad \nu \sim \mathcal{N}(0, \Sigma_\nu) \\ y = \mathbb{1}(y^* > 0) \end{cases}$$

where $\nu = A_\rho^{-1} B_\lambda^{-1} \varepsilon$, $A_\rho^{-1} = (I - \rho W_1)^{-1}$ and $B_\lambda^{-1} = (I - \lambda W_2)^{-1}$, and $\Sigma_\nu = A_\rho^{-1} B_\lambda^{-1} B_\lambda^{-1\prime} A_\rho^{-1\prime}$. Now, the inverse matrices, A_ρ^{-1} and B_λ^{-1}, can be written in terms of the infinite series expansion as

$$\begin{cases} A_\rho^{-1} = \sum_{i=0}^{\infty} \rho^i \, W_1^i = I + \rho W_1 + \rho^2 W_1^2 + \rho^3 W_1^3 + \cdots + \rho^\infty W_1^\infty \\ B_\lambda^{-1} = \sum_{i=0}^{\infty} \lambda^i \, W_2^i = I + \lambda W_2 + \lambda^2 W_2^2 + \lambda^3 W_2^3 + \cdots + \lambda^\infty W_2^\infty \end{cases}$$

where, focusing on the first expansion, the first right-hand-side term (the identity matrix I) represents a direct effect on y^* of a change in X due to the fact that its diagonal elements are ones and its off-diagonal elements are zero. On the contrary, the second right-hand-side term (ρW_1) represents an indirect effect on y^* of a change in X, also called *spillover effects*. The other terms are referred to the indirect effects of higher-orders, i.e. the spatial effects due to the neighbours of the neighbours, and so on. The term *spillover effects* is also referred to the autoregressive process in the error terms, focusing on the second finite series expansion, although the difference in the interpretation is that A_ρ^{-1} enters both in the mean and in the variance-covariance matrix of the conditional distribution of $y|X$, whereas B_λ^{-1} only in the variance-covariance matrix.

Higher-order effects have also another meaning in spatial econometrics. Indeed, the higher-order effects can be also defined through the orders of the autoregressive processes, as is typically done in time-series econometrics (consider e.g. an ARMA(1,1) vs. an ARMA(2,2)). Nevertheless, the specification and interpretation for these types of spatial linear models is not obvious as in time, and none of these specifications have been found with limited dependent variables. For details the reader is referred to Elhorst, Lacombe, and Piras (2012), Arbia, Bee, and Espa (2013), Debarsy and LeSage (2018), among others.

3.1. Marginal Effects

Marginal effects of nonlinear models differ substantially from the linear ones. Indeed, in nonlinear specifications these effects depend both on the estimated parameters and the level of the explanatory variable of interest x_h. In addition, when considering *spatial dependence*, total marginal effects have inside two distinct source of information: (i) *direct effects* due to the direct impact of the regressor x_{ih} in a specific region i on the dependent variable y_i and, (ii) *indirect effects* due to the presence of the spatial spillover effects, i.e. the impact from other regions $x_{jh} \forall j \neq i$, see J. P. LeSage et al. (2011). Note that the term *region* is a general definition of spatial data, and the observations $i = 1, \dots, n$ can also be referred to individuals, like e.g. economic agents.

In this way, the usual interpretation "*a unit variation of a specific regressor determines a specific and constant variation in the dependent variable*" is no more valid. To preserve the link between econometric models and economic theory, proper marginal effects should be always defined. In the following, we discuss the definition of the marginal effects for spatial probit models.

Let $x_{.h} = (x_{1h}, x_{2h}, \dots, x_{ih}, \dots, x_{nh})'$ be an n-dimensional vector of units referred to the h-th regressor, $h = 1, \dots, k$, and $x_{i.} = (x_{i1}, x_{i2}, \dots, x_{ih}, \dots, x_{ik})'$ be a k-dimensional vector of regressors referred to unit i. Billé and Leorato (2019), among others, propose the following specifications of the marginal effects

$$\frac{\partial \mathrm{P}(y_i = 1 \mid X_n)}{\partial x'_{.h}} \Big|_{\bar{x}} = \phi\left(\{\Sigma\}_{ii}^{-1/2}\{A_\rho^{-1}\overline{X}\}_{i.}\beta\right)\{\Sigma\}_{ii}^{-1/2}\{A_\rho^{-1}\}_{i.}\beta_h,$$

$$\frac{\partial \mathrm{P}(y_i = 1 \mid X_n)}{\partial x'_{.h}} \Big|_{x} = \phi\left(\{\Sigma\}_{ii}^{-1/2}\{A_\rho^{-1}X\}_{i.}\beta\right)\{\Sigma\}_{ii}^{-1/2}\{A_\rho^{-1}\}_{i.}\beta_h$$

where Σ is the variance-covariance matrix implied by the reduced form of the spatial probit model, $\overline{X}$ is an n by k matrix of regressor means, $\{\cdot\}_{i.}$ is the i-th row of the matrix inside, and $\{\cdot\}_{ii}$ is the i-th diagonal element of a square matrix. Note that Σ depends on ρ or λ if a SAR(1)-probit model or a SAE(1)-probit model is considered, respectively.

The first specification of the above equations explains the impact of a marginal change in the mean of the h-th regressor, i.e., $\overline{x}_{.h}$, on the conditional probability of $\{y_i = 1\}$, i.e., $\mathrm{P}(y_i = 1 \mid X_n)$, setting $\overline{x}_{.h'}$ for all the remaining regressors, $h' = 1, \ldots, k-1$. The second specification of the above equations considers instead the marginal impact evaluated at each single value of $x_{.h}$. This is particularly informative in space in terms of *spatial heterogeneity* due to the possibility of evaluating a marginal impact with respect to a particular region value x_{ih}. The results are two n-dimensional square matrices for $\{y_1, y_2, \ldots, y_n\}$. Finally, the average of the main diagonals provides a synthetic measure of the *direct impact*, while the average of the off-diagonal elements provide a synthetic measure of the *indirect impact*.

4. Fitting Spatial Econometric Nonlinear Models in R

In this section we explain how to fit spatial nonlinear (probit) models in R. There are at least three packages that one can use to estimate a SAR(1)-probit or a SAE(1)-probit model: (i) `McSpatial`, (ii) `ProbitSpatial`, (iii) `spatialprobit`. The first one has inside different functions to estimate a SAR(1)-probit model. The Linearized GMM proposed by Klier and McMillen (2008) and an MLE-based function. However, the linearized GMM approach is accurate as long as the true autocorrelation coefficient is relatively small. In addition, there is no discussion about the asymptotic behaviour of this estimator and no other forms of model specifications are allowed for. A recent variant of this GMM-based estimator is the one proposed by Santos and Proença (2019). The second package has the `SpatialProbitFit` function that fits both the SAR(1)-probit and the SAE(1)-probit model through an Approximate MLE approach, see Martinetti and Geniaux (2017). Although is computationally very fast for large datasets, see also Mozharovskyi and Vogler (2016) for a similar approach, this estimation procedure does not account for more general model specifications and the possible use of *dense* weighting matrices. In addition, there is no discussion about its asymptotic behaviour. Finally, the third one is based on the Bayesian approach, see e.g. J. P. LeSage et al. (2011). Alternative ML-based estimation approaches have been recently proposed by Wang, Iglesias, and Wooldridge (2013), Billé and Leorato (2019). Here, we focus the attention on the use of the `SpatialProbitFit` function in the `ProbitSpatial` package.

