LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

Master in

Applied Econometrics and Forecasting

Dissertation

Estimating a Knowledge Production Function and Knowledge Spillovers: A new two-step estimation procedure of a Spatial Autoregressive Poisson Model

Author:

Ludgero Miguel Carraça Glórias

Advisor:

Profª Doutora Isabel Proença

November 2020

To Mrs. Maria de Lurdes Alenquer dos Santos Carraça, because everything that I was, am and will be, I owe to her

"The human race needs an intellectual challenge. It must be boring to be God, and have nothing to discover"

Stephen Hawking

Abstract

Several econometric studies seek to explain the determinants of knowledge production using as dependent variable the number of patents in a given region. Some of these studies intend to capture the effects of knowledge spillovers through linear models with spatial autorregressive term. However, no study has been found that estimates the effect of such term while also considering the discrete nature of the dependent variable, which is a count variable.

This essay aims to fill this gap by proposing a new Two-step Limited Information Maximum Likelihood estimator for a Spatial Autorregressive Poisson model. The properties of this estimator are evaluated in a set of Monte Carlo Experiments. The simulation results suggest that, in general, this estimator presents lower Bias and lower RMSE than the alternative estimators proposed, only showing worse results when the spatial dependence is very close to the unit. An empirical example, using the new estimator and a set of alternative estimators for comparison, is executed, where the creation of knowledge in 234 NUTS II from 24 European countries is analyzed. The results show that there is a strong spatial dependence on the creation of innovation between regions. It is also concluded that the socio-economic environment is essential for the knowledge formation process and that, unlike public R&D institutions, private companies are efficient in producing innovation. It should also be noted that regions with less capacity to transform R&D expenses into new patents, have greater capacity for absorption and segregation of knowledge, which may show that neighboring regions less efficient in the production of knowledge tend to create strong relations with each other taking advantage of the knowledge sharing process.

Keywords: Spatial Econometrics; Poisson Regression; Knowledge Spillovers; Knowledge Production; Two-Step Limited Information Maximum Likelihood

JEL classifications : C21 C25 O31 R12

Resumo

Vários estudos econométricos procuram explicar os determinantes da criação de conhecimento usando como variável dependente o número de patenteamentos numa determinada região. Alguns destes estudos procuram captar os efeitos de *Knowledge Spillovers* através de modelos lineares que incorporam dependência espacial. No entanto, nenhum estudo foi encontrado que captasse este efeito, tendo ao mesmo tempo em atenção a natureza discreta da variável dependente, que neste caso é uma variável de contagem. Este artigo pretende preencher essa lacuna propondo um novo estimador de máxima verosimilhança de informação limitada a dois passos para um modelo Poisson Autorregressivo Espacial. As propriedades do estimador são avaliadas num conjunto de simulações de Monte Carlo. Os resultados da simulação sugerem que este estimador tem menor Bias e menor RMSE, na generalidade, que outros estimadores anteriormente propostos, sendo que apenas mostra piores resultados quando a dependência espacial é muito próxima da unidade. Um exemplo empírico, empregando o novo estimador e um conjunto de estimadores alternativos para comparação, é realizado, sendo que a criação de conhecimento em 234 NUTS II de 24 países europeus é analisada. Os resultados evidenciam que existe uma forte dependência espacial na criação de inovação entre as regiões. Conclui-se também que o ambiente socioeconómico é essencial para o processo de formação de conhecimento e que contrariamente às instituições públicas de R&D, as empresas privadas são eficientes na produção de inovação. É de realçar ainda, que regiões com menor capacidade em transformar despesas R&D em novas patentes apresentam maior capacidade de absorção e segregação de conhecimento, podendo evidenciar que, regiões vizinhas menos eficientes na produção de conhecimento tendem a criar relações fortalecidas relativamente à partilha de conhecimento.

Palavras-Chave: Econometria espacial; Regressão de Poisson; Externalidades de conhecimento; Máxima Verosimilhança de informação limitada a dois passos.

Classificação JEL: C21 C25 O31 R12

Acknowledgments

First, I would like to thank my Advisor, Professora Doutora Isabel Proença, for all the valuable suggestions and support throughout this journey.

Secondly, thank my parents and grandmother for the emotional and financial support during my academic life.

Also thank Beatriz Carvalho Alvadia for all the love, care and attention given throughout the execution of this dissertation, without which I would have been unable to finish this process.

I am also grateful to Professor Doutor Luís Silveira Santos for providing part of the R Script used in the execution of this dissertation, as well as some quite useful suggestions.

A special thanks to all the teachers I came across during my seventeen years as a student, once all contributed in part to this project.

Finally, thanks to all the friends and family who were at my side during the realization of this project.

Thank you all, once again.

Ludgero Glórias, 2020

Table of content

List of Tables and Figures

Glossary

APE – Averaged Partial Effects

DPE – Direct Partial Effect

EID – Euclidean Inverse Distance

FIML – Full Information Maximum Likelihood

GMM – Generalized Method of Moments

IPE – Indirect Partial Effect

IHS – Inverse Hyperbolic Sine

KPF – Knowledge Production Function

LIML – Limited Information Maximum Likelihood

LISA – Local Indicator of Spatial Association

ML – Maximum Likelihood

OLS – Ordinary Least Squares

P.P – Percentage Points

R&D – Research and Development

RMSE – Root-Mean-Square Error

SAR – Spatial Autoregressive

1. Introduction

Since the rise of the modern economy, economists have been focusing on competitiveness has a preponderant factor for the prosperity of regional, national and international markets. As such, understanding the determinants of competitiveness became a priority for economic and governmental decision makers. At the present moment, one of the predominant variables in competitiveness is the capacity for innovation, and Fritsch (2002) states that the production of knowledge is quite useful to compare the quality of regional innovation systems, being, therefore, a key variable in today's economy. OCDE (1999) states “Technological change and innovation are among the main determinants of productivity growth. Productivity is the key to increasing real income and competitiveness and is one of the most important yardsticks of industrial performance.”

As such, understanding the process of innovation is necessary to current political and economic decision-making. Hence, a large theoretical and empirical literature associated with the theme can be found. The vast majority tries to study the innovation process empirically through the number of new patents in a given region (Buesa *et al.*, 2010). Now, part of this literature proves the existence of externalities associated with the creation of knowledge, commonly known by Knowledge Spillovers. In an attempt to capture these externalities quantitatively, spatial econometrics mechanisms have been increasingly used. However, among this vast literature, there are few empirical studies that, in addition to using spatial econometrics, also pay attention to the discrete nature of the dependent variable. The reason for the scarce literature is the little exploration of spatial autoregressive models of counts.

One of the aims of this essay is to estimate a knowledge production function using spatial econometrics methodologies in order to capture the effects of Knowledge Spillovers in European countries. Given that the studied dependent variable is a count variable (number of new patents) it is decided to use a non-linear estimation process, in this case a Poisson regression. This leads to the second objective of this essay: introducing a new estimation process for the Spatial Autorregressive Poisson Model (SAR-Poisson) presented by Lambert *et al.* (2010). This new methodology aims to eliminate the bias generated in the estimation proposed by Lambert *et al.* (2010), by proposing a first-step Poisson Maximum Likelihood approach where, in the estimation of the the logarithm of the dependent variable, no computational transformation is needed to deal with the

possible problem of zero counts, nor is it necessary to resort to an estimation using a loglinear specification. This is a relevant innovation since it takes into account important results addressed by Santos Silva & Tenreyro (2006), which, when neglected, can cause biased estimates. In addition, if the results of Monte Carlo simulations are satisfactory, then this new methodology will be an important contribution to the estimation of count models with spatial dependence, given the still scarce literature related to the topic.

In the following section, a brief literature survey will be carried out on Knowledge Production Functions and the determinants of innovation, ending with a short summary of the existing spatial models of counts, and some of the problems associated with the estimation. In section 3, the SAR-Poisson and the relevant partial effects to be estimated are presented, followed by a detailed exposition of the new estimation process proposed in this essay. Section 4 presents the main results of a Monte Carlo simulation study, where the proposed new estimator is compared with three other estimators used to estimate count models, in an attempt to prove the benefits of using the first one. It is hypothesized that the new estimator is less biased resulting in more accurate estimates. In section 5, an application of the new estimator for Poisson models with spatial dependence is presented, aiming to estimate the impact of various socio-economic variables in the creation of knowledge, as well as quantifying the mechanisms of Knowledge Spillovers. In addition to the proposed estimator, and as a form of comparison, the same model is estimated with the other three estimators previously referenced. Section 6 elaborates the summary of the main results, followed by some concluding remarks, ending with a discussion on some possible extensions of this essay.

2. Literature Survey

The present literature survey will be divided into seven sub-chapters. The first five try to familiarize the reader with the relevant literature on the specification of the models and determinants that seek to explain the creation of knowledge through innovation, and how this can flow through space. The remaining explore the different estimation approaches used for spatial models, with particular attention to nonlinear spatial models.

2.1 Knowledge Production Function

Griliches (1979), in an attempt to model the knowledge production process, proposed a specification based on a Cobb-Douglas function. It is known as *Knowledge Production Function* (KPF) and describes the relation between knowledge creation and underlying factors, such as human resources, technology, and capital. This methodology has been widely reproduced, with Furman *et al.* (2002), Furman & Hayes (2004), Krammer (2009) and Buesa *et al.* (2010) being some examples. The theoretical function would be

$$Y = DC^{\alpha}L^{\beta}K^{\gamma}e^{(\lambda t+u)} \quad (1)$$

Where Y represents the output of the production function; *D* is a constant; C and L are the conventional inputs capital and labor, respectively; K is a measure of the current state of technical knowledge, measured by the R&D expenditures; t is the time index; u stands for all the unmeasured determinants of the knowledge production; e is the exponential function, and α, β, γ, and λ are the coefficients aimed to estimate.

2.2 Patents as a measure of Knowledge

There is a wide debate about which is the best variable to measure knowledge creation, as can be read in Smith (2005), European Commission (2001: 38), among others.

As stated earlier, knowledge creation is intertwined with the idea of innovation, therefore, it is common to use the number of new patents registered as the knowledge production proxy. However, some studies, such as Acs & Audretsch (1988), have calculated that only between 49% to 60% of patents actually become a product and consequently an innovation, or Arundel & Kabla (1998), who detected strong variations between industrial sectors in the percentage of patents that in fact become innovation, with this occurring, on average, 33% in the case of products, and only 20% in the case of services. Another associated problem is the inability to quantify the heterogeneity between each patent in the production of knowledge (Kleinknecht *et al.*, 2002). It is important to note that not all knowledge production is reflected in the form of a patent (e.g: Scientific articles), which is considered another big disadvantage of using this proxy. Finally, and perhaps the biggest disadvantage for studies that take international data into account, as the present, is that there are different propensities for patenting in different countries and different sectors, and this fact must be taken into account in the analysis of results (Buesa *et al.*, 2010). Other measures, such as the case of innovations sold or considering a variable formed by the sum of patents with innovation (Ferreira & Godinho, 2015) are also proposed as the proxy of knowledge creation, nevertheless, the difficulty

in finding data is seen as a major disadvantage (Buesa *et al.*, 2010). Despite all these disadvantages, patents are considered the best proxy found, since they guarantee a minimum level of originality, and are more likely to become innovations rather than the alternatives. On the other hand, OCDE (2004: 136) has ensured that the vast majority of inventions have been patented in recent decades. Finally, patents have the great advantage of being granted to the regions where they were developed, facilitating studies like this one (Buesa *et al.*, 2010).

2.3 R&D as a source of Knowledge Production

Mansfield (1965) was a pioneer in estimating the effects of R&D on innovation, and since then, the use of variables related to R&D have become almost mandatory when modeling knowledge production. Krammer (2009) concludes that in the innovation process, employment in R&D is crucial. Romer (1990) states that inputs in R&D constitute the most important variable in the creation of knowledge, since an increase of this factor will accelerate the stock of knowledge, promoting productivity and technological progress.

However, it is necessary to take into account that there are different institutions that invest in R&D. These have different objectives and research channels, for example, universities and research institutions can focus on a theoretical component that can later be a channel for innovation in companies (Jiao & Chen 2018). Given this condition, several studies preferred to divide both R&D expenditures and investment in R&D human capital between different sectors: Private, Public and University, such as Krammer (2009), Ferreira & Godinho (2015) and Zhang *et al.* (2020). The conclusion is that the expenses in the private sector are quite significant, in contrast to the university and public sector, probably arising from the inefficiency in the patenting process on the public and universities behalf (Zhang *et al.*, 2020). Ferreira & Godinho (2015) did not find robust results in relation to the university component, still, they were able to conclude that it is always less significant than investment in the private sector.

2.4 Regional determinants in Knowledge Production

Besides economic determinants, it is important to take into account variables connected with the regional environment. Studies such as Ferreira & Godinho (2015) and Acs *et al.* (2002) emphasize the level of education of the population, as it is expected that a higher level of education will represent a positive impact on efficiency at the time of knowledge production. With a higher level of education, greater scientific literacy and innovation

capacity are expected. Nevertheless, it is important to consider the technological and human capital sophistication in a given region. The production of knowledge depends to a great extent on what Buesa *et al* .(2010) refers to as the "innovation environment". Only regions with advanced financial and technological means are able to have a favorable environment for the production of knowledge. This is the reason why some authors take technological sophistication into account, as is the case of Furman *et al.* (2002) and Ferreira & Godinho (2015), which use GDP per capita as their proxy. Besides financial and technological advantages, the social conditions of the population are also determining factors for the quality of the "Innovative Environment" mentioned above. Ferreira & Godinho (2015) use the mortality rate for tuberculosis and violent crimes as proxies of the social conditions of the inhabitants of the region, noting that several studies link poverty to tuberculosis.

Other control variables such as population, investment and number of companies are used in several studies (Ferreira & Godinho, 2015)

2.5 Space and Mobility as determinants for Knowledge Creation

As previously mentioned, the diffusion of knowledge is one of the essential forms for innovation and growth (Lucas, 1988 and Romer, 1990). Therefore, research networks seem to be essential points in the dissemination of knowledge. Studies such as Di Cagno *et al.* (2016) and Miguèlez & Moreno (2013) conclude that regions with institutions that participate in the above-mentioned networks tend to present a higher level of innovation. Now, as a result of globalization, whether through new technologies or through personal meetings at innovation fairs or conferences, knowledge flows through space with ease, in what the literature calls Knowledge Spillovers.

Spatial econometrics has been expanding rapidly since the end of the last century, being increasingly considered for studies in applied economics. It is commonly used to capture the effects of Knowledge Spillovers when analyzing regional innovation, which goes hand in hand with the relationship that Marshall (1920) pointed out between innovation and space. Autant-Bernard (2012) points out two major reasons for the use of spatial econometrics when modeling knowledge creation. The first, because of the endogenous growth theory that argues that knowledge is similar to a public good, meaning that a new agent can use the knowledge of another without costs, or with costs lower than those used to produce it. This premise is the basis of the theory of growth and new geography that explains the clustering process and the unique distribution of economic activities, thus

implying spatial dependence. The second, comes for the strong spatial polarization of economic activities. This polarization means the existence of high spatial heterogeneity in knowledge production that should be accounted for. This spatial heterogeneity may be the reason for the existence of spatial dependence within the random error in econometric models (Autant-Bernard, 2012).

Anselin *et al.* (1997) were the pioneers in this theme, using regional R&D levels in conjunction with neighboring regions R&D levels, in an attempt to measure knowledge creation. However, it is important to bear in mind that, when considering interregional spillovers, assuming spatial dependence only on random disturbances can be misleading, therefore, justifying the importance of adding the spatial autoregressive term (Anselin *et al.*, 1997; Maggioni *et al.*, 2007).

The importance of incorporating spatial autoregressive term in modeling the knowledge creation is stressed in Autant-Bernard & LeSage (2011). First, the authors estimate an aspatial knowledge production model. Nevertheless, the presence of unobservable regional inputs in knowledge production leads them to estimate a Spatial Durbin Model, in which the spatial dependence is captured by the estimated spatial autoregressive coefficients related to both spatially lagged dependent and explanatory variables.

Autant-Bernard (2012) also exposes the benefits of introducing in the regression the spatial autoregressive term in this type of problem. Firstly, it is possible to capture the direct and indirect effects of an explanatory variable distinctly. The increases in knowledge derived from the variation of inputs in the region itself are called direct effects, while the impact in that same region caused by a variation of input in neighboring regions is formally known as indirect effects, the latter being in this case called Knowledge Spillovers. Secondly, with this type of spatial dependence it is possible to study the extent of space on knowledge spillovers, and how its proliferation decays with distance. Finally, it allows the adequate estimation of the model coefficients, since by neglecting spatial dependence we are estimating models with endogeneity, generating biased and inconsistent estimators (Anselin & Le Gallo, 2006).

A variety of studies have been replicated modelling the spatial dependence using a spatial autoregressive term, some examples are Furková (2019), Autant-Bernard & LeSage (2011), Zhang *et al.* (2020) and Caragliu & Nijkamp (2016). All conclude that there is a strong spatial dependence when modeling innovation, and that regional innovation has a spatial spillover effect, both at the level of the dependent variable and the independent variables themselves. Caragliu & Nijkamp (2016) stretched to the point

of not considering only distance as a proximity factor, also opting to take into account relational, social, cognitive and technological proximity.

2.6 Spatial Model for counts: why the non-linear approach.

Until now, the referenced studies which estimate the spatial dependence (Furková, 2019; Autant-Bernard & LeSage, 2011; Zhang *et al.*, 2020), have not taken into account the discrete nature of the dependent variable, which in this case is the number of patents in a given region in a given period. Therefore, we are in the presence of a discrete non-negative variable, that is, a count variable. For a more accurate estimation of the model, it is crucial to take into account these characteristics of the variable. The distribution of this type of variables is skewed to the right due to the high number of zeros and / or small values. Data of this nature is intrinsically heteroscedastic with the variance growing with the average. This last aspect leads to invalid inference, therefore, when estimating standard errors, heteroscedasticity must be considered. (Cameron & Trivedi, 2005).

However, there are a few studies that mutually incorporate the discrete nature of the data and take into consideration the spatial dependence. The still recent exploration of the spatial autoregressive model of counts and the additional complexity arising from the model estimation, largely due to the endogeneity caused by the spatial autocorrelation factor, are considered the main reasons for this scarce literature. In fact, the only study found does not directly estimate the spatial autoregressive term referred earlier. LeSage *et al.* (2007) use a Bayesian Hierarchical Poisson Spatial interaction model to measure the effects of interregional flows of knowledge, using Markov Chain Monte Carlo (MCMC) methods for the estimation.

One way to estimate count models taking into account spatial dependence is to use a loglinear model, making it possible to use the standard approach for estimating spatial linear models. However, considering the work of Santos Silva & Tenreyro (2006) a problem arises. According to the authors, when estimating a log-linearized specification with OLS it will generate biased estimators for elasticity in the presence of heteroscedasticity. For that reason, interpreting the estimated parameters of log-linearized models as elasticities can be improperly in this situation. This is due to Jensen's inequality that implies that the expected value of the logarithm of a random variable is different from the logarithm of its expected value ($E(\ln y) \neq \ln E(y)$). Alternatively, the authors suggest that constant-elasticity models should be estimated in their multiplicative form

using a Poisson Pseudo-Maximum-Likelihood (PPML) estimation technique. To guarantee the consistency of this estimator it is only necessary the correct specification of the conditional mean. Consequently, the dependent variable does not have to be Poisson, nor even to be an integer. Besides all, the implementation of this estimator is straightforward.

2.7 Spatial model for counts: Existing applications

Some authors addressed the estimation of count models with spatial dependence, as was the case of Kaiser & Cressie (1997), that presented a model that allows positive dependence, specifying the conditional distribution as a function of a probabilistic mass Winsorized Poisson; Schabenberger & Pierce (2002) analyse conditional autoregressive general linear models, introducing a conditional spatially autoregressive error model of counts; In addition to these, and as previously mentioned, LeSage *et al.* (2007) estimate a Bayesian hierarchical Poisson spatial interaction model. However, all these studies fail to estimate the spatial autoregressive term, not benefiting from the existence of this term in the model specification.

Reflecting on the theory and application of SAR models to count data, one can conclude that its advancement is quite limited, even when compared to other non-linear spatial models. Regarding binary dependent variables, there is a growing variety of studies based on the logit and probit models with spatial lag, estimated through non-linear GMM (NLGMM) (Pinkse & Slade, 1998; Klier & McMillen, 2008; among others). However, and based on the work of Klier & McMillen (2008), Hays & Franzese (2009) presented a Spatial-Lag Count Model estimated through the non-linear least squares and the GMM estimator. Nevertheless, the authors only present results from a simulation study for a small sample and for low or moderate spatial dependence coefficient values.

Lambert *et al.* (2010) propose another solution: a two-step Limited Information Maximum Likelihood (LIML) estimator for the Spatial Autoregressive Model of Counts. These authors introduce a new specification: An exponential model with the number of counts on location i, $i=1\ldots N$ as a function of, among other covariates, the spatially lagged logarithm of the conditional expected mean of counts on contiguous regions. This model, in addition to enabling the estimation of the SAR coefficient, has the advantage of being invertible, and consequently allows analytic calculation of the spatial partial effects. This essay will explore this specification more closely in the next section.

3. Spatial Autoregressive Model of Counts

The present chapter will be divided into three parts. The first and second part will introduce the Spatial Autoregressive Model of Counts, while in the third part a detailed description of the estimation process will be presented.

3.1 The Model

Beginning by stating the traditional linear spatial model

$$y_i = \rho \sum_{i \neq j} w_{ij} y_j + \mathbf{x_i}\boldsymbol{\beta} + \varepsilon_i; \quad i=1,2,\ldots,N \tag{2}$$

y_i is the dependent variable for the unit i and N denotes the number of spatial units; $\mathbf{x_i}$ represents a 1xK vector of exogenous variables for the unit i; the Kx1 vector $\boldsymbol{\beta}$ is the corresponding vector of unknown regression parameters; ρ denotes the unknown spatial autoregressive parameter; the coefficients $w_{i,j}$ are known non-negative scalars that refer to the *a priori* defined spatial weights of unit j on unit i, with $j \neq i$ and $j = 1, 2, \ldots, N$.; lastly, ε_i represents the i.i.d random error of the unit i.

The spatial lag model can be written in matrix form

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3}$$

being the "reduced-form"

$$\mathbf{y} = \mathbf{A^{-1}X\beta} + \mathbf{A^{-1}\varepsilon} \tag{4}$$

Where $\mathbf{y}= [y_1; y_2;...;y_N]^T$ and $\mathbf{X}=[\mathbf{x_1^T;x_2^T;\ldots;x_N^T}]^T$. The error is $\boldsymbol{\varepsilon} = [\varepsilon_1; \varepsilon_2;...; \varepsilon_N]^T$ and the spatial autoregressive operator is $\mathbf{A} = (\mathbf{I} - \rho\mathbf{W})$. In this case, $\mathbf{A^{-1}}$ reflects the "Leontief Inverse", used in order to quantify the global feedback effects between spatial units. $\mathbf{W}$ is the $N \times N$ spatial weights matrix, with generic element $w_{i,j}$, where $w_{i,j}$,= 0 when $j=i$ **;** $\mathbf{I}$ is the $N \times N$ identity matrix.

However, here the main objective is to model count variables, and thus it is necessary to resort to a non-linear specification. Therefore, following an extension of the linear model, Lambert *et al.* (2010) begins by presenting a specification inspired by the exponential feedback model for time-series (Blundell *et al.*, 1995), which is equal to

$$\mu_i = \exp\left(\rho \sum_{j \neq i} w_{ij} y_j + \mathbf{x_i}\boldsymbol{\beta}\right); \tag{5}$$

Even so, Lambert *et al.* (2010) adverts for the fact that while using this specification, it is impossible to obtain the spatial autoregressive operator inverse, $\mathbf{A^{-1}}$. To solve this problem, the authors propose an alternative specification based on the multiplicative AR models of Zeger & Qaqish (1998),

$$E(y_i|\mathbf{x}_i) \equiv \mu_i = \exp(\mathbf{x_i}\boldsymbol{\beta}) \prod_{i \neq j} E(y_j|\mathbf{x}_i)^{\rho w_{ij}} \tag{6}$$

Which is equivalent to

$$\mu_i = \exp\left[\rho \sum_{i \neq j} w_{ij} \log\left(\mu_j\right) + \mathbf{x_i \beta}\right] \quad (7)$$

This model is invertible, which makes it possible to calculate the "Leontief Inverse", which is essential for the calculation of partial spatial effects.

Equation (7) can also be written in the reduced matrix form

$$\mu_i = \exp\left(\mathbf{A_i^{-1} X \beta}\right) \quad (8)$$

Where $\mathbf{A_i}^{-1}$ is the **i-th** row of the Leontief Inverse.

3.2 Partial Effects

It is important to bear in mind that in the presence of spatial dependence between spatial units, variations in variables in region *i* (x_i), can impact counts in region *i* (μ_i) as well as counts in neighboring regions (μ_j). On that account, LeSage and Pace (2009) proposed the decomposition of partial effects between direct and indirect effects. The Direct Partial Effects (DPE) measure the variation of *y* in region *i*, given a variation of x_{ik} in the same region *i*; while the Indirect Partial Effects (IPE) measure the variation of *y* in a region *i*, given a variation of input in a neighboring region.

Lambert *et al.* (2010) show that the partial derivatives of SAR-Poisson take the following form for any *k* variable, where $\mathbf{x_k}$ is the vector with all the observations of the k^{th} variable.

$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{x}'_\mathbf{k}} = \begin{bmatrix} \partial\mu_1/\partial X_{1k} & \cdots & \partial\mu_1/\partial X_{Nk} \\ \vdots & \ddots & \vdots \\ \partial\mu_N/\partial X_{1k} & \cdots & \partial\mu_n/\partial X_{Nk} \end{bmatrix} = \mathbf{A^{-1}} \boldsymbol{\mu}^{diag} \beta_k \quad (9)$$

Where $\boldsymbol{\mu}^{diag}$ is a diagonal matrix of order *n* with elements μ_i.

LeSage & Chih (2016) state that the elements in the main diagonal of matrix (9) represent the DPE, while the elements off-diagonal are interpreted as the IPE. Nevertheless, the authors went further by also decomposing the IPE in two parcels: spillout and spillin.

The sum of off-diagonal elements in each row of matrix (9) produce a region-specific cumulative spillin effect. These are showing how variations in neighboring *j* regions inputs produce a spillin impact on each region *i*'s output, for example, $\left(\frac{\partial \mu_i}{\partial x_{jk}, i \neq j}\right)$. The region-specific cumulative spillout effect is the sum of off-diagonals elements of each

column of matrix (9). These measure how changes in region *i* knowledge inputs impact neighboring regions *j* outputs, for example, $(\frac{\partial \mu_j}{\partial x_{ik}, i \neq j})$.

With the average partial effects being

$$\boldsymbol{Average\ Direct\ Partial\ Effects} = \frac{\beta_k}{N} \sum_{i=1}^{N} a_{ii}^{-1} \mu_i \quad (10)$$

$$\boldsymbol{Average\ Spillin\ Effects} = \frac{\beta_k}{N} \sum_{i=1}^{N} \sum_{i \neq j} a_{ij}^{-1} \mu_i \quad (11)$$

$$\boldsymbol{Average\ Spillout\ Effects} = \frac{\beta_k}{N} \sum_{j=1}^{N} \sum_{i \neq j} a_{ij}^{-1} \mu_i \quad (12)$$

Where $\boldsymbol{a_{ii}^{-1}}$ refers to the elements on the diagonal of the matrix $\mathbf{A^{-1}}$ and $\boldsymbol{a_{ij}^{-1}}$refers to the off-diagonal elements.

3.3 Estimation

In this sub-chapter, the estimation process for SAR-Poisson applied in Lambert *et al.* (2010) will be analyzed. Later, a new estimation process for the SAR-Poisson will be presented. This new approach tries to solve some identified problems in the first: avoiding the estimation of *log(μ)* in the first-step, in order to not resort to a purely computational transformation to solve the problem of the zeros, while taking in to account the work of Santos Silva & Tenreyro (2006).

Lambert *et al.* (2010) suggests estimating the eq(7) using a two-step Limited Information Maximum Likelihood (LIML).

The two-step LIML estimation process was proposed by Murphy & Topel (1985) as an alternative to FIML[1], since the last needs the derivation of the joint distribution, which is known to be quite demanding. Besides that, maximizing the joint log-likelihood can be numerically difficult, (Greene, 2003)[2].

Traditionally, in spatial econometrics, the problem with the estimation of autoregressive models has to do with the fact that the spatially lagged dependent variable is endogenous. However, with the specification presented by Lambert *et al.* (2010), the spatially lagged variable is the expected mean of counts in neighboring regions *j* (μ_j), and since this is not an observable variable, it must be estimated *a priori*. For this reason, a Two-Step LIML is used.

In the first step proposed by Lambert *et al.* (2010) a set of instrumental variables, Q, are regressed over the observable variable $\sum_{j \neq i} w_{ij} \log(y_j)$, with Q = {X, WX, W^2X}, obtaining the vector of predicted values,

1. Full Information Maximum Likelihood
2. For more details regarding the two-step maximum likelihood estimation, see Greene (2003), chapter 14.7, pages 576-582.

$$Q\hat{\delta}\ with\ \delta = (\mathbf{Q'Q})^{-1}\mathbf{Q'W}log(y_j) \quad (13)$$

In the second step, Lambert *et al.* (2010) uses the values previously predicted as the proxies for the unobserved spatial lagged variable, $Wlog(\mu_j)$, performing a maximum likelihood estimation assuming a Poisson distribution. However, this process has some weaknesses.

Given the nature of the logarithmic function, when performing the first step, y_j can only adopt strictly positive values. This fact is quite restrictive, particularly when zeros are expected to be observable. Therefore, in order to solve this constraint, the authors suggest replacing y_j with the logged-transformed values approximating neighborhood counts, [log(y_j^*)]. Three suggestions are declared: 1) adding an ad hoc constant c to y_j, when y_j is zero, leading to y_j^* = max{c, y_j}; 2) estimating the constant c simultaneously with the other parameters; 3) using an inverse hyperbolic sine (IHS) transformation to the neighboring counts. Nevertheless, these transformations can be computationally demanding, especially when addressing the IHS, while also allowing the creation of bias in the estimation. Furthermore, considering the work of Santos Silva & Tenreyro (2006) another problem arises. As stated before, estimating a loglinear model using OLS will generate biased estimators for the elasticities. However, this is in fact the procedure chosen by Lambert *et al.* (2010) in the first step.

As such, in the present essay, it is proposed a first-step approach where in the estimation of the unobserved spatial lagged variable no computational transformation is needed to deal with the possible problem of zero counts, nor is it necessary to resort to an estimation using a loglinear specification.

For this approach, in the first step, μ_j is estimated, and posteriorly logarithmized. To avoid non-positive predicted values, a Poisson regression is applied in the first-step, forcing the predicted values of $\widehat{\mu_j}$ to always be greater than zero. The use of a Maximum Likelihood Poisson estimation instead of OLS for the loglinear specification, meets the results of Santos Silva & Tenreyro (2006) previously exposed.

In the case proposed, the Poisson probabilistic density function of the first-step is defined by

$$f_1(y_j|Q;\alpha) = \frac{exp(\mathbf{Q\alpha})^{y_j} exp\,(-exp(\mathbf{Q\alpha}))}{y_j!} \quad (14)$$

Where **α** represents a vector of parameters and **Q=[X, WX,W^{2}X]** is the instrument matrix used. The corresponding log-likelihood function is

$$\ln L_1 = \sum_{j=1}^{N} y_j(\mathbf{Q\alpha}) - exp(\mathbf{Q\alpha}) - ln\, y_j! \quad (15)$$

The second-step starts by logarithmizing the predicted values estimated in the first, log($\widehat{\mu_j}$), where $\widehat{\mu_j} = exp\left(Q_i'\widehat{\alpha}\right)$, followed by the multiplication of these with the matrix of spatial weights **W**.