A brief description of the Approximate MLE is as follows. First of all, one of the main problems in estimating both linear and nonlinear spatial models, especially with large dataset as it is typically the case in microeconometrics, is related to the repeated calculations of the determinats of n-dimensional matrices, see e.g. Smirnov and Anselin (2001) and Pace and LeSage (2004). In addition, within spatial nonlinear probit models, a n-dimensional integral problem for the estimation of the parameters arises. In view of these features, there are two different approaches that has been proposed by Martinetti and Geniaux (2017): (i) maximization of the full log likelihood function by means of a multi-dimensional optimisation algorithm (ii) maximisation of the log likelihood conditional to ρ. They propose to approximate the multidimensional integral by the product of the univariate conditional probabilites $\Phi_n(x_1 \in A_1, \dots, x_n \in A_n) = \mathcal{P}(x_1 \in A_1) \prod_{i=2}^{n} \mathcal{P}\,(x_i \in A_i \mid \{x_1 \in A_1, \dots, x_{i-1} \in A_{i-1}\})$. Then, by considering the Cholesky decomposition of the variance-covariance matrix $\Sigma_{v(\rho,\lambda)} = CC'$, the interval limits are transformed $S_i = (a_i', b_i')$ by taking advantage of the lower triangular matrix C. The algorithm iteratively substitutes the univariate conditional probabilities with the quantities $\tilde{z}_i = \frac{\phi(a_i') - \phi(b_i')}{\Phi(b_i') - \Phi(a_i')}$ and it ends when the probability of the last random variable is computed and the approximation $\Phi_n(x_1 \in A_1, \dots, x_n \in A_n) \approx \prod_{i=1}^{n} \left(\Phi(b_i') - \Phi(a_i') \right)$ is reached.

The algorithm starts with the computation of $X^* = \frac{(I - \rho W_n)^{-1} X}{\sqrt{(\sum_{i=1}^{n} \Sigma_{ii})}}$ where the matrix $(I - \rho W_n)^{-1}$ is computed by a truncation of its Taylor approximation and a starting value of ρ. Then, the interval limits (a_i, b_i) and $\left(\Phi\left(\frac{b_i}{\sqrt{(\Sigma_{ii})}}\right) - \Phi\left(\frac{a_i}{\sqrt{(\Sigma_{ii})}}\right) \right)$ are computed for each conditional probabilities starting from a random variable. By using the lower triangular matrix of Σ they compute the transformed interval limits $\left(\hat{b}_i, \hat{a}_i\right)$ and the quantities $U_i = \Phi\left(\hat{b}_i\right) - \Phi(\hat{a}_i)$. Finally, the log likelihood function is simply the sum of each contribution $\log(U_i)$.

In the following subsection we implement the above estimation procedure in R to fit spatial probit models and to empirically explain their spatial marginal effects.

4.1. Empirical application in Labour Economics

The analysis of employment/unemployment rates at a more disaggregated geographical level has begun at least 20 years ago, see e.g. Pischke and Velling (1997). Not too long after, several researchers started to introduce in Labour Economics the concept of *socio-economic distance* and to use the spatial and spatio-temporal econometric techniques to deal with interdependences between countries/regions of the employment/unemployment rates, see Conley and Topa (2002), Patacchini and Zenou (2007), Schanne, Wapler, and Weyh (2010), Cueto, Mayor, and Suárez (2015), Halleck Vega and Elhorst (2016), Watson and Deller (2017), and Kosfeld and Dreger (2019), among others.

In this section we analyse a data set that contain several information free available at the Eurostat website `https://ec.europa.eu/eurostat/data/database`. After some data manipulation, the data set consists of 312 observations (European regions at NUTS2 level) referred to 2016 year, with information on employment rates (*employment*, in %), GDP at current market price (*gdp*, measured in Euro per Inhabitant), education level of population aged 25-64 (in %) for three different levels: (i) secondary school or lower (*isced_02*), (ii) high school or lower (*isced_34*), and (iii) degree, post degree, Ph.D. (*isced_58*). The involved countries are the following: Austria (AT), Belgium (BE), Bulgaria (BG), Cyprus (CY), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Greece (EL), Spain (ES), Finland (FI), France (FR), Croatia (HR), Hungary (HU), Ireland (IE), Italy (IT), Lithuania (LT), Luxembourg (LU), Latvia (LV), Montenegro (ME), Macedonia (MK), Malta (MT), Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Romania (RO), Serbia (RS), Sweden (SE), Slovenia (SI), Slovakia (SK), Turkey (TR), United Kingdom (UK). The aim of this section is to study the impact of the above-mentioned explanatory variables on a new binary variable, i.e. the employment status (*bin.emp*), at a regional level in 2016. In a recent study (Gazzola and Mazzacani 2019), the impact of language skills on the individual employment status within some European countries has been analysed through a probit model. In their paper there are also details on the use of the *isced* levels for education.

Once the data from the Eurostat website have been read in R and after some data manipulation, we saved a dataframe with extension `.RData`. The function `load` is able to load the dataset called "`Dataset.RData`" specifying the right directory (path), while the function `nrow` calculates the number of rows of the dataframe that corresponds to the sample size n. The output can be seen by using the function `head`, which shows the first six rows.

```
load("C:/Users/Anna Gloria/Dropbox/Handbook/Data/Dataset.RData")
n <- nrow(data)
n

## [1] 312

head(data)

##   nuts_id country_id employment   gdp isced_02 isced_34 isced_58  long
## 1    AT11         AT       74.1 28600     16.5     55.7     27.9 16.52
## 2    AT12         AT       77.4 33300     14.7     55.1     30.2 15.75
## 3    AT13         AT       69.7 49200     16.9     43.2     39.9 16.39
## 4    AT21         AT       73.6 34200     10.6     60.0     29.4 13.88
## 5    AT22         AT       75.3 36500     14.0     58.6     27.4 14.98
## 6    AT31         AT       79.3 41600     15.7     55.3     29.1 13.95
##     lat
## 1 47.53
## 2 48.24
## 3 48.19
## 4 46.77
## 5 47.26
## 6 48.15
```