The result is then incorporated in second-step Poisson's probability density function

$$\boldsymbol{f_2}\left(y_i|\mathbf{x_i}, Wlog(\widehat{\mu_j}); \boldsymbol{\beta}, \rho\right) = \frac{exp\left(\mathbf{x_i\beta}+\rho\sum_{j\neq i} w_{ij}\, log(\widehat{\mu_j})\right)^{yi} exp\,(-exp(\mathbf{x_i\beta}+\rho\sum_{j\neq i} w_{ij}\, log(\widehat{\mu_j}))}{yi!} \quad (16)$$

With the following log-likelihood function

$$\mathbf{ln}\, \boldsymbol{L_2} = \sum_{i=1}^{N} y_i(\mathbf{x_i\beta} + \rho\sum_{j\neq i} w_{ij}\ log\left(\widehat{\mu_j}\right)) - exp\left(\mathbf{x_i\beta} + \rho\sum_{j\neq i} w_{ij}\ \log\left(\widehat{\mu_j}\right)\right) - ln\, y_i! \quad (17)$$

This step is quite similar to the first considering the fact that in both a Poisson regression is used.

It should also be noted that when using this methodology, the inference of the second step estimation is invalid. Wooldridge (2002) warns of the fact that standard errors and test statistics obtained from a two-step regression are generally invalid because they ignore the sampling variation in the in the coefficients estimated in the first step. One way to overcome these problems is to estimate standard errors using bootstrap estimation methods.

It should be noted that, when estimating the second-step, if one chooses to use a Pseudo Maximum Likelihood estimator, it is not necessary to guarantee that the dependent variable follows a Poisson distribution, and more relevantly, the dependent variable does not need to be an integer.

The Newton-Raphson algorithm is applied in this essay to maximize the log likelihood function of equations (15) and (17).

4. Monte Carlo Simulations

In this chapter, a series of Monte Carlo simulations are presented, with the intent of comparing various estimation methods proposed for modeling count data with spatial dependence. The proposed SAR-Poisson 1stStep-ML estimator is compared with the SAR-Poisson 1stStep-OLS estimator presented by Lambert *et al.* (2010) where the pre-defined first-step transformation (adding constant, c=1) is used. These estimators are also compared with the Aspatial ML Poisson estimator (ρ=0) and with the SAR-LogLinear,

considering $\log(y^*)$ as dependent variable, where $y^* = \max\{0.5, y\}$, in order to solve a possible zero´s problem.

The comparisons will be based in the Bias and Root Mean Squared Errors results.

4.1 Simulation Design

The present simulation design follows the suggested design by Lambert *et al.* (2010). Therefore, it will be closely related to other spatial econometric simulation studies, such as Kelejian & Prucha (2007) and Klier & McMillen (2008).

The random dependent variable was generated as $\tilde{y}_i \sim$ Poisson (μ_i^{SAR}) with $\mu_i^{SAR}=\exp\left(\mathbf{A_i^{-1}X\beta}\right)$, where $\mathbf{A_i^{-1}}$ is the **i-th** row of $(\mathbf{I}-\rho\mathbf{W})^{-1}$. The design matrix X includes two covariates, X_1 and X_2, not including a intercept. The first was randomly generated from a normal distribution, $X_1 \sim N(1,2)$. Following what Santos Silva & Tenreyro (2006) point out in their simulation study, econometric studies generally incorporate a mix of continuous and dummy variables, thereby, in the present study, a dummy variable was included as covariate, randomly generated from the Bernoulli distribution, $X_2 \sim Bern(0.5)$.

The spatial weights matrix, W, is built using the same two-step process found in other spatial econometrics simulation studies, as it is the case of Silveira Santos & Proença (2019). First, N space units are generated within the unit circle. Secondly, and taking into account the chosen criterion, a matrix W_0 is constructed, and later normalized by rows, so that the sum of all elements of each row is one. In the present study, three different criteria were used in the construction of the matrix W. W_1 is a contiguity matrix created using the nearest neighbor criterion, where it is computationally defined that each unit *i* will have **seven** units *j* as neighbors, these being the seven units *j* closest to *i*. W_2 is created based on an inverse distance criterion, using the Euclidean distance between unit *i* and unit *j*, with $i,j=1,2\ldots N$. The same Monte Carlo experiment is performed using a third matrix W. W_3 is a contiguity matrix created using the nearest neighbor criterion, where it is computationally defined that each unit *i* will have **four** units *j* as neighbors, these being the four units *j* closest to *i*. This contiguity matrix is similar to the contiguity matrix used in the empirical application presented in chapter 5 of this essay.

The matrix W_2 is said to be denser than the matrix W_1, since W_2 contains more nonzeros entries. W_2 contains *N* zeros (main diagonal), while W_1 contains *N(N-7)* zeros (each row has seven nonzero values). On the other hand, matrix W_1 is denser than matrix W_3 (each row has four nonzero values).

The Monte Carlo simulations were conducted for each design of W and for each of the four estimators described above. The sample size, N, varies over the set: 100; 250; 500; 750; 1000. The spatial autoregressive parameter, ρ, varies over the set: 0; 0.2; 0.4; 0.6; 0.8. The parameters associated with variables X_1 and X_2, β_1 and β_2 respectively, are held fixed at 0.5. For each experiment, 1000 replications are used.

The *Bias* of β´s is calculated by subtracting the estimated value from the true value of the coefficient ($\hat{\beta}_j - \beta_j,\ j = 1,2$), or ($\hat{\rho} - \rho$) for the SAR parameter, with the analyzed value being the average of the 1000 replications. RMSE is also calculated for each β coefficient, given that: $RMSE = \sqrt{Bias^2_{\hat{\beta}_j} + Var(\hat{\beta}_j)}$, where $Bias^2_{\hat{\beta}_j}$ is the square of the *Bias* for $\hat{\beta}_j$ calculated before, where j=1,2. $Var(\hat{\beta}_j)$ is the empirical variance in the 1000 replications of the estimated coefficient. RMSE is also calculated for the SAR parameter, ρ: $RMSE = \sqrt{Bias^2_{\hat{\rho}} + Var(\hat{\rho})}$, where $Bias^2_{\hat{\rho}}$ is the square of the *Bias* for $\hat{\rho}$, and $Var(\hat{\rho})$ is the empirical variance in the 1000 replications of the estimated coefficient.

4.2 Monte Carlo Results

It should be noted that the results between W_1 and W_3 are quite similar. This suggests that estimators should not be considerably sensitive to the density of the matrix W, when using the queen contiguity criterion. For this reason, for the remaining results, the analysis will focus only on experiments related to the use of W_1 and W_2 matrices. The results for W_3 can be found in tables A5 and A6 of the appendix.

Table A1, found in the appendix, shows the results for the Bias of the estimated coefficients, β_1 and β_2, for each estimation method and, for both W_1 and W_2 construction criteria of the matrix W. Both SAR-Poisson estimators show similar and quite satisfactory results, with the SAR-Poisson 1stStep-ML presenting lower Bias, in absolute value, for lower levels of spatial dependence, while the SAR-Poisson 1stStep-OLS appears to behave better for ρ values closer to unit. It is worth noting, that both estimators have lower Bias, in absolute value, associated to the continuous variable than to the dummy variable. Note also, that when ρ increases both estimators present a smaller Bias in absolute value when using the matrix W_2 compared to the matrix W_1, nevertheless this difference is residual, especially for a large N. When analyzing the results for the SAR-LogLinear estimator, it is possible to realize that for lower values of spatial dependence,

the estimator is significantly downward biased, and as ρ approaches the unit it becomes upwards biased. In general, in this estimator, β_1 was also found to be less biased, than β_2. Finally, and as expected, the Poisson ML estimator shows progressively worse results as ρ increases, with these being much more pronounced in the estimating dummy coefficient. Even so, it should be noted that when there is no spatial dependence (ρ=0), this estimation method is slightly better than the SAR methods.

In table A2 of the appendix, it is possible to compare the results obtained for the Bias of the spatial autoregressive coefficient, ρ. Globally, the SAR-Poisson 1stStep-ML presents smaller Bias, in absolute value, than the remaining estimators, especially when N is large. However, for ρ=0.8 it shows a higher Bias, in absolute value, particularly in the W_2 matrix. Although slightly worse than the first, the SAR-Poisson 1stStep-OLS presents satisfactory results, namely for high ρ levels. For extreme values of ρ, the SAR-LogLinear presents highly biased results. It is interesting to emphasize that, with the exception of the SAR-Poisson 1stStep-OLS, the Poisson estimators evidence that, as ρ increases, bias grow in absolute value, which can mean that higher levels of spatial dependence imply greater distortion in the estimation of this coefficient. Nevertheless, it is important to stress that, excluding the SAR-Poisson 1stStep-OLS, the use of W_2 matrix results in extra biased estimations.

Table A4 in the appendix, shows the results referring to β´s RMSE. As previously stated, these results take into account not only the Bias of the estimation, but also the sample variance of the estimated coefficients. From a general point of view, and regarding β_1, the SAR-Poisson 1stStep-ML presents the best results, particularly for W_1. However the SAR-Poisson 1stStep-OLS shows a more desirable set of results for higher ρ values. In both estimators, it is noted that as ρ and N increase, the RMSE decreases, showing that the larger the sample, and higher the spatial dependence, the smaller the variance in the estimates. This result is only slightly contradicted when ρ = 0.8. On the other hand, the SAR-LogLinear estimator, presents much higher RMSE´s results, while maintaining the trend of decreasing these as N and ρ rise, only approaching the values of the other two estimators when N = 1000 and ρ = 0.8. As expected, the aspatial ML estimator only shows satisfactory results when ρ = 0. As for β_2, the conclusions are quite similar to β_1, with the disclaimer that the RMSE´s for this coefficient are much higher, especially for smaller N. The W_1 matrix shows slightly better results. The SAR-LogLinear estimator does not present satisfactory results, as it never approaches its peers, even when N and ρ present

high values. Lastly, the aspatial estimator is, again, quite far from the results of the other estimators, showing even more inefficiency in estimating the dummy´s coefficient.

Table A3, in the appendix, reflects the RMSE values regarding the estimation of the coefficient of spatial dependence ρ. Both SAR-Poisson estimators present quite similar results, with the SAR-Poisson 1stStep-ML showing better results as the sample increases. It is also important to note that the SAR-Poisson 1stStep-ML exhibits higher RMSE for matrix W_2 for high levels of spatial dependence, when compared to SAR-Poisson 1stStep-OLS. However, in general, and as mentioned for β´s, the use of W_1 seems to trigger better results. The gradual decrease in RMSE observed in β´s is also noted here. On its turn, the SAR-LogLinear estimator shows, once more, worse results than the other two estimators, especially when ρ takes extreme values.

In summary, by generally analyzing the results and taking into account other simulation studies such as Lambert e*t al.* (2010), Silveira Santos & Proença (2019), Anselin & Le Gallo (2006), Klier & McMillen (2008) and Santos Silva & Tenreyro (2006), it is possible to draw some conclusions. First, it should be noted that the estimator SAR-Poisson 1stStep-ML presents better results than its counterparts, with the exception of high spatial dependence cases, that is ρ=0.8. Since the only difference between this and the estimator proposed by Lambert *et al.* (2010) happens in the non-transformation of the dependent variable and the use of a Poisson regression instead of a loglinear estimation when estimating the first step, this result seems to be in agreement with that found by Santos Silva & Tenreyro (2006). Another interesting result is that there is a greater distortion for the estimated coefficient of the dummy variable compared to the estimated coefficient of the continuous variable, allowing the deduction that the distribution of the explanatory variables can be a condition of its performance, a conclusion that Lambert *et al.* (2010) also finds. Another common conclusion between studies is the fact that the RMSE decreases as the spatial dependence and sample size increase. Another fact already mentioned is that the use of different W matrices produces different results. Several studies have already addressed this issue, with Silveira Santos & Proença (2019) being one of them, where impacts were found in the estimation of the coefficients, for a spatial Probit, given the density of the W matrix. However, the RMSE´s, of both β´s and ρ, appear to be generally higher for the W_2 matrix, suggesting that the variance of the estimated coefficients may, somehow, be related to the density of the spatial weights matrix chosen. Nonetheless, this aspect should be comprehensively studied in the future. Another expected conclusion was the poor performance of the Aspatial ML Poisson estimator in

the presence of spatial dependence, which presented an accentuated upward Bias for both coefficients. This result is in agreement with Anselin and Le Gallo (2006) who found biased and inconsistent estimators when spatial dependence was not taken into consideration. Finally, there is a significant increase in the importance of Bias when the SAR-LogLinear is used in the estimation, this agrees with the possibility of biased estimations being produced when using a linear model to explain count variables. In addition, it is interesting to note that the distortion of results is more significant for values of ρ near the unit, which is in line with the results of Klier and McMillen (2008). These infer that for high levels of spatial dependence, the estimation methods proposed for linearized spatial models obtain unsatisfactory results when compared with lower ρ values.

5. Empirical Example

In this chapter, an empirical example will be presented, where a knowledge production function will be estimated. Given the satisfactory results of the previous chapter, the SAR-Poisson 1stStep-ML estimator proposed in this essay will be used to estimate the model. For comparative purposes, the same model will be estimated using the three alternative estimators proposed in the simulation study in the previous chapter. In addition to the above, for each estimator, two models with different spatial weights matrix are estimated: the first using the Queen contiguity criterion and the second using a Euclidean Inverse Distance (EID) matrix.

5.1 Exploratory Data Analysis

The data was retrieved from Eurostat *(Eurostat regional database)*. The database created by the author contains data of 234 NUTS II regions, split between 24 European Countries, of which 22 belong to European Union, with the addition of the United Kingdom and Norway. All data refers to 2012. More detailed information can be found in the Appendix B notes.

The objective of the essay is to study the production of knowledge, therefore, it was decided to follow the suggestion of Buesa *et al.* (2010) and use the number of patents in a given region per million inhabitants as a proxy for the creation of innovation. The amount was rounded to the nearest integer in order to obtain a discrete variable.

The description set of used variables in the present study, and the expected sign of the estimated coefficient associated, can be found in the Table B1 of the Appendix. The descriptive statistics of these variables are shown in table B2 of the Appendix. In addition to these, it is possible to find in the table B3 in the Appendix the correlation matrix of the variables used in the study.

As previously announced, there are several channels of knowledge production. Therefore, the data related to expenditure in R&D and to the number of people working full-time in R&D, were divided in three sources: the first refers only to the private initiative; the second is linked only to the public sector; and finally, the third portion refers only to the Universities. This division will allow the deduction of the different impacts of R&D investment in the creation of innovation, based on the channel used, enabling a more refined analysis (Zhang *et al.*, 2020). Theoretically it would be expected that the sign of the estimated coefficients related to these variables would all be positive, since more R&D expenditure, as well as more full-time R&D employees, should trigger an increase in knowledge creation. However, the literature suggests that this happens only for the private sector. Both the public sector and universities, the signal often appears to be negative. Although for the second case the explanation lies in the fact that the great university contribution to knowledge creation arises in the form of scientific articles, for the public sector the explanation presented is that the public sector is inefficient in the production of knowledge (Zhang *et al.*, 2020; Ferreira & Godinho, 2015).

To capture the effect of the "innovative environment", data on the percentage of graduates in the population between 25 and 65 years old was also collected, using this measure as a proxy for the level of education of the population in the region. GDP *per capita* was used as a proxy for technological sophistication, while the tuberculosis mortality rate was considered as a proxy for the level of poverty of the inhabitants, once several studies relate tuberculosis with poverty (Ferreira & Godinho, 2015). In addition to these, the number of inhabitants was defined as the control variable. It is expected that a better socio-economic environment will boost innovation (Ferreira & Godinho, 2015).