The dataset `data` already contains the spatial coordinates, i.e. longitude (*long*) and latitude (*lat*), of the spatial polygons (European regions). Since the spatial observations are polygons/areas, then the spatial coordinates correspond here to their *centroids*. After having installed a package through the code `install.packages("name of the package")`, we can easily upload it through the use of the function `library`. The function `readOGR` in the package `rgdal` is particularly useful to read and transform different types of OGR vector maps (shapefiles) into Spatial objects. Then, the function `coordinates` into the package `sp` is able to calculate the centroids from a Spatial (polygon) object as in the following

```
library(rgdal)
eu                <- readOGR("C:/Users/Anna
Gloria/Dropbox/Handbook/Data/Shapefile/EU_reg.shp")

## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\Anna
Gloria\Dropbox\Handbook\Data\Shapefile\EU_reg.shp", layer: "EU_reg"
## with 312 features
## It has 5 fields
```

```
coords             <- coordinates(eu)
colnames(coords) <- c("long","lat")
head(coords)
```

```
##          long      lat
## 0  4.583496 50.66619
## 1  4.028426 50.44180
## 2  5.744833 50.51087
## 3  5.529503 49.95737
## 4 16.523298 47.53059
## 5 15.747114 48.24467
```

Finally, the function `colnames` provides column names. Figure 2 shows the employment rates (%) in 2016 for all the considered European regions. As we can observe, the majority of the highest values of the employment rates are associated to Northern and Central Europe, especially Germany, UK, Scandinavia.

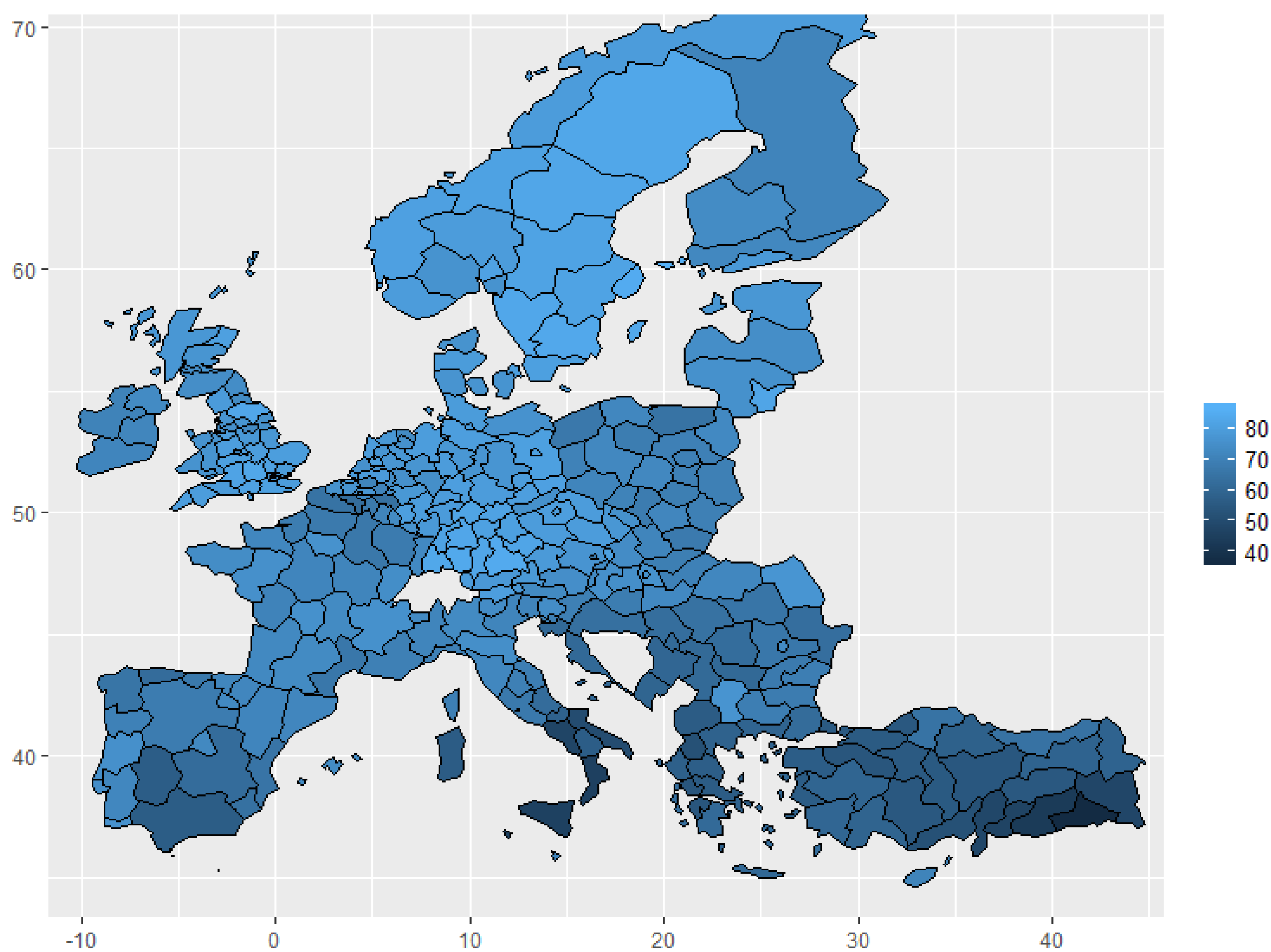

Figure 2 Employment Rates (%) in 2016

Let us now move to the estimation procedure. To estimate a SAR(1)-probit model and a SAE(1)-probit model of section 3, we consider the function `SpatialProbitFit` in the `ProbitSpatial` package. The dependent variable *employment* is a continuous variable, so we need to create a dicothomous dependent variable before estimating the spatial models. To this purpose, we use the mean of the European Area employment rate (EA19) as a threshold, that was equal to 70% in 2016. So, if the employment rate in one European region is greater of equal than the EA19 mean value, then its binary dependent variable takes the value 1, otherwise 0. The definition of the binary variable (*bin.emp*) can be easily obtained by using the function `ifelse` in R.

```
bin.emp <- ifelse(data$employment >= 70, 1, 0)
data    <- data.frame(data,bin.emp)
```

We first regress a standard probit model without the intercept by using the `glm` function and see the output through the function `summary`.

```
fit     <- glm(bin.emp ~ gdp + isced_02 + isced_34 + isced_58 + 0,
               family = binomial(link="probit"),data=data)
summary(fit)

##
## Call:
## glm(formula = bin.emp ~ gdp + isced_02 + isced_34 + isced_58 +
##     0, family = binomial(link = "probit"), data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.7458  -0.1808   0.2046   0.5790   2.4590
##
## Coefficients:
##            Estimate Std. Error z value Pr(>|z|)
## gdp       6.697e-05  1.110e-05   6.036 1.58e-09 ***
## isced_02 -4.707e-02  6.361e-03  -7.401 1.36e-13 ***
## isced_34  2.503e-03  4.219e-03   0.593    0.553
## isced_58  2.832e-03  1.069e-02   0.265    0.791
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
##     Null deviance: 432.52  on 312  degrees of freedom
## Residual deviance: 203.84  on 308  degrees of freedom
## AIC: 211.84
##
## Number of Fisher Scoring iterations: 7
```

Now consider the spatial probit model specifications. In order to fit a spatial model, we need to build our spatial weighting matrix. In this case we exploit the information of the geographical coordinates (centroids) of the spatial polygons, i.e. of the European regions. For this application we use the k-nearest neighbor criterion defined as follow. Let $W = \{w_{ij}\}$ be the spatial weighting matrix with elements equal to the weights among pairs of random variables y_i, y_j for $i, j = 1, \ldots, n$, with n the sample size, then

$$\begin{cases} w_{ij} = 1 & iff \quad y_j \in \mathcal{N}_k \\ w_{ij} = 0 & otherwise \end{cases}$$

where $\mathcal{N}_k$ is the set of nearest random variables y_j to y_i defined by k. Finally, W is row-normalized such that $\sum_j w_{ij} = 1, \forall i$.