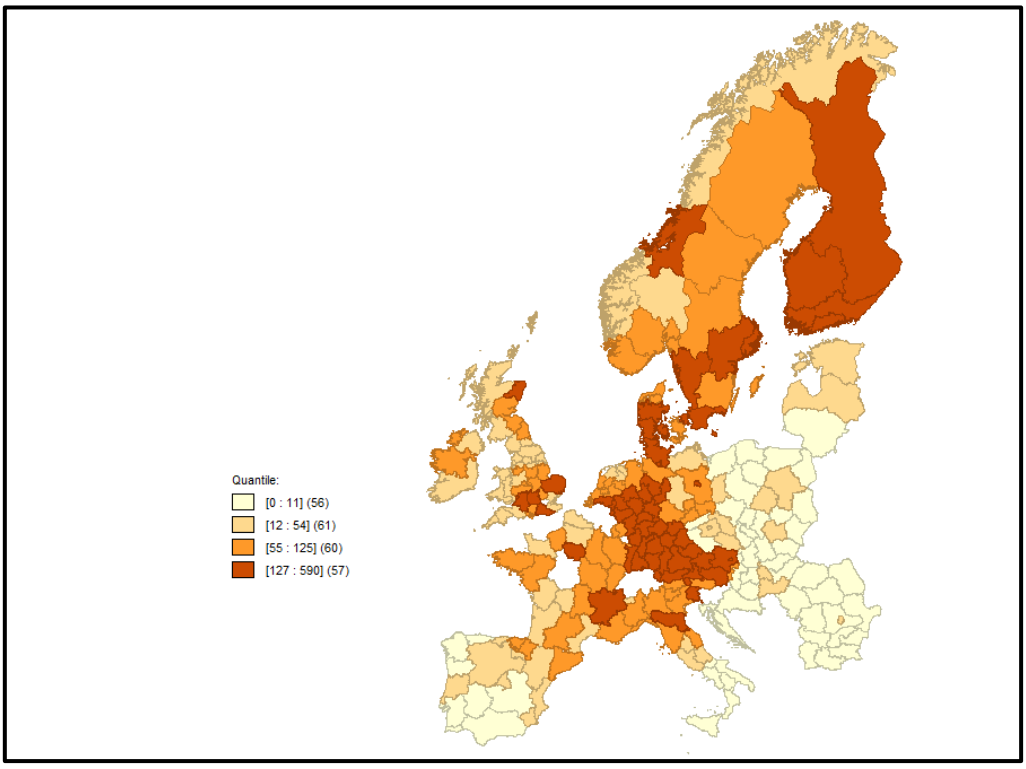
In table B2 of the Appendix it is possible to analyze the mean, standard deviation, and the quartiles of the studied variables. It is observable that the number of regions with 0 patents is equivalent to 6% of the all sample, a trait that is characteristic of count variables. In fact, 25% of the sample has between 0 and 12 patents, which shows the right skewed distribution typical of this type of variable. As for the R&D channels, it is worth noting a greater investment, in average terms, in the private sector than in the alternatives, which

may indicate a greater importance of this sector. Likewise, there are more full-time-Time workers in the R&D process in the private sector. Regarding GDP *per capita*, and it is possible to see that the value of the 1st quantile is 16.74, with the 3rd quartile having the value of 33.876, and the maximum of 84.047. This asymmetry, represented by a right skewed distribution, is typical of variables referring to the income, showing, once more, the wealth gap between countries, even when they belong to the same economic "integration region".

5.2 Exploratory Spatial Analysis

Analyzing the Spatial Distribution Map of the variable *Pat* per quartile in Figure 1 it is possible to verify the existence of a cluster effect, given the concentration of patenting taking place in Central Europe, South England and Scandinavia, with the number of new patents in southern and eastern Europe being residual. Likewise, the spatial correlogram (Appendix figure B4) shows that there is a strong spatial correlation, in relation to the number of new patents, between close regions, decreasing as the distance between regions increases.

Figure 1: Spatial Distribution Map of the variable *Pat* per quartile - Year 2012



Source: Eurostat, author calculations; Software: QGIS

Through the Moran index test, which tests for the presence of global spatial dependence, and the analysis of the Moran diagram, it is possible to draw a more careful conclusion about the problem.

However, this index is sensitive to the spatial weight matrix used. As stated earlier, in this essay two different W matrices will be used: A Queen contiguity matrix and an EID matrix. Regarding the first, the histogram of the number of neighbors is presented in figure B5 in the Appendix. With this matrix, 17 regions have the minimum number of neighbors (1), while 1 region has the maximum number (12), with the average number of neighbors being 4.42. For both matrices, Moran's I Test for Spatial Autocorrelation shows a positive and significant spatial dependence, as can be seen in table B6 of the Appendix. This conclusion is supported by the Moran diagrams (figures B7 and B8 in the Appendix). Analyzing the latter, it is worth noting that most of the observations are in the 1st and 3rd quadrant, and therefore, the majority of regions with more (less) new patents have neighboring regions also with a greater (less) number of new patents. However, the sole analysis of the Moran index can present distorted results, therefore, it is now imperative to look at LISA[3]. In figure B8 of the Appendix, this indicator is visible for the two matrices, both of which detect the presence of two highly patenting clusters in central Europe and Scandinavia, and the presence of low patenting clusters in the Iberian Peninsula and Eastern Europe. There are also two more low-patent clusters in northern Britain and southern Italy, mostly prominent in the inverse distance matrix. Figure B10 of the Appendix shows the LISA Significance Map, inferring that the results are more significant for the Central European cluster and for the Iberian Peninsula and Eastern Europe clusters. Regarding the Bivariate analysis, the variable Patents is spatially related to the other variables studied. Figures B11 and B12 in the Appendix refer to the set of Moran dispersion diagrams for the Queen and EID matrix, respectively. In these, the relationship between the variable *Pat* (abscissa axis) and the spatially lagged covariates (ordinate axis) is analyzed. It is possible to observe that the only variable that has a negative Bivariate Moran´s index is the mortality rate due to tuberculosis. It should also be noted that both the number of Full-Time workers in R&D in government institutions and the number of Full-Time workers in R&D in Universities have a Moran index close to zero, thus, it can be interpreted that there is no spatial correlation between the number of patents in a given region, and the number of Full-Time R&D workers in non-private institutions in neighboring regions.

3. Local Indicator of Spatial Association

5.3 Estimation of coefficients

This section presents the results of the estimation of the proposed model. As mentioned above, four different estimators will be used. The results for the SAR-Poisson estimations are shown in Table 1 and the results of the Log-Linear and Aspatial Poisson ML estimations are visible in Table 2.

As referenced earlier, the inference of these estimation processes is invalid. Therefore, to solve this problem, the bootstrap method was used to estimate the standard errors.

Table 1: SAR-Poisson coefficients and APE estimations

SAR-Poisson 1stStep-ML

Variable	Coefficients	Bootstrap SE	Average Partial Effects		
			Direct	Spillin	Spillout
ρ	6,81E-01 ***	0,06838			
R&D_B	8,91E-04 ***	0,00034	0,0934	0,1743	0,1683
R&D_G	-2,15E-03 *	0,00130	-0,2250	-0,4200	-0,4054
R&D_U	-3,21E-04	0,00079	-0,0336	-0,0628	-0,0606
Pers_B	-1,33E-05	0,00003	-0,0014	-0,0026	-0,0025
Pers_G	2,75E-05	0,00006	0,0029	0,0054	0,0052
Pers_U	5,07E-05	0,00005	0,0053	0,0099	0,0096
Educ	2,58E-04	0,01165	0,0270	0,0504	0,0487
Pop	-3,21E-09	9,62E-08	-3,36E-07	-6,28E-07	-6,06E-07
GDP	3,81E-02 ***	0,01003	3,9933	7,4536	7,1953
Mort	-1,95E-01 **	0,09624	-20,4060	-38,0886	-36,7686
Log Likelihood	-6557,154				
W	Queen				
N	234				

Variable	Coefficients	Bootstrap SE	Average Partial Effects		
			Direct	Spillin	Spillout
ρ	9,15E-01 ***	0,07167			
R&D_B	1,06E-03 ***	0,00027	0,1071	1,1106	0,9980
R&D_G	-2,14E-03 *	0,00112	-0,2165	-2,2452	-2,0174
R&D_U	-8,18E-06	0,00058	-0,0008	-0,0086	-0,0077
Pers_B	-3,53E-06	0,00002	-0,0004	-0,0037	-0,0033
Pers_G	3,39E-05	0,00005	0,0034	0,0355	0,0319
Pers_U	4,18E-05	0,00004	0,0042	0,0438	0,0393
Educ	-5,16E-03	0,00883	-0,5215	-5,4085	-4,8598
Pop	-6,79E-08	8,17E-08	-6,87E-06	-7,12E-05	-6,40E-05
GDP	2,39E-02 ***	0,00759	2,4161	14,0572	13,5154
Mort	-4,49E-01 ***	0,09965	-45,4119	-70,9559	-63,1820
Log Likelihood	-5274,879				
W	Inverse distance				
N	234				

SAR-Poisson 1stStep-OLS ad hoc constant c=1

Variable	Coefficients	Bootstrap SE	Average Partial Effects		
			Direct	Spillin	Spillout
ρ	6,19E-01 ***	0,07821			
R&D_B	9,18E-04 ***	0,00035	0,0878	0,1126	0,1093
R&D_G	-2,01E-03	0,00157	-0,1921	-0,2464	-0,2391
R&D_U	-5,85E-04	0,00083	-0,0560	-0,0718	-0,0697
Pers_B	-1,59E-05	0,00003	-0,0015	-0,0020	-0,0019
Pers_G	3,43E-05	0,00006	0,0033	0,0042	0,0041
Pers_U	1,86E-05	0,00005	0,0018	0,0023	0,0022
Educ	7,55E-03	0,01279	0,7221	0,9258	0,8984
Pop	6,10E-08	1,14E-07	5,84E-06	7,49E-06	7,27E-06
GDP	4,42E-02 ***	0,01127	4,2288	5,4222	5,2617
Mort	-6,72E-02	0,09071	-6,4270	-8,2409	-7,9969
Log Likelihood	-7704,048				
W	Queen				
N	234				

Variable	Coefficients	Bootstrap SE	Average Partial Effects		
			Direct	Spillin	Spillout
ρ	9,49E-01 ***	0,08829			
R&D_B	1,17E-03 ***	0,00032	0,0997	1,6657	1,4728
R&D_G	-2,85E-03 **	0,00132	-0,2180	-3,6447	-3,2227
R&D_U	-1,02E-05	0,00066	-0,0635	-1,0618	-0,9389
Pers_B	-8,47E-06	0,00002	-0,0017	-0,0289	-0,0256
Pers_G	5,29E-05	0,00006	0,0037	0,0622	0,0550
Pers_U	2,17E-05	0,00004	0,0020	0,0337	0,0298
Educ	6,36E-03	0,01044	0,8194	13,6962	12,1102
Pop	-2,99E-08	8,98E-08	-6,63E-06	-1,11E-04	-9,79E-05
GDP	2,93E-02 ***	0,00864	4,7988	80,2133	70,9246
Mort	-4,08E-01 ***	0,10894	-7,2934	-121,9107	-107,7934
Log Likelihood	-5992.116				
W	Inverse distance				
N	234				

Notes:

1) Standart errors were computed using Bootstrap method.
2) Significance levels: *10%, **5%, ***1%.
3) SAR-Poisson 1stStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients are estimated using a poisson regression.
4) SAR-Poisson 1stStep-OLS is estimated using a two-step process. In the first step, the unobservable variable $Wlog(\mu_j)$ is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_j=0, and in the second step, the coefficients are estimated using a poisson regression.
5) All estimations were computed using the software R.

In a first estimation, all variables contained in table B1 were used. However, given the individual non-significance of the coefficients referring to the total R&D personnel and researchers full-time variables (*Pers_B; Pers_G; Pers_U*), regressions of restricted models, not containing these variables, were performed. Nevertheless, using LR tests to test for joint significance, the variables proved to be jointly significant at 5%, and as such the final models presented are the non-restricted version (Table 1 and Table 2). The estimations of the coefficients and averaged partial effects of SAR-Poisson restricted regressions can be found in the Table B13 of the Appendix. Table B14 of the Appendix presents the coefficients and averaged partial effects of SAR-LogLinear and Aspatial Poisson restricted regressions. The results for the Likelihood Ratio test for joint significance of the variables Pers_B, Pers_G and Pers_U can also be found in tables B13 and B14 of the Appendix.

In all estimations, the coefficient related to the spatially lagged variable is quite significant (P-value <0.01). This coefficient is always positive in all models in which it was estimated, thus inferring that there is a strong positive spatial dependence (0.68 and 0.61 in Queen contiguity SAR-Poisson 1stStep-ML and Queen contiguity SAR-Poisson 1stStep-OLS, respectively) between the regions regarding the number of patents (an increase in patents in neighboring regions means an increase in the number patents in the region itself), which meets the results of Zhang *et al.* (2020) and Furková (2019). The LogLinear estimate (Table 2) has a lower ρ, although this may be due to the use of a loglinear estimation, it is important to remind here that for intermediate values of spatial dependence, the simulation study found a downward bias for this estimator. It should also be noted in Table 1 that the spatial autoregressive coefficient is extremely high in the estimates using an EID matrix. This result should be analyzed with caution, remembering the upward bias found when using this matrix W for high spatial dependence values. However, it should be noted that in most studies on the theme of knowledge creation, the EID matrix is excluded from empirical examples, largely because it generates questionable results.

As for the remaining explanatory variables, the variable *R&D_B* appears to be significant at 1% in all estimates with spatial dependence (Table 1 and Table 2), and significant at 10% in aspatial estimation (Table 2), always with a positive sign. In contrast, *R&D_U* is not significant, which can be explained by the fact that university contributions are mostly in the form of scientific articles and not patents. On the other hand, *R&D_G* is significant at 10% in five out of the seven estimated models, however, presents a

negative sign. These results converge with those of Krammer (2009), Zhang *et al.* (2020) and Ferreira & Godinho (2015) which also conclude the existence of inefficiency in the public R&D sector. In addition, these authors also conclude that it is R&D expenditures in the private sector that trigger greater knowledge creation.

Table 2: SAR-LogLinear & Aspatial Poisson ML coefficients and APE estimations

SAR-LogLinear ad hoc constant c=0.5											
			Average Partial Effects						Average Partial Effects		
Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout	Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout
ρ	4,43E-01 ***	0,00684				ρ	4,65E-01 ***	0,09088			
R&D_B	1,12E-03 ***	0,00007	0,3769	0,2717	0,2608	R&D_B	1,24E-03 ***	0,00041	0,4006	0,3418	0,3392
R&D_G	-2,06E-03 ***	0,00024	-0,1304	-0,0933	-0,0902	R&D_G	-2,28E-03	0,00175	-0,1379	-0,1216	-0,1171
R&D_U	-2,30E-04	0,00014	-0,0332	-0,0237	-0,0228	R&D_U	-1,69E-04	0,00091	-0,0232	-0,0197	-0,0196
Pers_B	-2,19E-07	0,00001	-0,0013	-0,0010	-0,0009	Pers_B	-2,15E-06	0,00003	-0,0126	-0,0105	-0,0105
Pers_G	1,29E-05	0,00001	0,0202	0,0141	0,0139	Pers_G	2,90E-05	0,00007	0,0433	0,0373	0,0362
Pers_U	1,59E-05 **	0,00001	0,0559	0,0394	0,0380	Pers_U	1,36E-05	0,00006	0,0456	0,0371	0,0381
Educ	9,07E-03 ***	0,00087	0,2630	0,1833	0,1819	Educ	6,84E-03	0,01280	0,1896	0,1614	0,1599
Pop	3,69E-08 **	1,47E-08	0,0779	0,0536	0,0533	Pop	3,27E-08	1,25E-07	0,0658	0,0547	0,0553
GDP	5,52E-02 ***	0,00122	1,5770	1,1086	1,0887	GDP	5,47E-02 ***	0,01114	1,4910	1,2743	1,2603
Mort	-2,55E-01 ***	0,01300	-0,2717	-0,1799	-0,1896	Mort	-3,13E-01 ***	0,06760	-0,3196	-0,2592	-0,2696
Log Likelihood	-321.9729					Log Likelihood	-324.8753				
W	Queen					W	Inverse distance				
N	234					N	234				

Aspatial Poisson ML			
Variable	Coefficients	Bootstrap SE	Average Partial Effects
R&D_B	7,93E-04 *	0,00048	0,06147
R&D_G	-3,61E-03	0,00280	-0,28006
R&D_U	2,90E-04	0,00117	0,02251
Pers_B	-1,60E-05	0,00003	-0,00124
Pers_G	4,99E-05	0,00009	0,00386
Pers_U	-1,62E-04 *	0,00006	-0,01259
Educ	6,70E-02 ***	0,01348	5,19633
Pop	5,13E-07 ***	1,07E-07	0,00004
GDP	4,93E-02 ***	0,01578	3,82501
Mort	-2,49E-03	0,08504	-0,26660
Log Likelihood	-11994.63		
W			
N	234		

Notes:

1) Standart errors were computed using Bootstrap method.
2) Significance levels: *10%, **5%, ***1%.
3) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients ρ is estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the second step is zero.
4) All estimations were computed using the software R.