To build our weighting matrix W we first load the package `spdep`, see Bivand and Piras (2015). The function `knearneigh` provides a list of class *knn* with the information into the first member of the region number ids to define the nearest neighbours for each random variable (in this case $k = 11$). The `knn2nb` function transforms the object of class *knn* into an object of class *nb* (neighbour list). The argument `sym=T` forces the weighting matrix to be symmmetric. The `nb2mat` function transforms, instead, an object of class *nb* into an n-dimensional weighting matrix. The argument `style = "W"` directly row-normalizes the weights, while the function `as(,"CsparseMatrix")` defines the weighting matrix to be sparse. Finally, the function `dim` provides information about the dimension of the spatial weighting matrix.

```r
library(spdep)
knn11       <- knn2nb(knearneigh(cbind(data$long,data$lat), k=11),sym=T)
Wknn_sparse <- as(nb2mat(knn11,style="W"),"CsparseMatrix")
dim(Wknn_sparse)
```

```
## [1] 312 312
```

We can now load the package `ProbitSpatial` and use the function `SpatialProbitFit` inside. This function is able to fit the SAR(1)-probit model or the SAE(1)-probit (also SEM-probit) model defined in section 3. The argument `DGP=` specifies the type of model, while `W=` and `method=` provide information on the weighting matrix and the estimation method, respectively. For details use the help function in R as `?SpatialProbitFit`.

```r
library(ProbitSpatial)
fit.sp1 <- SpatialProbitFit(bin.emp ~ gdp + isced_02 + isced_34 + isced_58 + 0,
                             data=data, W=Wknn_sparse, DGP='SAR', method="full-
lik")
```

```
## St. dev. of beta conditional on rho and Lik-ratio of rho
##                 Estimate   Std. Error     z-value     Pr(>|z|)
## gdp         3.754825e-05 6.839877e-06  5.4896086 4.028254e-08
## isced_02   -3.114469e-02 4.231689e-03 -7.3598718 1.840750e-13
## isced_34   -1.523355e-03 2.993939e-03 -0.5088129 6.108834e-01
## isced_58    9.102693e-03 7.938271e-03  1.1466846 2.515120e-01
## lambda      5.084473e-01           NA 19.6485634 9.307326e-06
```

```r
fit.sp2 <- SpatialProbitFit(bin.emp ~ gdp + isced_02 + isced_34 + isced_58 + 0,
                             data=data, W=Wknn_sparse, DGP='SEM', method="full-
lik")
```

```
## St. dev. of beta conditional on rho and Lik-ratio of rho
##                 Estimate   Std. Error     z-value     Pr(>|z|)
## gdp         7.476664e-05 1.414341e-05  5.2863230 1.247994e-07
## isced_02   -5.220912e-02 7.059491e-03 -7.3955937 1.407763e-13
## isced_34    1.445296e-02 4.171657e-03  3.4645615 5.310968e-04
## isced_58   -9.685959e-03 1.192018e-02 -0.8125681 4.164657e-01
## rho         6.372373e-01           NA 13.5257220 2.353158e-04
```

Table 1 shows the estimation results of the standard probit model, the SAR(1)-probit model and the SAE(1)-probit model, respectively. Regardless of the specific European region, the variables *gdp* and *isced_02* are statistically significant at $\alpha = 0.001$ significance level for all the three model specifications. This reflects on an higher probability of being employed for higher values of *gdp* due to the positive sign of its coefficient and a lower probability of being employed for higher percentage values of people with a secondary school education level at most due to its negative coefficient.

Table 1 Estimation Results

	Standard probit	SAR(1)-probit	SAE(1)-probit
gdp	0.00007	0.00004	0.00007
s.e.	0.00001	0.00001	0.00001
isced_02	-0.04707	-0.03114	-0.05221
s.e.	0.00636	0.00423	0.00704
isced_34	0.00250	-0.00152	0.01445
s.e.	0.00422	0.00299	0.00415
isced_58	0.00283	0.00910	-0.00969
s.e.	0.01069	0.00794	0.01186
rho	NA	0.50845	NA
s.e.	NA	NA	NA
lambda	NA	NA	0.63724
s.e.	NA	NA	NA
loglik	-101.92212	92.09784	95.15926

Specifically for the spatial models, both the estimated autoregressive coefficients $\left(\hat{\rho}, \hat{\lambda}\right)$ are significant. Their estimates are quite high with respect to the upper bound 1, highlighting an important role played by spatial dependence. Moreover, due to their positive signs, both the autoregressive processes in the dependent variables and in the error terms are not inhibitory, leading to a positive interconnection (clustering process) between units over space. From an economic point of view, the higher is the probability of being employed in one region the more is the probability of being employed in its neighbourhood, and viceversa. Then, these effects propagate with a decreasing magnitude through all the other regions.

Let now consider the marginal effects. As explained in subsection 3.1, one way to calculate the marginal effects in probit models consists in evaluating the average of the individual (local) marginal effects with respect to each regressor of interest. In a spatial context, these local effects depend on the estimated coefficients $\left(\hat{\beta}, \hat{\rho}, \hat{\lambda}\right)$, according to the type of model specification. The result with respect to each regressor is a square matrix of impacts, where the diagonal elements are the *direct effects* whereas the off-diagonal elements are the *indirect effects*. The average of these elements provide a summary measure of the impact due directly to the regressor and a summary measure due to the neighbouring dependent variables, respectively. In the following we can observe these average effects for the SAR(1)-probit model with respect to each regressors.

```
m.effects.sar <- rbind(av.dir.eff,av.ind.eff,av.tot.eff)
rownames(m.effects.sar) <- c("direct","indirect","total")
colnames(m.effects.sar) <- c("gdp","isced_02","isced_34","isced_58")
m.effects.sar

##                   gdp     isced_02      isced_34    isced_58
## direct   5.836276e-06 -0.004840945 -0.0002367812 0.001414868
## indirect 5.699937e-06 -0.004727858 -0.0002312499 0.001381816
## total    1.153621e-05 -0.009568804 -0.0004680311 0.002796685
```

There is a clear balance between average direct and average indirect effects for all the considered regressors, while the greatest total impact in absolute value is due to *isced_02*. Note that, although some regressor coefficients are not significant, like *isced_34* and *isced_58* in this case, the average direct, indirect and total impacts should be considered since that the autoregressive coefficent ρ is instead significant and directly affects all the marginal effects.