Regarding the variables related to the "Innovative Environment", *Educ* and *Pop* appear significant at 5%, only in the SAR-Loglinear model using a Queen Contiguity matrix and in the Aspatial Poisson model, both with a positive sign (Table 2). The *GDP* variable is statistically significant at 1% in all models, with a positive sign. Finally, the mortality rate appears significant at 5% in most of the estimated models, but this time with a negative sign. These results are in line with expectations, as a better level of education for the population, added to greater technological sophistication and associated with lower levels of poverty and higher quality of life are factors that, generally, foster the growth of innovation in a region. These results corroborate studies such as Ferreira & Godinho

(2015) and Acs *et al.* (2002) who conclude that the "Innovative Environment" is an essential mechanism to increase knowledge creation.

5.4 Estimation of Averaged Partial Effects

Given the non-linearity of the model, it is through the average partial effects (APE) that it is possible to quantify the impact of the variation of the explanatory variables on the dependent variable, on average, *ceteris paribus*.

As for the direct effects, it should be noted that the estimated values among all models are, in general, similar, being higher in the SAR estimators. This happens due to the spillovers mediated through the spatial multiplier, a result also found by Lambert *et al.* (2010). The variables referring to the “Innovative Environment” have a very high weight in the creation of knowledge. In the case of SAR-Poisson 1stStep-ML with the W Queen matrix, the increase of 1 P.P (percentage points) in the tuberculosis mortality rate in the region results, on average, a drop of 20.4060 patents per million inhabitant, *ceteris paribus*. On the other hand, an increase in GDP *per capita* of just 100 euros in the region, may trigger an increase, on average, of 0.4 patents per million inhabitants in their own, ceteris paribus (Table 1).

Regarding the variables of expenditure on R&D, these can present the most interesting results for economic decision makers. An increase of 10 euros *per capita* in a region in public R&D entities means, on average, a decrease of 2,25 patents in that region per million inhabitant, *ceteris paribus*. Now, given the inefficiency inferred there, a policy maker must transfer the financial resources of these institutions to private R&D companies, since these, for each increase of 10 euros *per capita* in R&D expenses trigger an increase of approximately 1 patenting per million inhabitant, on average, *ceteris paribus* (Table 1). This result is corroborated by almost all the estimated models. The spatial distribution map of SAR-Poisson 1stStep-ML Queen Contiguity Direct Partial Effect (DPE) *per quartil* related to the variable *R&D_B* is visible in figure B15 of the Appendix. It is visible that the regions with the most efficient companies in transforming R&D expenses into patents are located in the Center of Europe, in the South of Great Britain and in Scandinavia. Therefore, regions in Eastern Europe and Southern Europe must employ a reform in the private R&D creation system, seeking an increase in its efficiency. These reforms undergo the recruitment of more qualified personnel and the investment in more sophisticated technology.

As for the indirect effects, in general, the estimated coefficients are higher when using the EID matrix. This result is expected given the considerable increase in neighbors. Once again, the variables related to the "innovative environment" appear to be quite striking, showing that not only the socioeconomic situation of the region is central to the creation of knowledge, but also the interregional environment.

Analyzing the results shown in table 1, referring to the SAR-Poisson 1stStep-ML estimation, using the Queen contiguity matrix, as for the R&D expenditure variables, investment in government R&D institutions also does not benefit neighboring regions the knowledge creation process, since both the spillout effect and spillin effect are negative. On the other hand, investment in private R&D in one region will have a very positive impact in neighboring regions: with a variation of 10 euros per inhabitant in private R&D expenditure in all neighboring *j* regions, results in an increase of 1.74, on average, new patents in the region *i*, *ceteris paribus*. Conversely, the increase of 10 euros per inhabitant in the region *i* in expenditure on private R&D result in an increase, on average, of 1.68 in the set of all neighboring *j* regions. This fact highlights the presence of Knowledge Spillovers between regions. Figures B16 and B17 of the Appendix refer the spatial distribution map of SAR-Poisson 1stStep-ML Queen Contiguity spillin and spillout effect *per quartil*, respectively, of the variable *R&D_B.* It can be concluded that in addition to the central European cluster that shows a strong relationship in the creation of knowledge, regions in southern and eastern Europe, as well as some regions in southern England, have a remarkable capacity for absorbing innovation. Regarding the spillout effects, the European Center and Scandinavia cluster present themselves as the biggest "exporters" of Knowledge Spillovers. Interestingly, some regions that present less DPE with the investment in private R&D, as is the case with the regions of Eastern Europe and the North of the United Kingdom, present higher values of spillout and spillin. Therefore, it is possible to conclude that despite having a lower capacity for innovation, these regions show a strong interconnection between them, which leads to high levels of knowledge spillovers. This can be explained by a possible commitment of companies to strong interregional cooperation links, making the investment in one company positively reflected in the others. These links can be explained as a strategy to overcome the difficulty of competing solo against regions with high levels of patenting. As such, political-economic decision-makers in regions with less patent capabilities should create incentives for the creation of knowledge-sharing networks, thus enabling increased competitiveness.

6. Conclusion

The present essay provides some analysis on the main determinants of knowledge creation, while also quantifying the mechanisms of Knowledge Spillovers between different European NUTS II regions. On the other hand, it introduces a refinement in the estimation procedure of a new SAR-Poisson estimator, which despite being similar to the methodology proposed by Lambert *et al.* (2010), aims to eliminate the bias generated in the estimation proposed by the seconds. Contrary to these authors who carry out an OLS estimation when performing the first step, in this essay, it is presented a first-step Poisson Maximum Likelihood approach in the estimation of the unobserved spatial lagged variable - the expected mean of counts in neighboring regions. In the present, no computational transformation is needed to deal with the possible problem of zero counts, thus avoiding the undesirable creation of bias in the estimation, and at the same time, it is taken into account the work of Santos Silva & Tenreyro (2006) who state that estimating a loglinear regression using OLS can generate biased estimators.

The performance of the new SAR-Poisson 1stStep-ML estimator was evaluated through a Monte Carlo simulation study it was concluded that it was better behaved than the alternative estimators for both small and large samples, and only at very high levels of spatial dependence ($\rho = 0.8$) did the new estimator present higher Bias and RMSE. Other conclusions to highlight are: the existence of a greater bias in the estimation of dummy variables compared to continuous variables; the RMSE of the estimated coefficients is mostly higher for the EID matrix in comparison to the Queen matrix, thus assuming that the sample variance may somehow be related to the density of the chosen weight matrix; Both the aspatial ML estimator and the Loglinear estimator have unsatisfactory performances, being quite biased in comparison to the SAR-Poisson estimators, showing the consequences of not considering the existence of spatial dependence or ignoring the nature of the dependent variable, respectively.

In respect to the results of the empirical application, it is possible to infer that the hypothesis of the existence of spatial dependence on the creation of innovation in Europe cannot be rejected. Regions with a greater number of new patents are surrounded by regions with a major number of new patents. In addition, it is inferred that social and economic factors are determinant in the creation of knowledge, as it is the case of quality life standards and technological sophistication. It also appears that public R&D institutions are inefficient, contrary to private institutions, the latter being the major

promoters of innovation creation in the analyzed regions. It is also inferred that the increase in R&D expenditure by private institutions positively influences the creation of innovation in neighboring regions. Finally, it is concluded that regions with low levels of knowledge creation try to overcome this obstacle by strengthening relations with neighboring regions, increasing the absorptive and segregative capacity for innovation, thus creating strong clusters of knowledge sharing.

Given these conclusions, political and economic decision-makers are advised to:

1) Seek to develop a fruitful regional environment for the creation of innovation, investing in the fight against poverty, in the education of the population and in technological sophistication.
2) Reallocate investment in public R&D institutions to private initiative institutions, enabling them to become even more efficient in creating knowledge.
3) Promote interregional relations between companies, benefiting the flow of innovation, and facilitating the progression of knowledge, especially in regions with difficulties to do it solo.

Some suggestions are presented that could be of interest to investigate below:

1) It was concluded that the possible problem of zeros in the procedure proposed by Lambert *et al*. (2010) is one of the sources of bias in the estimation, as such, it will be interesting to understand how the different proposed estimators behave when the number of zeros in the sample increases.

2) A GMM estimator can be applied, as an alternative to ML, where no assumption about the distribution is made. It would then be interesting to study also through a set of Monte Carlo experiments, how the GMM estimator would behave. The study should focus on different levels of spatial dependence, different sample sizes and different contiguity matrices.

3) Regarding the theme of Knowledge Spillovers, and given the inconclusive results concerning the importance of the number of full-time people in the R&D process, a deeper analysis would be interesting in an attempt to understand if it is the characteristic of companies or if it is the natural talent of the inventors the real driver of innovation creation.

References

Acs,Z., Anselin, L. and Varga, A. (2002). Patents and innovation counts as measures of regional production of new Knowledge. *Research Policy,* 31(7), 1069-1085.

Acs, Z., Audretsch, D.B. (1988). Innovation in large and small firms: na empirical analysis. *The American Economic Review,* 78, 678-690.

Anselin, L., Varga, A. & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics,* 42(3), 422–448.

Anselin, L. & Le Gallo, J. (2006). Interpolation of air quality measures in hedonic house price models: spatial aspe*cts. Spatial Economic Analysis,* 1(1), 31–52.

Arundel, A., Kabla, I. (1998). What percentage of innovations are patented? Empirical estimates for European firms. *Research Policy,* 27(2), 127-141.

Autant-Bernard, C. (2012). Spatial Econometrics of Innovation: Recent Contributions and Research Perspectives. *Spatial Economic Analysis,* 7(4), 403-419.

Autant-Bernard, C., LeSage, J. (2011). Quantifying knowledge spillovers using spatial econometric tools. *Journal of Regional Science,* 51(3), 471–496.

Blundell, R., Griffith, R., & Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of econometrics,* 108(1), 113-131.

Buesa,M., Heijis, S. and Baumert, T. (2010). The determinants of regional innovation in Europe: A combined factorial and a regression knowledge production function approach. *Research Policy* 39(6), 722-735.

Cameron, A. C., & Trivedi, P. K. (2005). "*Microeconometrics: methods and applications*." Cambridge university press.

Caragliu, A., & Nijkamp, P. (2016). Space and knowledge spillovers in European regions: the impact of different forms of proximity on spatial knowledge diffusion. *Journal of Economic Geography,* 16(3), 749-774.

Di Cagno, D., Fabrizi, A., Meliciani, V., & Wanzenböck, I. (2016). The impact of relational spillovers from joint research projects on knowledge creation across European regions. *Technological Forecasting and Social Change*, *108*, 83-94.

European Commission. (2001). Recherche et développement: statistiques annuelles. Luxembourg.

Ferreira, V., & Godinho, M. M. (2015). 14. The determinants of innovation. Dynamics of Knowledge Intensive Entrepreneurship. *Business Strategy and Public Policy*, 304.

Fritsch,M. (2002). Measuring the quality of regional innovation systems: a knowledge production function approach. *International Regional Science Review,* 25(1), 86-101.

Furková, A. (2019). Spatial spillovers and European Union regional innovation activities. *Central European Journal of Operations Research* 27(3), 815–834.

Furman, J. and Hayes, R. (2004). Catching up or standing still? National innovation productivity among follower nations. 1978-1999. *Research Policy*, 33(9), 1329-1354.

Furman, J., Porter, M. and Stern, S. (2002). The determinants of national innovation capacity. *Research Policy*, 31(6), 889-933.

Greene, W. H. (2003). "*Econometric analysis."* Pearson Education India.

Griliches, Z. (1979). Issues in assessing the contribution of R&D productivity growth. *Bell Journal of Economics*, 92-116.

Hays, J.C., Franzese, R.J. (2009). A Comparison of the Small-Sample Properties of Several Estimators for Spatial-Lag Count Models. *Unpublished Working Paper, University of Illinois Champaign-Urbana*.

Jiao,C.H., Chen,Y.F. (2018). R&D resource allocation, spatial correlation and regional TFP growth. *Stud. Sci. Sci*, 36, 81-92.

Kaiser, M.S., Cressie, N. (1997). Modeling Poisson variables with positive spatial dependence. *Statistics and Probability* 35(4), 423–432.

Kelejian, H.H., Prucha, I.R. (2007). HAC estimation in spatial framework. *Journal of Econometrics* 140(1), 131-154.

Klier, T., McMillen, D.P. (2008). Clustering of auto supplier plants in the United States: generalized method of moments spatial logit for large samples. *Journal of Business and Economic Statistics*, 26(4), 460–471.

Kleinknecht, A., van Montfort, K, Brouwer, E. (2002). The non-trivial choice between innovation indicators. *Economics of Innovation and New Technology*, 11(2), 109-121.

Krammer, S. (2009). Drivers of national innovation in transition: Evidence from a panel of Eastern European countries. *Research Policy*, 38(5), 845-860.

Lambert, Dayton M., Brown, Jason P., Floriax, Raymond J.G.M. (2010). A two-step estimator for a spatial lag model of counts: Theory, small sample performance and na application. *Regional Science and Urban Economics*, 40(4) 241-252.

LeSage, J. P., & Chih, Y.-Y. (2016). Interpreting heterogeneous coefficient spatial autoregressive panel models. *Economics Letters*, 142, 1–5.

LeSage, J. P., Fischer, M. M., & Scherngell, T. (2007). Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects. *Papers in Regional Science*, 86(3), 393-421.

LeSage, J.P., Pace, R.K. (2009). *"Introduction to Spatial Econometrics."* CRC Press, Taylor and Francis Group.

Lucas R. E. (1988). On the Mechanics of economic development. *Journal of Monetary Economics* 22(1), 3-42.

Maggioni, M., Nosvelli, M. & Uberti, E. (2007). Space vs. networks in the geography of innovation: a European analysis. *Papers in Regional Science*, 86(3), 471–493.

Mansfield, E. (1965). Rates of return from industrial research and development. *American Economic Review,* 55, 310–322.

Marshall, A. (1920). Principes d´économie politique. Paris, London e New York, Gordon & Breach, reprint 1971, vol.2, 576 pp.

Miguèlez, Ernest, Moreno, Rosina. (2013). Research Networks and Inventors Mobility as Drivers of Innovation: Evidence from Europe. *Regional Studies,* 47(10), 1668-1685.

Murphy. K., Topel, R. (1985). Estimation and inference in two step econometric models. *Journal of Business and Economic Statistics*, 3(1), 370-379.

OCDE. (2004). Compendium of Patent Statistics, Paris.

OCDE. (1999). Managing National Innovation Systems, Paris.

Pinske, J., Slade, M.E. (1998). Contracting in space: an application of spatial statistics to discrete-choice model. *Journal of Econometrics*, 85(1), 125–154.

Romer, P.M. (1990). Endogenous Technological Change. *J. Polit. Econ*. 98, 71-102.

Silveira Santos, L. Proença, I. (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.

Santos Silva, JMC, Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics,* 88(4), 641–58.

Schabenberger, O., & Pierce, F. J. (2001). "*Contemporary statistical models for the plant and soil sciences."* CRC press.

Smith,K. (2005). Measuring Innovation.

Wooldridge, J. M. (2002). "*Econometric analysis of cross section and panel data"*. MIT Press. Cambridge, MA, 108.

Zeger, S.L., Qaqish, B., (1988). Markov regression models for time series: a quasilikelihood approach. *Biometrics,* 44, 1019–1031.

Zhang, F., Wang, Y., Liu., W. (2020). Science and Technology Resource Allocation, Spatial Association, and Regional Innovation. *Sustainability,* 12(2), 694.

Appendix A

Table A1:
Bias: SAR-Poisson, SAR-LogLinear and Aspatial ML Poisson estimates: β_1, β_2

β1-SAR-Poisson 1stStep-ML

W1

Rho/n	100	250	500	750	1000
0.0	-0,0028	-0,0011	-0,0007	0,0000	0.0000
0.2	-0,0013	0,0000	-0,0007	0,0001	-0,0004
0.4	-0,0006	0,0000	-0,0006	0,0002	0,0002
0.6	0,0003	0,0007	0,0015	0,0012	0,0015
0.8	0,0049	0,0046	0,0040	0,0032	0,0043

W2

Rho/n	100	250	500	750	1000
0.0	-0,0006	-0,0010	0,0001	-0,0001	0,0003
0.2	-0,0017	-0,0002	-0,0001	-0,0005	-0,0004
0.4	0,0003	0,0002	-0,0006	0,0003	-0,0002
0.6	0,0012	-0,0005	-0,0003	0,0001	-0,0003
0.8	0,0025	-0,0015	-0,0016	-0,0017	-0,0015

β1-SAR-Poisson 1stStep-OLS

W1

Rho/n	100	250	500	750	1000
0.0	-0,0011	-0,0010	-0,0006	-0,0003	-0,0003
0.2	-0,0047	-0,0041	-0,0044	0,0040	-0,0036
0.4	-0,0032	-0,0047	-0,0046	-0,0044	-0,0046
0.6	-0,0013	-0,0019	-0,0020	-0,0022	-0,0020
0.8	0,0021	0,0014	0,0017	0,0005	0,0004

W2

Rho/n	100	250	500	750	1000
0.0	-0,0001	-0,0010	0,0001	-0,0002	0,0002
0.2	-0,0019	-0,0007	-0,0004	-0,0006	-0,0005
0.4	-0,0007	-0,0002	-0,0008	0,0001	-0,0003
0.6	0,0008	-0,0004	-0,0002	0,0002	-0,0001
0.8	0,0003	0,0002	0,0000	0,0000	0,0001

β1-SAR-LogLinear

W1

Rho/n	100	250	500	750	1000
0.0	-0,0733	-0,0754	-0,0752	-0,0748	-0,0745
0.2	-0,0538	-0,0555	-0,0555	-0,0556	-0,0550
0.4	-0,0295	-0,0291	-0,0298	-0,0300	-0,0295
0.6	0,0008	0,0005	-0,0005	-0,0006	-0,0001
0.8	0,0054	0,0057	0,0064	0,0070	0,0060

W2

Rho/n	100	250	500	750	1000
0.0	-0,0758	-0,0784	-0,0787	-0,0784	-0,0782
0.2	-0,0549	-0,0568	-0,0569	-0,0560	-0,0562
0.4	-0,0260	-0,0260	-0,0271	-0,0258	-0,0260
0.6	0,0064	0,0094	0,0101	0,0099	0,0109
0.8	0,0043	0,0059	0,0065	0,0078	0,0082

β1-Aspatial Poisson ML

W1

Rho/n	100	250	500	750	1000
0.0	0,0003	-0,0004	-0,0005	0,0000	0,0000
0.2	0,0311	0,0312	0,0312	0,0310	0,0308
0.4	0,0891	0,0868	0,0872	0,0872	0,0870
0.6	0,2021	0,1979	0,1988	0,1969	0,1971
0.8	0,4957	0,4846	0,4886	0,4771	0,4861

W2

Rho/n	100	250	500	750	1000
0.0	-0,0008	-0,0004	0,0001	0,0000	0,0001
0.2	0,0290	0,0291	0,0285	0,0285	0,0286
0.4	0,0769	0,0761	0,0754	0,0750	0,0740
0.6	0,1711	0,1662	0,1609	0,1609	0,1592
0.8	0,4010	0,3873	0,3743	0,3630	0,3577

β2-SAR-Poisson 1stStep-ML

W1

Rho/n	100	250	500	750	1000
0.0	-0,0062	-0,0028	0,0026	0,0004	0,0003
0.2	-0,0016	0,0012	0,0014	-0,0004	0,0003
0.4	0,0024	0,0009	0,0031	-0,0036	0,0043
0.6	0,0017	0,0023	0,0040	-0,0001	0,0027
0.8	-0,0015	-0,0022	-0,0058	0,0077	-0,0078

W2

Rho/n	100	250	500	750	1000
0.0	-0,0012	0,0010	-0,0027	0,0005	-0,0004
0.2	-0,0038	-0,0021	-0,0011	-0,0003	0,0001
0.4	0,0031	0,0020	-0,0003	-0,0003	0,0008
0.6	0,0045	0,0018	0,0013	-0,0007	0,0005
0.8	0,0397	0,0172	0,0109	-0,0069	0,0081

β2-SAR-Poisson 1stStep-OLS

W1

Rho/n	100	250	500	750	1000
0.0	-0,0105	-0,0021	-0,0010	-0,0015	-0,0015
0.2	-0,0132	-0,0118	-0,0156	-0,0138	-0,0130
0.4	-0,0143	-0,0182	-0,0188	-0,0191	-0,0179
0.6	-0,0052	-0,0051	-0,0097	-0,0094	-0,0098
0.8	0,0072	0,0066	0,0065	0,0017	0,0017

W2

Rho/n	100	250	500	750	1000
0.0	-0,0005	0,0010	-0,0029	-0,0009	-0,0005
0.2	-0,0058	-0,0035	-0,0019	-0,0003	-0,0003
0.4	-0,0006	-0,0002	-0,0012	-0,0003	0,0002
0.6	-0,0007	0,0010	0,0015	0,0009	0,0008
0.8	0,0015	0,0006	-0,0001	0,0000	0,0004

β2-SAR-LogLinear

W1

Rho/n	100	250	500	750	1000
0.0	-0,0535	-0,0501	-0,0482	-0,0458	-0,0462
0.2	-0,0340	-0,0398	-0,0346	-0,0358	-0,0374
0.4	-0,0181	-0,0188	-0,0192	-0,0188	-0,0175
0.6	0,0402	0,0247	0,0235	0,0166	0,0183
0.8	0,3381	0,3131	0,3028	0,3295	0,3058

W2

Rho/n	100	250	500	750	1000
0.0	-0,0785	-0,0767	-0,0777	-0,0756	-0,0767
0.2	-0,0520	-0,0542	-0,0539	-0,0539	-0,0550
0.4	-0,0219	0,0241	-0,0235	-0,0238	-0,0227
0.6	0,0143	0,0111	0,0112	0,0123	0,0130
0.8	0,0364	0,0227	0,0175	0,0148	0,0163

β2-Aspatial Poisson ML

W1

Rho/n	100	250	500	750	1000
0.0	-0,0064	-0,0004	-0,0003	-0,0004	-0,0007
0.2	0,1018	0,1035	0,1028	0,1027	0,1036
0.4	0,2737	0,2768	0,2760	0,2789	0,2799
0.6	0,6201	0,6327	0,6391	0,6363	0,6383
0.8	1,7530	1,7635	1,7964	1,8182	1,8423

W2

Rho/n	100	250	500	750	1000
0.0	-0,0014	-0,0011	0,0004	-0,0005	0,0000
0.2	0,0991	0,0998	0,1022	0,1009	0,1014
0.4	0,2671	0,2663	0,2682	0,2670	0,2693
0.6	0,5905	0,6000	0,6034	0,6027	0,6015
0.8	1,6308	1,6558	1,6638	1,6922	1,7034

Notes:

1) Bias is estimated as $\hat{\beta} - \beta_0$, the difference between the paramater estimate and its true value. Entries are calculated as the average of 1000 simulations.
2) SAR-Poisson 1stStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1 and β_2 are estimated using a Poisson regression.
3) SAR-Poisson 1stStep-OLS is estimated using a two-step process. In the first step, the unobservable variable Wlog(μ_j) is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_j=0, and in the second step, the coefficients β_1 and β_2 are estimated using a Poisson regression.
4) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1 and β_2 are estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
5) Dark shaded entries denote cases where the Bias of the SAR-Poisson 1stStep-ML were smaller than the SAR-Poisson 1stStep-OLS.
6) Bright shaded entries denote cases where the Bias of SAR-LogLinear or aspatial ML Poisson were smaller or equal than the SAR-Poisson estimators.
7) W_1 is a continguity matrix created using the nearest neighbour criterion, where it is computationally defined that each unit i will have **seven** units j as neighbors, these being the seven units j closest to i. W2 is created based on an inverse distance criterion, using the Euclidean distance between unit i and unit j, with i,j=1,2…N.

Table A2:
Bias: SAR-Poisson, SAR-LogLinear estimates: β_1, β_2

Rho-SAR-Poisson 1stStep-ML

W1

Rho/n	100	250	500	750	1000
0.0	0,0042	0,0015	0,0005	-0,0012	-0,0008
0.2	0,0012	-0,0025	0,0006	-0,0003	0,0003
0.4	-0,0036	-0,0014	-0,0012	-0,0023	-0,0027
0.6	-0,0002	-0,0004	-0,0018	-0,0016	-0,0010
0.8	0,0047	0,0039	0,0048	0,0050	0,0050

W2

Rho/n	100	250	500	750	1000
0.0	-0,0152	-0,0020	0,0001	-0,0010	-0,0017
0.2	-0,0033	0,0006	0,0007	0,0009	0,0015
0.4	-0,0038	-0,0009	0,0024	0,0002	0,0007
0.6	0,0124	0,0072	0,0040	0,0021	0,0026
0.8	0,0501	0,0341	0,0236	0,0166	0,0143

Rho-SAR-Poisson 1stStep-OLS

W1

Rho/n	100	250	500	750	1000
0.0	0,0037	0,0018	-0,0003	0,0018	0,0014
0.2	0,0079	0,0105	0,0149	0,0131	0,0119
0.4	0,0169	0,0234	0,0246	0,0247	0,0249
0.6	0,0160	0,0181	0,0204	0,0208	0,0209
0.8	0,0018	0,0034	0,0042	0,0061	0,0082

W2

Rho/n	100	250	500	750	1000
0.0	-0,0082	0,0007	0,0018	0,0010	-0,0006
0.2	-0,0122	-0,0108	-0,0115	-0,0121	-0,0119
0.4	-0,0064	-0,0030	0,0000	-0,0026	-0,0016
0.6	0,0126	0,0144	0,0152	0,0152	0,0154
0.8	0,0047	0,0044	0,0045	0,0045	0,0043

Rho-SAR-LogLinear

W1

Rho/n	100	250	500	750	1000
0.0	0,0768	0,0779	0,0772	0,0751	0,0749
0.2	0,0151	0,0207	0,0199	0,0216	0,0228
0.4	-0,0189	-0,0205	-0,0199	-0,0196	-0,0206
0.6	-0,0199	-0,0216	-0,0228	-0,0233	-0,0241
0.8	0,0595	0,0501	0,0513	0,0471	0,0523

W2

Rho/n	100	250	500	750	1000
0.0	0,1133	0,1164	0,1197	0,1179	0,1184
0.2	0,0294	0,0351	0,0349	0,0339	0,0350
0.4	-0,0285	-0,0301	-0,0285	-0,0300	-0,0302
0.6	-0,0378	-0,0445	-0,0488	-0,0493	-0,0515
0.8	0,0980	0,0905	0,0638	0,0555	0,0667

Notes:

1) Bias is estimated as $\hat{\rho} - \rho_0$, the difference between the paramater estimate and its true value. Entries are calculated as the average of 1000 simulations;
2) SAR-Poisson 1ST Step-ML is estimated using a two-step process. In the first step, the unobservable variable μ_i is estimated using a Poisson regression, and in the second step, the coefficients β_1 and β_2 are estimated using a poisson regression.
3) SAR-Poisson 1ST Step-OLS is estimated using a two-step process. In the first step, the unobservable variable Wlog(μ_i) is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_i=0, and in the second step, the coefficients β_1 and β_2 are estimated using a Poisson regression.
4) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_i is estimated using a Poisson regression, and in the second step, the coefficients β_1 and β_2 are estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
5) Dark shaded entries denote cases where the Bias of the SAR-Poisson 1ST Step-ML were smaller than the SAR-Poisson 1ST Step-OLS
6) Bright shaded entries denote cases where the Bias of SAR-LogLinear or aspatial ML Poissn were smaller or equal than the SAR-Poisson estimators
7) W_1 is a continguity matrix created using the nearest neighbour criterion, where it is computationally defined that each unit i will have **seven** units j as neighbors, these being the seven units j closest to i. W2 is created based on an inverse distance criterion, using the Euclidean distance between unit i and unit j, with i,j=1,2…N.

Table A3:
RMSE: SAR-Poisson and SAR-LogLinear estimates: ρ

Rho-SAR-Poisson 1stStep-ML

W1

Rho/n	100	250	500	750	1000
0.0	0,1112	0,0676	0,0474	0,0383	0,0320
0.2	0,0792	0,0483	0,0330	0,0276	0,0239
0.4	0,0512	0,0289	0,0214	0,0173	0,0149
0.6	0,0237	0,0148	0,0122	0,0110	0,0097
0.8	0,0142	0,0118	0,0106	0,0097	0,0094

W2

Rho/n	100	250	500	750	1000
0.0	0,2196	0,1283	0,0934	0,0803	0,0687
0.2	0,1419	0,0860	0,0636	0,0484	0,0420
0.4	0,0814	0,0513	0,0360	0,0287	0,0245
0.6	0,0571	0,0377	0,0264	0,0187	0,0157
0.8	0,0840	0,0493	0,0290	0,0276	0,0245

Rho-SAR-Poisson 1stStep-OLS

W1

Rho/n	100	250	500	750	1000
0.0	0,1191	0,0734	0,0532	0,0398	0,0357
0.2	0,0873	0,0542	0,0395	0,0318	0,0279
0.4	0,0560	0,0390	0,0332	0,0307	0,0294
0.6	0,0311	0,0244	0,0237	0,0230	0,0224
0.8	0,0152	0,0146	0,0131	0,0137	0,0162

W2

Rho/n	100	250	500	750	1000
0.0	0,1933	0,1123	0,0815	0,0695	0,0597
0.2	0,1388	0,0840	0,0634	0,0485	0,0426
0.4	0,0875	0,0546	0,0387	0,0310	0,0256
0.6	0,0423	0,0275	0,0216	0,0196	0,0184
0.8	0,0096	0,0062	0,0054	0,0052	0,0047

Rho-SAR-LogLinear

W1

Rho/n	100	250	500	750	1000
0.0	0,1516	0,1117	0,0934	0,0850	0,0825
0.2	0,1180	0,0731	0,0544	0,0467	0,0425
0.4	0,0845	0,0594	0,0415	0,0376	0,0343
0.6	0,0651	0,0442	0,0376	0,0326	0,0305
0.8	0,1071	0,0800	0,0679	0,0626	0,0663

W2

Rho/n	100	250	500	750	1000
0.0	0,1953	0,1516	0,1365	0,1272	0,1260
0.2	0,1350	0,0912	0,0654	0,0585	0,0523
0.4	0,0957	0,0665	0,0495	0,0439	0,0408
0.6	0,0796	0,0615	0,0561	0,0559	0,0545
0.8	0,2070	0,1897	0,1459	0,1305	0,1546

Notes:

1) RMSE is estimated as $\sqrt{Bias^2 + Var(\widehat{\rho})}$ where $Bias_\rho^2$ is the square of the averaged bias ρ calculated after 1000 replications and $Var(\widehat{\rho})$ is the empirical variance of the estimated coefficient.
2) SAR-Poisson 1STStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_i is estimated using a Poisson regression, and in the second step, the coefficients ρ is estimated using a poisson regression.
3) SAR-Poisson 1STStep-OLS is estimated using a two-step process. In the first step, the unobservable variable Wlog(μ_i) is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_i=0, and in the second step, the coefficients ρ is estimated using a Poisson regression.
4) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_i is estimated using a Poisson regression, and in the second step, the coefficients ρ is estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
5) Dark shaded entries denote cases where the Bias of the SAR-Poisson 1stStep-ML were smaller than the SAR-Poisson 1STStep-OLS
6) Bright shaded entries denote cases where the Bias of SAR-LogLinear or aspatial ML Poisson were smaller or equal than the SAR-Poisson estimators
7) W_1 is a continguity matrix created using the nearest neighbour criterion, where it is computationally defined that each unit i will have **seven** units j as neighbors, these being the seven units j closest to i. W2 is created based on an inverse distance criterion, using the Euclidean distance between unit i and unit j, with i,j=1,2…N.