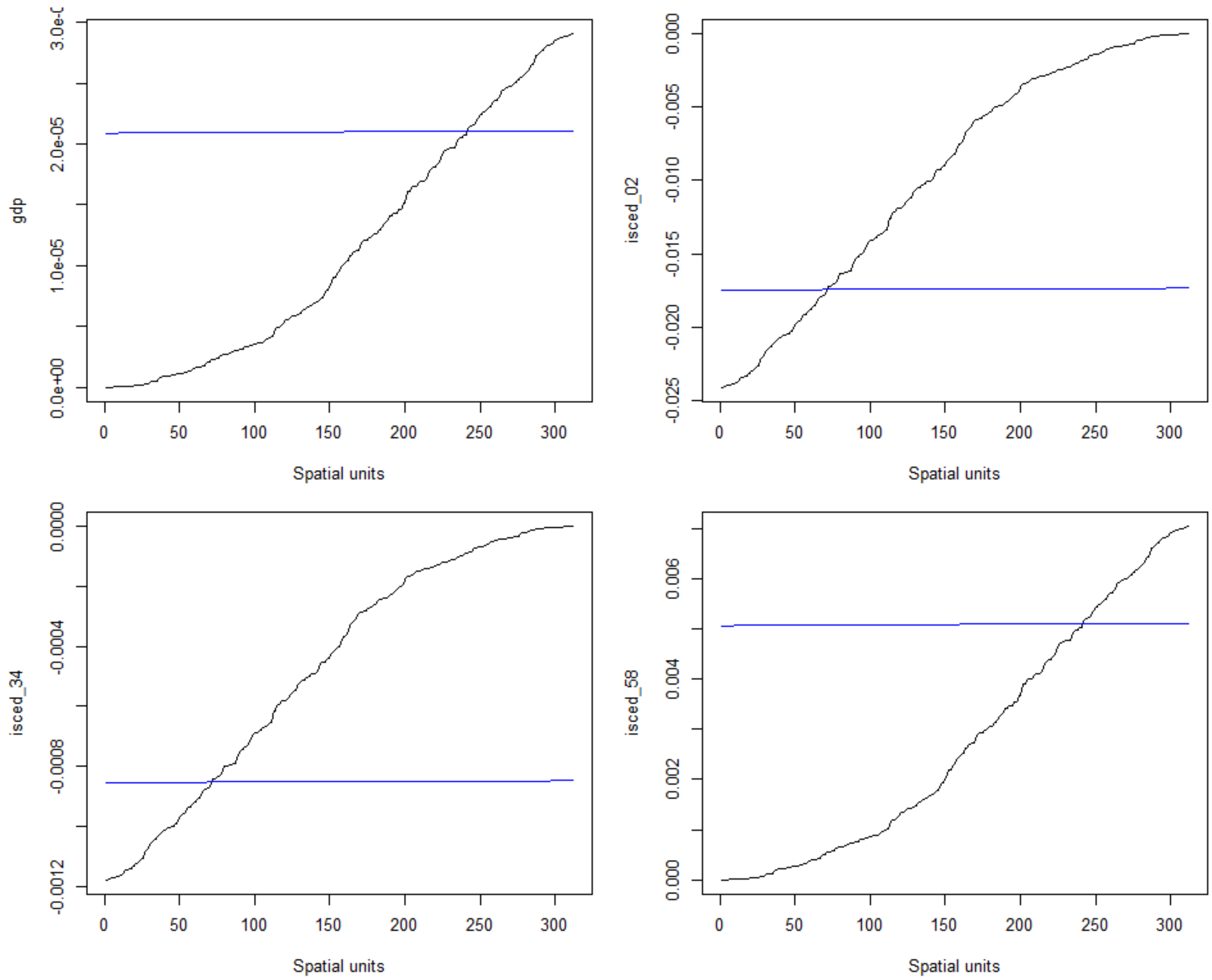

Figure 3 Local Marginal Effects from the SAR(1)-probit model

The same occurs for the SAE(1)-probit specification, where in this case the impacts due to the spillover effects seem to be more pronounced.

```
m.effects.sae <- rbind(av.dir.eff,av.ind.eff,av.tot.eff)
rownames(m.effects.sae) <- c("direct","indirect","total")
colnames(m.effects.sae) <- c("gdp","isced_02","isced_34","isced_58")
m.effects.sae
```

```
##                    gdp     isced_02     isced_34      isced_58
## direct   3.985610e-06 -0.002783129 0.0007704489 -0.0005163326
## indirect 6.395970e-06 -0.004466270 0.0012363899 -0.0008285928
## total    1.038158e-05 -0.007249399 0.0020068387 -0.0013449255
```