Table A4:
RMSE: SAR-Poisson, SAR-LogLinear and Aspatial ML Poisson estimates: β_1, β_2

β1-SAR-Poisson 1stStep-ML					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0242	0,0154	0,0102	0,0083	0,0069
0.2	0,0251	0,0152	0,0107	0,0087	0,0071
0.4	0,0223	0,0137	0,0102	0,0077	0,0070
0.6	0,0176	0,0106	0,0080	0,0065	0,0061
0.8	0,0161	0,0134	0,0124	0,0113	0,0106
W2					
Rho/n	100	250	500	750	1000
0.0	0,0387	0,0232	0,0165	0,0142	0,0114
0.2	0,0363	0,0207	0,0151	0,0121	0,0110
0.4	0,0304	0,0187	0,0136	0,0103	0,0088
0.6	0,0241	0,0134	0,0093	0,0076	0,0066
0.8	0,0305	0,0156	0,0123	0,0102	0,0092

β1-SAR-Poisson 1stStep-OLS					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0270	0,0163	0,0117	0,0090	0,0077
0.2	0,0271	0,0165	0,0119	0,0092	0,0083
0.4	0,0228	0,0141	0,0106	0,0089	0,0080
0.6	0,0181	0,0113	0,0079	0,0065	0,0056
0.8	0,0174	0,0155	0,0123	0,0120	0,0118
W2					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0402	0,0236	0,0167	0,0143	0,0114
0.2	0,0360	0,0207	0,0151	0,0121	0,0110
0.4	0,0300	0,0186	0,0136	0,0103	0,0088
0.6	0,0233	0,0134	0,0093	0,0076	0,0065
0.8	0,0100	0,0056	0,0038	0,0031	0,0026

β1-SAR-LogLinear					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0798	0,0777	0,0764	0,0756	0,0751
0.2	0,0655	0,0593	0,0572	0,0567	0,0559
0.4	0,0487	0,0382	0,0337	0,0324	0,0312
0.6	0,0401	0,0263	0,0184	0,0141	0,0128
0.8	0,0425	0,0311	0,0256	0,0226	0,0209
W2					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0822	0,0807	0,0799	0,0792	0,0787
0.2	0,0662	0,0602	0,0585	0,0572	0,0569
0.4	0,0470	0,0354	0,0313	0,0285	0,0280
0.6	0,0378	0,0247	0,0181	0,0158	0,0153
0.8	0,0305	0,0220	0,0159	0,0142	0,0134

β1-Aspatial Poisson ML					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0171	0,0174	0,0122	0,0099	0,0084
0.2	0,0424	0,0353	0,0330	0,0321	0,0316
0.4	0,0946	0,0893	0,0885	0,0880	0,0877
0.6	0,2122	0,2027	0,2017	0,1989	0,1987
0.8	0,5380	0,5063	0,5029	0,4887	0,4954
W2					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0276	0,0178	0,0120	0,0095	0,0082
0.2	0,0386	0,0327	0,0300	0,0296	0,0294
0.4	0,0820	0,0781	0,0763	0,0757	0,0745
0.6	0,1784	0,1694	0,1625	0,1621	0,1603
0.8	0,4256	0,4019	0,3816	0,3692	0,3626

β2-SAR-Poisson 1stStep-ML					
W1					
Rho/n	100	250	500	750	1000
0.0	0,1034	0,0649	0,0454	0,0357	0,0314
0.2	0,0912	0,0567	0,0401	0,0333	0,0287
0.4	0,0811	0,0486	0,0353	0,0300	0,0261
0.6	0,0606	0,0395	0,0305	0,0277	0,0238
0.8	0,0602	0,0548	0,0558	0,0502	0,0476
W2					
Rho/n	100	250	500	750	1000
0.0	0,1487	0,0914	0,0669	0,0541	0,0488
0.2	0,1368	0,0843	0,0624	0,0495	0,0421
0.4	0,1182	0,0758	0,0515	0,0412	0,0362
0.6	0,0902	0,0527	0,0387	0,0302	0,0274
0.8	0,1310	0,0613	0,0402	0,0270	0,0277

β2-SAR-Poisson 1stStep-OLS					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,1085	0,0648	0,0495	0,0382	0,0330
0.2	0,0999	0,0631	0,0437	0,0357	0,0316
0.4	0,0849	0,0531	0,0411	0,0354	0,0309
0.6	0,0643	0,0403	0,0292	0,0242	0,0215
0.8	0,0589	0,0487	0,0427	0,0404	0,0428
W2					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,1520	0,0924	0,0674	0,0543	0,0489
0.2	0,1376	0,0845	0,0625	0,0496	0,0422
0.4	0,1186	0,0761	0,0515	0,0411	0,0361
0.6	0,0894	0,0529	0,0385	0,0300	0,0273
0.8	0,0376	0,0216	0,0156	0,0126	0,0105

β2-SAR-LogLinear					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,1572	0,1086	0,0812	0,0695	0,0629
0.2	0,1498	0,1016	0,0738	0,0647	0,0602
0.4	0,1502	0,0959	0,0692	0,0634	0,0540
0.6	0,1590	0,1091	0,0882	0,0699	0,0625
0.8	0,5243	0,4711	0,4432	0,4651	0,4352
W2					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,1783	0,1277	0,1037	0,0930	0,0883
0.2	0,1680	0,1163	0,0853	0,0774	0,0716
0.4	0,1585	0,0989	0,0733	0,0622	0,0526
0.6	0,1388	0,0856	0,0601	0,0505	0,0443
0.8	0,1082	0,0615	0,0381	0,0304	0,0306

β2-Aspatial Poisson ML					
W1					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0716	0,0752	0,0564	0,0449	0,0382
0.2	0,1588	0,1263	0,1132	0,1089	0,1089
0.4	0,2967	0,2854	0,2800	0,2819	0,2821
0.6	0,6483	0,6454	0,6465	0,6423	0,6432
0.8	1,9053	1,8426	1,8515	1,8604	1,8784
W2					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,1224	0,0797	0,0551	0,0436	0,0394
0.2	0,1512	0,1238	0,1109	0,1095	0,1075
0.4	0,2866	0,2732	0,2714	0,2691	0,2709
0.6	0,6054	0,6064	0,6061	0,6050	0,6043
0.8	1,6870	1,6898	1,6781	1,7039	1,7125

Notes:

1) RMSE is estimated as $\sqrt{Bias_{\beta_j}^2 + Var(\widehat{\beta_j})}$, where $Bias_{\beta_j}^2$ is the square of the averaged bias of β_j calculated after 1000 replications, where j=1,2. $Var(\widehat{\beta_j})$ is the empirical variance of the estimated coefficient.
2) SAR-Poisson 1stStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1 and β_2 are estimated using a Poisson regression.
3) SAR-Poisson 1stStep-OLS is estimated using a two-step process. In the first step, the unobservable variable Wlog(μ_j) is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_j=0, and in the second step, the coefficients β_1 and β_2 are estimated using a Poisson regression.
4) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1 and β_2 are estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
5) Dark shaded entries denote cases where the Bias of the SAR-Poisson 1stStep-ML were smaller than the SAR-Poisson 1stStep-OLS
6) Bright shaded entries denote cases where the Bias of SAR-LogLinear or aspatial ML Poisson were smaller or equal than the SAR-Poisson estimators
7) W_1 is a continguity matrix created using the nearest neighbour criterion, where it is computationally defined that each unit i will have **seven** units j as neighbors, these being the seven units j closest to i. W2 is created based on an inverse distance criterion, using the Euclidean distance between unit i and unit j, with i,j=1,2…N.

Table A5:
Bias: SAR-Poisson, SAR-LogLinear and Aspatial ML Poisson estimates: β_1, β_2 and ρ

β1-SAR-Poisson 1stStep-ML

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0013	-0,0009	-0,0006	0,0000	0,0000
0.2	-0,0002	-0,0003	0,0001	-0,0003	0,0002
0.4	0,0018	0,0031	0,0030	0,0033	0,0031
0.6	0,0053	0,0055	0,0049	0,0045	0,0076
0.8	0,0333	0,0256	0,0271	0,0269	0,0258

β1-SAR-Poisson 1stStep-OLS

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0015	-0,0004	-0,0002	-0,0007	0,0002
0.2	-0,0068	-0,0057	-0,0049	-0,0051	-0,0049
0.4	-0,0074	-0,0067	-0,0061	-0,0066	-0,0059
0.6	-0,0043	-0,0036	-0,0040	-0,0047	-0,0041
0.8	0,0024	0,0008	0,0008	-0,0021	-0,0002

β1-SAR-LogLinear

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0733	-0,0744	-0,0745	-0,0750	-0,0746
0.2	-0,0546	-0,0541	-0,0556	-0,0552	-0,0555
0.4	-0,0297	-0,0293	-0,0293	-0,0296	-0,0294
0.6	-0,0003	0,0007	-0,0008	-0,0006	-0,0007
0.8	0,0051	0,0058	0,0056	0,0044	0,0071

β1-Aspatial Poisson ML

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0008	-0,0004	-0,0004	0,0000	0,0001
0.2	0,0334	0,0320	0,0328	0,0327	0,0327
0.4	0,0945	0,0955	0,0961	0,0950	0,0954
0.6	0,2162	0,2243	0,2249	0,2249	0,2227
0.8	0,5565	0,5744	0,5710	0,5781	0,5701

β2-SAR-Poisson 1stStep-ML

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0016	-0,0020	0,0028	-0,0006	0,0001
0.2	0,0020	0,0014	0,0053	0,0043	0,0041
0.4	0,0175	0,0206	0,0198	0,0221	0,0212
0.6	0,0382	0,0308	0,0296	0,0263	0,0248
0.8	0,0978	0,0919	0,0982	0,0490	0,0311

β2-SAR-Poisson 1stStep-OLS

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0069	-0,0032	-0,0038	-0,0011	-0,0013
0.2	-0,0173	-0,0166	-0,0191	-0,0190	-0,0184
0.4	-0,0227	-0,0236	-0,0241	-0,0259	-0,0267
0.6	-0,0144	-0,0147	-0,0169	-0,0191	-0,0196
0.8	0,0067	0,0086	-0,0031	-0,0136	-0,0153

β2-SAR-LogLinear

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0535	-0,0491	-0,0421	-0,0449	-0,0445
0.2	-0,0344	-0,0374	-0,0366	-0,0343	-0,0365
0.4	-0,0184	-0,0182	-0,0182	-0,0180	-0,0170
0.6	0,0383	0,0236	0,0200	0,0198	0,0160
0.8	0,3167	0,3313	0,3078	0,3140	0,3109

β2-Aspatial Poisson ML

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	-0,0014	-0,0011	0,0034	-0,0005	0,0000
0.2	0,1017	0,1032	0,1052	0,1047	0,1046
0.4	0,2822	0,2859	0,2864	0,2869	0,2874
0.6	0,6469	0,6572	0,6612	0,6640	0,6653
0.8	1,8195	1,9117	1,9474	1,9597	1,9623

Rho-SAR-Poisson 1stStep-ML

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0006	0,0000	0,0002	-0,0013	-0,0002
0.2	-0,0008	0,0015	-0,0006	0,0007	0,0002
0.4	-0,0050	-0,0086	-0,0081	-0,0088	-0,0090
0.6	0,0002	-0,0041	-0,0048	-0,0059	-0,0114
0.8	0,0118	0,0060	0,0003	-0,0034	-0,0015

Rho-SAR-Poisson 1stStep-OLS

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0043	0,0006	0,0018	0,0015	-0,0005
0.2	0,0193	0,0199	0,0215	0,0227	0,0219
0.4	0,0303	0,0329	0,0336	0,0358	0,0346
0.6	0,0236	0,0247	0,0275	0,0296	0,0288
0.8	0,0114	0,0112	0,0158	0,0239	0,0231

Rho-SAR-LogLinear

W3					
Rho/n	**100**	**250**	**500**	**750**	**1000**
0.0	0,0768	0,0761	0,0739	0,0740	0,0740
0.2	0,0175	0,0184	0,0220	0,0216	0,0231
0.4	-0,0185	-0,0204	-0,0199	-0,0197	-0,0205
0.6	-0,0190	-0,0228	-0,0235	-0,0245	-0,0230
0.8	0,0664	0,0507	0,0506	0,0491	0,0510

Notes:

1) Bias is estimated as the difference between the paramater estimate and its true value. Entries are calculated as the average of 1000 simulations;
2) SAR-Poisson 1stStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1, β_2, and ρ are estimated using a poisson regression.
3) SAR-Poisson 1stStep-OLS is estimated using a two-step process. In the first step, the unobservable variable $Wlog(\mu_j)$ is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_j=0, and in the second step, the coefficients β_1, β_2 and ρ are estimated using a poisson regression.
4) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1,β_2 and ρ are estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
5) W_3 is a continguity matrix created using the nearest neighbour criterion, where it is computationally defined that each unit i will have **four** units j as neighbors, these being the four units j closest to i.