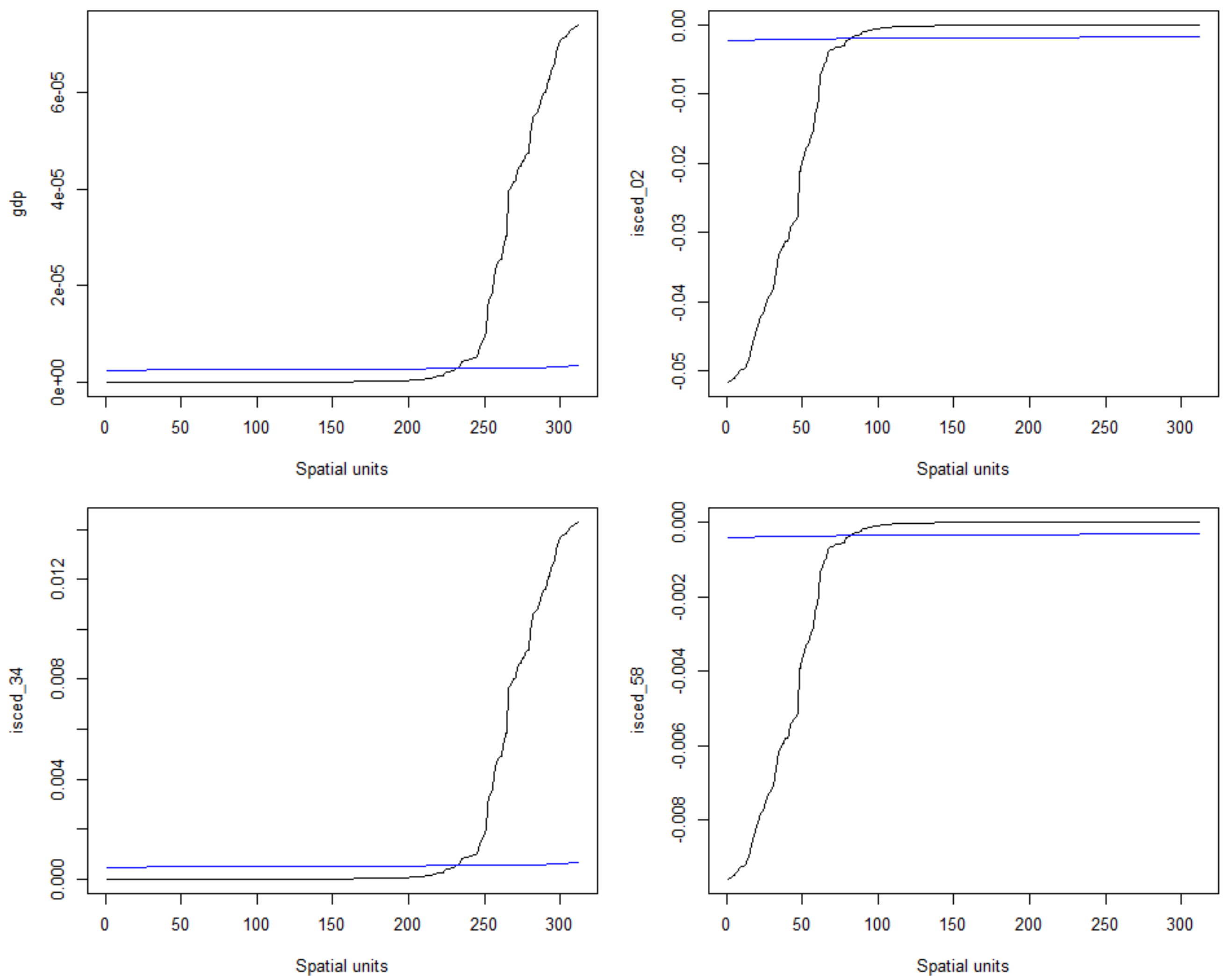

Figure 4 Local Marginal Effects from the SAE(1)-probit model

Figures 3 and 4 show the local (total) marginal effects from a SAR(1)-probit model and a SAE(1)-probit model, respectively, sorted from low to high values. In this way, we are able to evaluate the range of variation of the total marginal effects for different regions and compare these values with the average one (blue line). For instance, it is interesting to note that in Figure 4 the distributions of the marginal effects are more asymmetric than the ones in Figure 3.

Figures 5 and 6 show, instead, the local marginal effects from a SAR(1)-probit model and a SAE(1)-probit model, respectively, on the different European regions. The point in this case is to geographically identify regions with greater or lower marginal impacts. As we can observe from Figure 5, the positive total impact of the *gdp* on employment rates reach its lowest values in the regions within Scandinavia, UK, Germany and Turkey. Although the magnitude is different according to different regressors, the same occurs for the variable *isced_58*. On the contrary, quite the opposite can be found for the variable *isced_02* and *isced_34*, because of the negative sign in these cases. A plausible conclusion might be that both the direct effects from the regressors and the indirect effects due to neighboring regions have generally a lower impact in Scandinavia, UK, Germany and Turkey, with respect to the other European countries. It is finally worth noting that, one can be interested in evaluating only the direct or the indirect effects, although both of them depend on the spatial dependence coefficient. Careful attention should be paid on the proper interpretation of the marginal effects, since they are the link between estimation results and policy interventions.

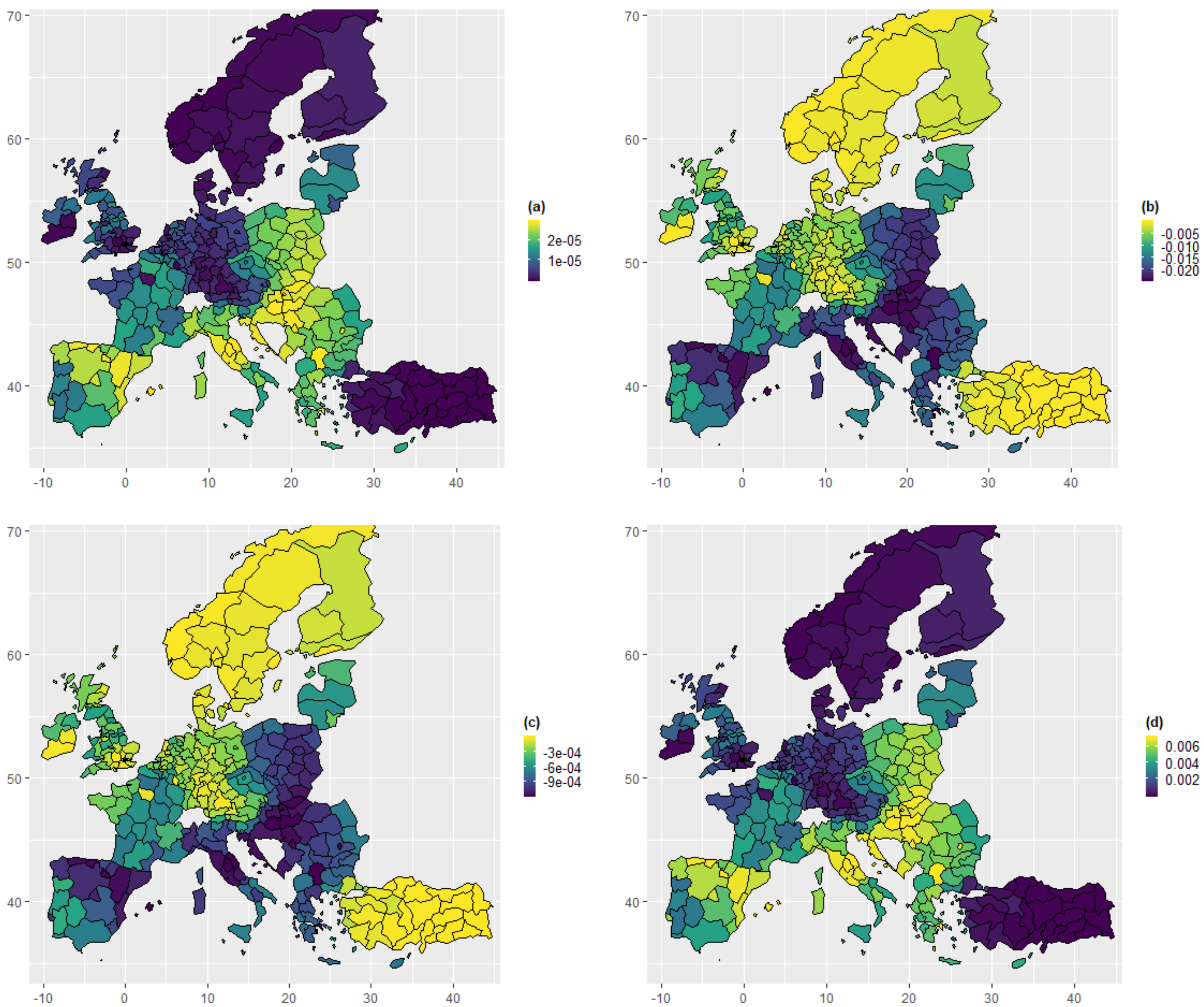

Figure 5 Local Total Marginal Effects from the SAR(1)-probit model in Europe: (a) gdp, (b) isced_02, (c) isced_34, (d) isced_58.

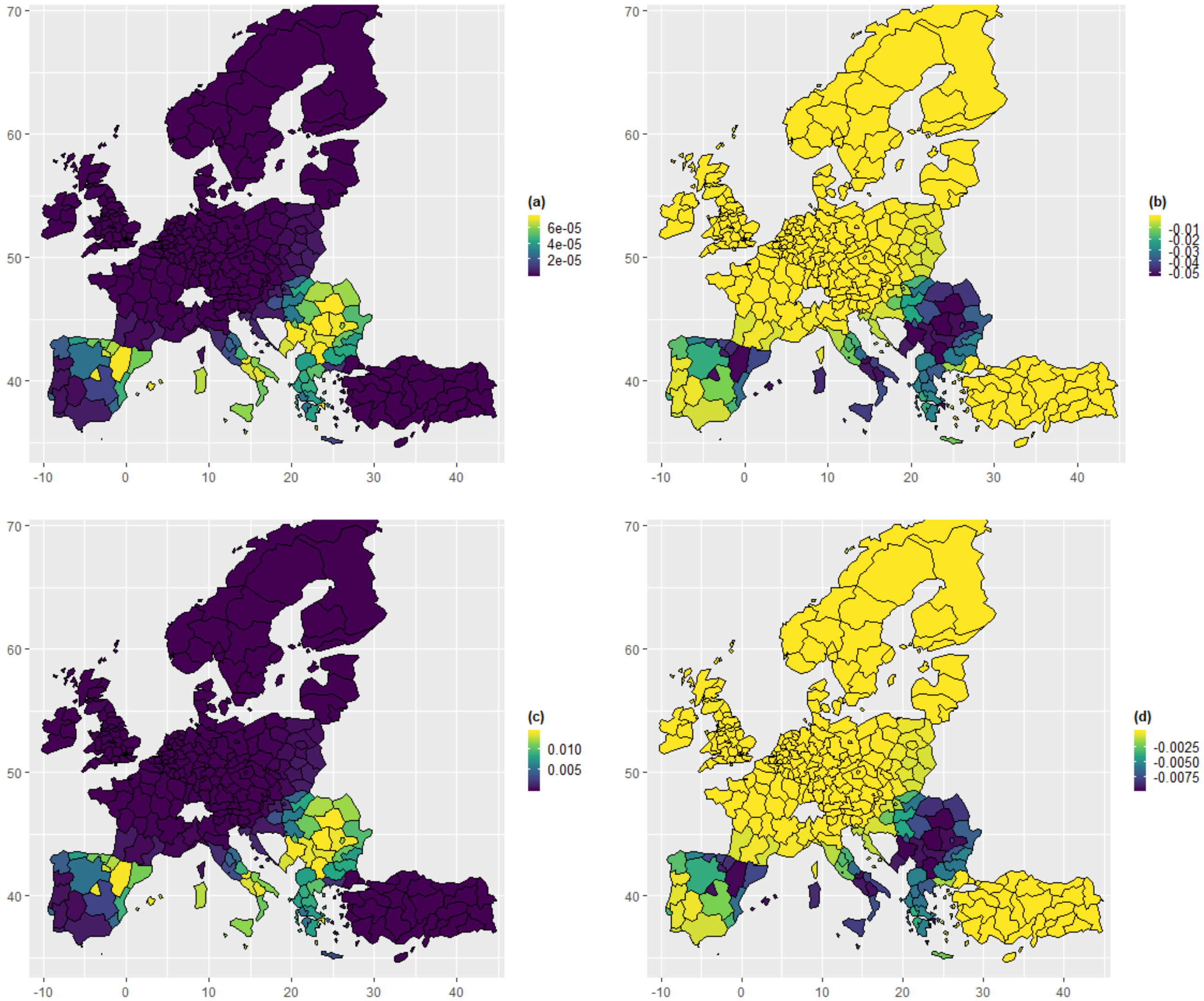

Figure 6 Local Total Marginal Effects from the SAE(1)-probit model in Europe: (a) gdp, (b) isced_02, (c) isced_34, (d) isced_58.

5. References

Ahrens, Achim, and Arnab Bhattacharjee. 2015. "Two-Step Lasso Estimation of the Spatial Weights Matrix." *Econometrics* 3 (1): 128–55.

Amemiya, Takeshi. 1985. *Advanced Econometrics*. Harvard university press.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Vol. 4. Springer Science & Business Media.

Anselin, Luc, and Anil K Bera. 1998. "Introduction to Spatial Econometrics." *Handbook of Applied Economic Statistics* 237.

Arbia, Giuseppe. 2014. *A Primer for Spatial Econometrics: With Applications in R*. Springer.

Arbia, Giuseppe, Marco Bee, and Giuseppe Espa. 2013. "Testing Isotropy in Spatial Econometric Models." *Spatial Economic Analysis* 8 (3): 228–40.

Arbia, Giuseppe, Giuseppe Espa, and Diego Giuliani. 2018. *Spatial Microeconometrics*. Taylor & Francis.

Bhattacharjee, Arnab, and Chris Jensen-Butler. 2013. "Estimation of the Spatial Weights Matrix Under Structural Constraints." *Regional Science and Urban Economics* 43 (4): 617–34.

Billé, Anna Gloria. 2013. "Computational Issues in the Estimation of the Spatial Probit Model: A Comparison of Various Estimators." *Review of Regional Studies* 43 (2, 3): 131–54.

Billé, Anna Gloria, and Giuseppe Arbia. 2019. "Spatial Limited Dependent Variable Models: A Review Focused on Specification, Estimation and Health Economics Applications." *Journal of Economic Surveys* 33 (5): 1531–1554.

Billé, Anna Gloria, and Leopoldo Catania. 2018. "Dynamic Spatial Autoregressive Models with Time-Varying Spatial Weighting Matrices." *Available at SSRN 3241470*.

Billé, Anna Gloria, and Samantha Leorato. 2019. "Partial ML estimation for spatial autoregressive nonlinear probit models with autoregressive disturbances." *Econometric Reviews* 39 (5): 1–39.

Bivand, Roger, and Gianfranco Piras. 2015. "Comparing Implementations of Estimation Methods for Spatial Econometrics." In. American Statistical Association.

Cameron, A Colin, and Pravin K Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge university press.

Case, Anne C, Harvey S Rosen, and James R Hines Jr. 1993. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics* 52 (3): 285–307.

Catania, Leopoldo, and Anna Gloria Billé. 2017. "Dynamic Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances." *Journal of Applied Econometrics* 32 (6): 1178–96.

Conley, Timothy G. 2010. "Spatial Econometrics." In *Microeconometrics*, 303–13. Springer.

Conley, Timothy G, and Giorgio Topa. 2002. "Socio-Economic Distance and Spatial Patterns in Unemployment." *Journal of Applied Econometrics* 17 (4): 303–27.

Corrado, Luisa, and Bernard Fingleton. 2012. "Where Is the Economics in Spatial Econometrics?" *Journal of Regional Science* 52 (2): 210–39.

Cueto, Begoña, Matías Mayor, and Patricia Suárez. 2015. "Entrepreneurship and Unemployment in Spain: A Regional Analysis." *Applied Economics Letters* 22 (15): 1230–5.

Debarsy, Nicolas, and James LeSage. 2018. "Flexible Dependence Modeling Using Convex Combinations of Different Types of Connectivity Structures." *Regional Science and Urban Economics* 69: 48–68.

Elhorst, J Paul. 2014. *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. Vol. 479. Springer.

Elhorst, J Paul, Donald J Lacombe, and Gianfranco Piras. 2012. "On Model Specification and Parameter Space Definitions in Higher Order Spatial Econometric Models." *Regional Science and Urban Economics* 42 (1-2): 211–20.

Fingleton, B, and G Arbia. 2008. "New Spatial Econometric Techniques and Applications in Regional Science." *Papers in Regional Science* 87 (3): 311–17.

Gazzola, Michele, and Daniele Mazzacani. 2019. "Foreign Language Skills and Employment Status of European Natives: Evidence from Germany, Italy and Spain." *Empirica*, September. https://doi.org/10.1007/s10663-019-09460-7.

Getis, Arthur, and Jared Aldstadt. 2004. "Constructing the Spatial Weights Matrix Using a Local Statistic." *Geographical Analysis* 36 (2): 90–104.

Halleck Vega, Solmaria, and J Paul Elhorst. 2016. "A Regional Unemployment Model Simultaneously Accounting for Serial Dynamics, Spatial Dependence and Common Factors." *Regional Science and Urban Economics* 60: 85–95.

Kelejian, Harry H, and Gianfranco 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: 140–49.

———. 2017. *Spatial Econometrics*. Academic Press.

Kelejian, Harry H, and Ingmar R Prucha. 1998. “A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.” *The Journal of Real Estate Finance and Economics* 17 (1): 99–121.

———. 2010. “Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances.” *Journal of Econometrics* 157 (1): 53–67.

Klier, Thomas, and Daniel P McMillen. 2008. “Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples.” *Journal of Business & Economic Statistics* 26 (4): 460–71.

Kosfeld, Reinhold, and Christian Dreger. 2019. “Towards an East German Wage Curve-Nuts Boundaries, Labour Market Regions and Unemployment Spillovers.” *Regional Science and Urban Economics* 76: 115–24.

Le Gallo, Julie, and Bernard Fingleton. 2012. “Measurement Errors in a Spatial Context.” *Regional Science and Urban Economics* 42 (1-2): 114–25.

Lee, Lung-Fei. 2004. “Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models.” *Econometrica* 72 (6): 1899–1925.

LeSage, James P. 2014. “What Regional Scientists Need to Know About Spatial Econometrics.” *The Review of Regional Studies* 44 (1): 13.

LeSage, James P, R Kelley Pace, Nina Lam, Richard Campanella, and Xingjian Liu. 2011. “New Orleans Business Recovery in the Aftermath of Hurricane Katrina.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 174 (4): 1007–27.

LeSage, James, and R Pace. 2014. “The Biggest Myth in Spatial Econometrics.” *Econometrics* 2 (4): 217–49.

LeSage, JP, and R Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman &Hall/CRC.

Maddala, Gangadharrao S. 1986. *Limited-Dependent and Qualitative Variables in Econometrics*. 3. Cambridge university press.

Martinetti, Davide, and Ghislain Geniaux. 2017. “Approximate Likelihood Estimation of Spatial Probit Models.” *Regional Science and Urban Economics* 64: 30–45.

McFadden, Daniel L. 1984. “Econometric Analysis of Qualitative Response Models.” In *Handbook of Econometrics*, 2:1395–1457. Elsevier.

Mozharovskyi, Pavlo, and Jan Vogler. 2016. “Composite Marginal Likelihood Estimation of Spatial Autoregressive Probit Models Feasible in Very Large Samples.” *Economics Letters* 148: 87–90.

Otto, Philipp, and Rick Steinert. 2018. “Estimation of the Spatial Weighting Matrix for Spatiotemporal Data Under the Presence of Structural Breaks.” *arXiv Preprint arXiv:1810.06940*.

Pace, R Kelley, and Ronald Barry. 1997. “Sparse Spatial Autoregressions.” *Statistics & Probability Letters* 33 (3): 291–97.

Pace, R Kelley, and James P LeSage. 2004. “Chebyshev Approximation of Log-Determinants of Spatial Weight Matrices.” *Computational Statistics & Data Analysis* 45 (2): 179–96.

Patacchini, Eleonora, and Yves Zenou. 2007. “Spatial Dependence in Local Unemployment Rates.” *Journal of Economic Geography* 7 (2): 169–91.

Pinkse, Joris, and Margaret E Slade. 2010. "The Future of Spatial Econometrics." *Journal of Regional Science* 50 (1): 103–17.

Pinkse, Joris, Margaret E Slade, and Craig Brett. 2002. "Spatial Price Competition: A Semiparametric Approach." *Econometrica* 70 (3): 1111–53.

Pischke, Jörn-Steffen, and Johannes Velling. 1997. "Employment Effects of Immigration to Germany: An Analysis Based on Local Labor Markets." *Review of Economics and Statistics* 79 (4): 594–604.

Qu, Xi, and Lung-fei Lee. 2015. "Estimating a Spatial Autoregressive Model with an Endogenous Spatial Weight Matrix." *Journal of Econometrics* 184 (2): 209–32.

Santos, Luís Silveira, and Isabel Proença. 2019. "The Inversion of the Spatial Lag Operator in Binary Choice Models: Fast Computation and a Closed Formula Approximation." *Regional Science and Urban Economics* 76: 74–102.

Schanne, Norbert, Rüdiger Wapler, and Antje Weyh. 2010. "Regional Unemployment Forecasts with Spatial Interdependencies." *International Journal of Forecasting* 26 (4): 908–26.

Smirnov, Oleg, and Luc Anselin. 2001. "Fast Maximum Likelihood Estimation of Very Large Spatial Autoregressive Models: A Characteristic Polynomial Approach." *Computational Statistics & Data Analysis* 35 (3): 301–19.

Wang, Honglin, Emma M Iglesias, and Jeffrey M Wooldridge. 2013. "Partial Maximum Likelihood Estimation of Spatial Probit Models." *Journal of Econometrics* 172 (1): 77–89.

Watson, Philip, and Steven Deller. 2017. "Economic Diversity, Unemployment and the Great Recession." *The Quarterly Review of Economics and Finance* 64: 1–11.