Table A6:
RMSE: SAR-Poisson, SAR-LogLinear and Aspatial ML Poisson estimates: β_1, β_2 and ρ

β1-SAR-Poisson 1stStep-ML					
W3					
Rho/n	100	250	500	750	1000
0.0	0,0315	0,0198	0,0134	0,0109	0,0090
0.2	0,0326	0,0198	0,0133	0,0117	0,0099
0.4	0,0316	0,0197	0,0139	0,0118	0,0107
0.6	0,0271	0,0174	0,0125	0,0103	0,0134
0.8	0,0655	0,0502	0,0472	0,0455	0,0439

β1-SAR-Poisson 1stStep-OLS					
W3					
Rho/n	100	250	500	750	1000
0.0	0,0358	0,0212	0,0148	0,0119	0,0100
0.2	0,0356	0,0212	0,0147	0,0118	0,0109
0.4	0,0324	0,0199	0,0144	0,0119	0,0106
0.6	0,0266	0,0160	0,0119	0,0104	0,0087
0.8	0,0494	0,0361	0,0327	0,0344	0,0324

β1-SAR-LogLinear					
W3					
Rho/n	100	250	500	750	1000
0.0	0,0798	0,0767	0,0757	0,0758	0,0751
0.2	0,0646	0,0578	0,0573	0,0563	0,0563
0.4	0,0490	0,0378	0,0332	0,0321	0,0314
0.6	0,0383	0,0258	0,0176	0,0141	0,0120
0.8	0,0444	0,0319	0,0239	0,0218	0,0227

β1-Aspatial Poisson ML					
W3					
Rho/n	100	250	500	750	1000
0.0	0,0276	0,0178	0,0119	0,0095	0,0082
0.2	0,0427	0,0367	0,0345	0,0339	0,0336
0.4	0,1015	0,0983	0,0976	0,0959	0,0962
0.6	0,2288	0,2300	0,2284	0,2275	0,2246
0.8	0,6065	0,6061	0,5963	0,5972	0,5870

β2-SAR-Poisson 1stStep-ML					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1356	0,0855	0,0592	0,0474	0,0413
0.2	0,1226	0,0759	0,0525	0,0437	0,0398
0.4	0,1143	0,0788	0,0645	0,0619	0,0558
0.6	0,1184	0,0851	0,0721	0,0582	0,0311
0.8	0,2134	0,1867	0,1864	0,1999	0,1949

β2-SAR-Poisson 1stStep-OLS					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1408	0,0865	0,0648	0,0509	0,0433
0.2	0,1295	0,0801	0,0573	0,0458	0,0423
0.4	0,1152	0,0718	0,0521	0,0456	0,0425
0.6	0,0898	0,0571	0,0394	0,0360	0,0341
0.8	0,1701	0,1297	0,1155	0,1217	0,1164

β2-SAR-LogLinear					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1572	0,1055	0,0754	0,0682	0,0630
0.2	0,1510	0,1002	0,0744	0,0627	0,0603
0.4	0,1502	0,0949	0,0722	0,0619	0,0558
0.6	0,1605	0,1066	0,0829	0,0722	0,0627
0.8	0,5046	0,4894	0,4407	0,4563	0,4405

β2-Aspatial Poisson ML					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1224	0,0797	0,0540	0,0436	0,0394
0.2	0,1544	0,1242	0,1146	0,1126	0,1109
0.4	0,3083	0,2969	0,2915	0,2902	0,2899
0.6	0,6891	0,6764	0,6717	0,6715	0,6705
0.8	2,0317	2,0655	2,1007	2,0681	2,0419

Rho-SAR-Poisson 1stStep-ML					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1272	0,0784	0,0547	0,0436	0,0368
0.2	0,0983	0,0554	0,0380	0,0320	0,0274
0.4	0,0641	0,0377	0,0294	0,0268	0,0244
0.6	0,0440	0,0308	0,0254	0,0203	0,0248
0.8	0,0774	0,0542	0,0431	0,0417	0,0416

Rho-SAR-Poisson 1stStep-OLS					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1414	0,0885	0,0621	0,0478	0,0441
0.2	0,1116	0,0647	0,0482	0,0400	0,0368
0.4	0,0717	0,0501	0,0419	0,0404	0,0382
0.6	0,0422	0,0324	0,0308	0,0322	0,0307
0.8	0,0429	0,036563	0,0342	0,0400	0,0387

Rho-SAR-LogLinear					
W3					
Rho/n	100	250	500	750	1000
0.0	0,1516	0,1101	0,0926	0,0845	0,0819
0.2	0,1143	0,0734	0,0549	0,0470	0,0420
0.4	0,0860	0,0571	0,0439	0,0385	0,0344
0.6	0,0614	0,0430	0,0353	0,0320	0,0294
0.8	0,1112	0,0790	0,0648	0,0639	0,0630

Notes:

1) RMSE for β´s is estimated as $\sqrt{Bias^2_{\beta_j} + Var(\widehat{\beta_j})}$, where $Bias^2_{\beta_j}$ is the square of the averaged bias of β_j calculated after 1000 replications, where j=1,2. $Var(\widehat{\beta_j})$ is the empirical variance of the estimated coefficient. RMSE for ρ is estimated as $\sqrt{Bias^2 + Var(\widehat{\rho})}$ where $Bias^2_{\rho}$ is the square of the averaged bias ρ calculated after 1000 replications and $Var(\widehat{\rho})$ is the empirical variance of the estimated coefficient.
2) SAR-Poisson 1stStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1, β_2, and ρ are estimated using a Poisson regression.
3) SAR-Poisson 1stStep-OLS is estimated using a two-step process. In the first step, the unobservable variable Wlog(μ_j) is estimated using na OLS regression, adding a ad hoc constant (c=1) when y_j=0, and in the second step, the coefficients β_1, β_2 and ρ are estimated using a Poisson regression.
4) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1,β_2 and ρ are estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
5) W_3 is a continguity matrix created using the nearest neighbour criterion, where it is computationally defined that each unit i will have **four** units j as neighbors, these being the four units j closest to i.

Appendix B

Table B1: Variable definitions and expected signal

Variables	Abbreviations	UNIT	Expected signal
The number of patents registered	Pat	Unit per million inhabitants	
Intramural Expenditure on R&D by Private Business	R&D_B	Euro per inhabitant	+
Intramural Expenditure on R&D by the Government	R&D_G	Euro per inhabitant	Ambiguous
Intramural Expenditure on R&D by Universities	R&D_U	Euro per inhabitant	Ambiguous
Total R&D personnel and researchers in Private Business	Pers_B	No. of workers Full-time	+
Total R&D personnel and researchers in the Government	Pers_G	No. of workers Full-time	Ambiguous
Total R&D personnel and researchers in Universities	Pers_U	No. of workers Full-time	Ambiguous
% Population aged 25-64 with bachelor´s degree	Educ	Percentage	+
Population	Pop	Number of inhabitants	-
GDP *per Capita*	GDP	Thousand Euro *per capita*	+
Tuberculosis mortality	Mort	Rate per 100 thousand inhabitants	-

Table B2: Descriptive Statistics of the variables

	Pat	R&D_B	R&D_G	R&D_U	Pers_B	Pers_G	Pers_U	Educ	Pop	GDP	Mort
N	234	234	234	234	234	234	234	234	234	234	234
Mean	89,171	318,248	59,819	135,917	5744,342	1467,979	3294,923	27,334	1982780,504	26,922	1,009
Std Dev	106,045	382,444	87,684	152,388	9291,554	2628,740	3558,347	8,700	1563839,620	13,874	1,375
Max	590,000	2441,700	480,600	891,700	97982,000	17934,000	34836,000	50,100	11898502,000	84,047	8,800
Min	0	0	0	0	0	0	0	11,2	126620	3,561	0,1
1º Quartile	12,000	63,675	7,075	37,425	1181,000	140,250	1025,500	19,900	1073943,500	16,740	0,400
Median	54,500	181,500	24,700	92,300	2969,000	513,000	2238,000	27,400	1575968,000	27,003	0,600
3º Quartile	125,000	418,900	70,125	159,200	7144,750	1585,000	4454,000	33,200	2411857,250	33,876	1,000
% 0´s	6%										

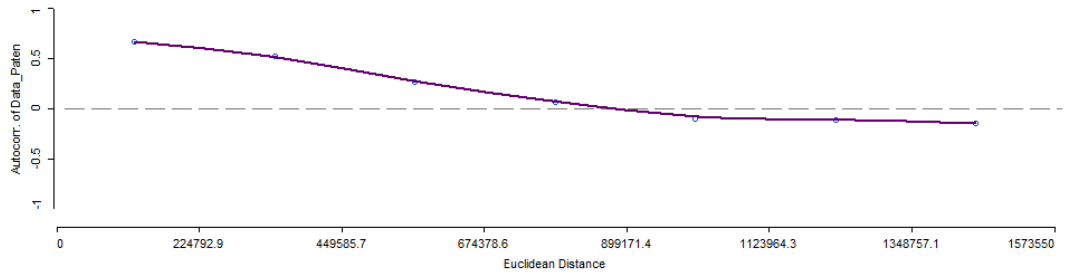
Source: Eurostat, author calculations

Table B3: Correlation Matrix of the variables

	Pat	R&D_B	R&D_G	R&D_U	Pers_B	Pers_G	Pers_U	Educ	Pop	GDP	Mort
Pat	1,000										
R&D_B	0,717	1,000									
R&D_G	0,301	0,438	1,000								
R&D_U	0,390	0,533	0,510	1,000							
Pers_B	0,474	0,601	0,332	0,211	1,000						
Pers_G	0,153	0,215	0,591	0,117	0,622	1,000					
Pers_U	0,164	0,286	0,329	0,261	0,746	0,674	1,000				
Educ	0,263	0,450	0,408	0,440	0,296	0,253	0,355	1,000			
Pop	0,056	0,082	0,119	-0,088	0,663	0,651	0,775	0,010	1,000		
GDP	0,573	0,639	0,519	0,642	0,368	0,148	0,226	0,559	-0,043	1,000	
Mort	-0,285	-0,248	-0,176	-0,266	-0,141	-0,017	-0,099	-0,275	0,082	-0,466	1,000

Source: Eurostat, author calculations

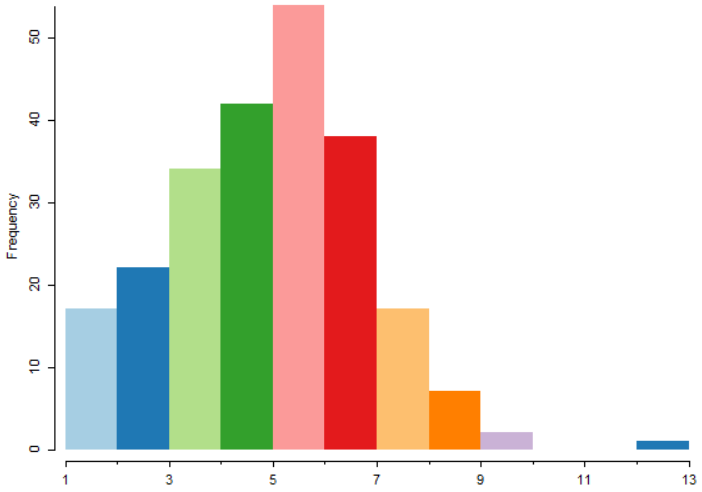
Figure B4: Spatial Distance Correlogram- Variable *Pat*



Note: The value of the spatial autocorrelation is given in order to the Eucledian distance between regions

Source: Eurostat, author calculations; Software: GeoDa

Figure B5: Queen contiguity matrix-histogram of the number of neighbors



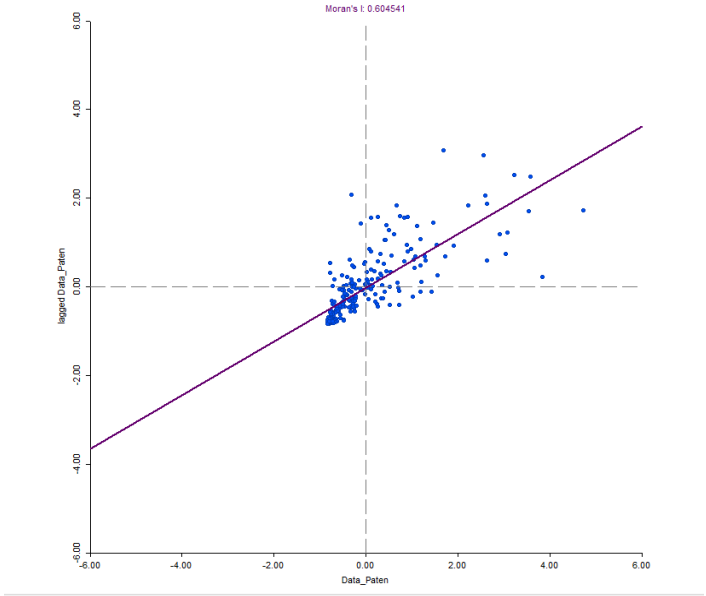
Source: Eurostat, author calculations; Software: GeoDa

Table B6: Moran´s I Test for Spatial Autocorrelation

Moran's I Test For Spatial Autocorrelation		
Matrix type	Queen	Euclidean Inverse Distance
Moran's I Test Statistic	0.6045	0.2999
P-Value	0.0000	0.0000

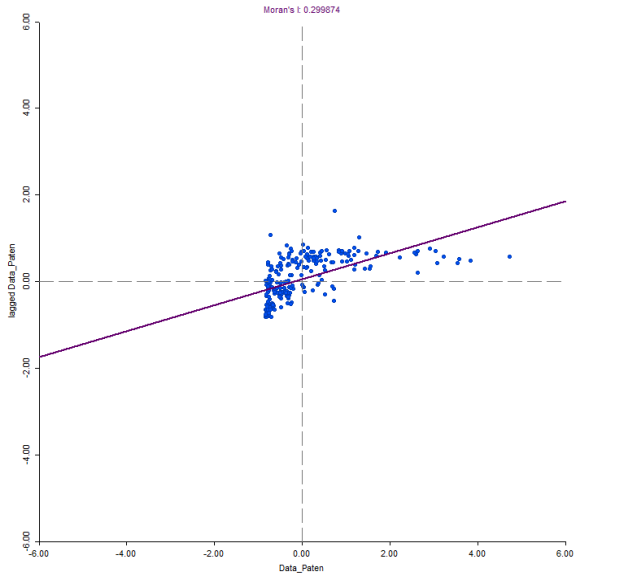
Source: author calculations; Software: R

Figure B7: Moran diagrams for variable *Pat*, Queen matrix- year 2012



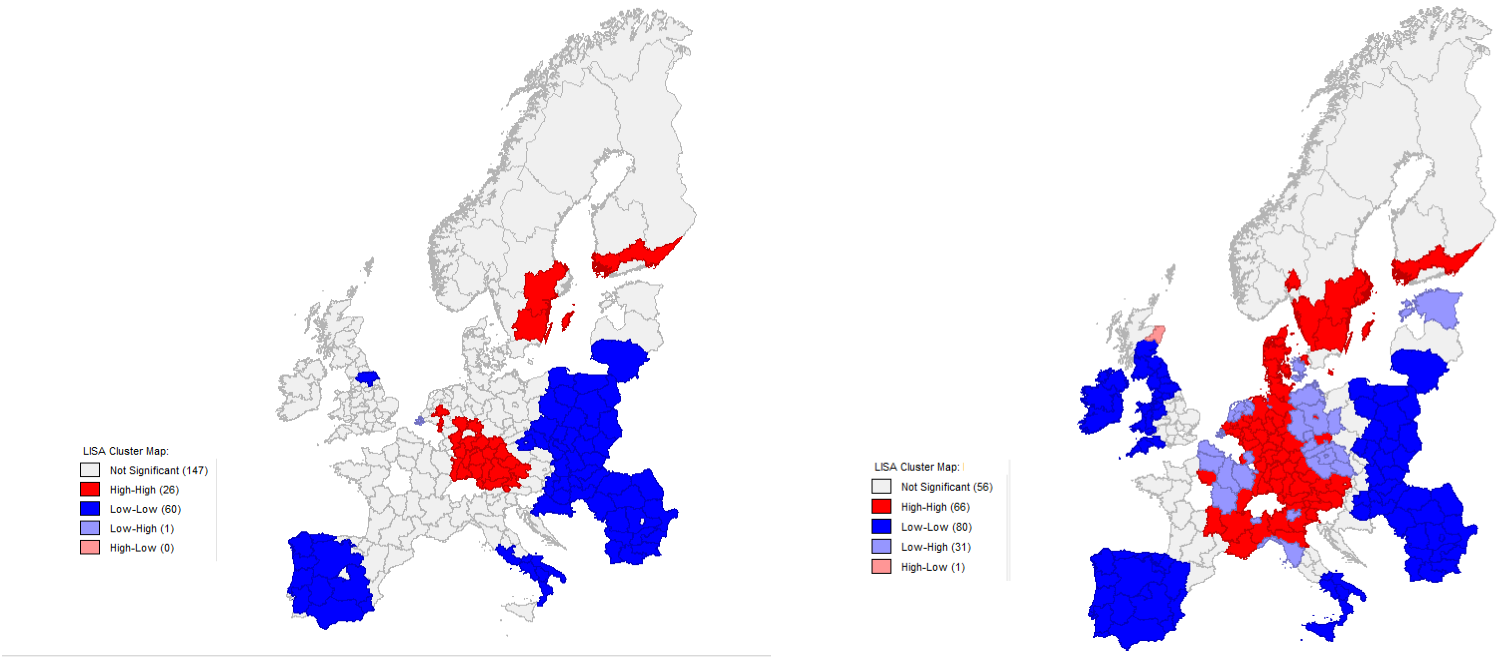
Source: Eurostat, author calculations; Software: GeoDa

Figure B8: Moran diagrams for variable *Pat*, EID matrix- year 2012



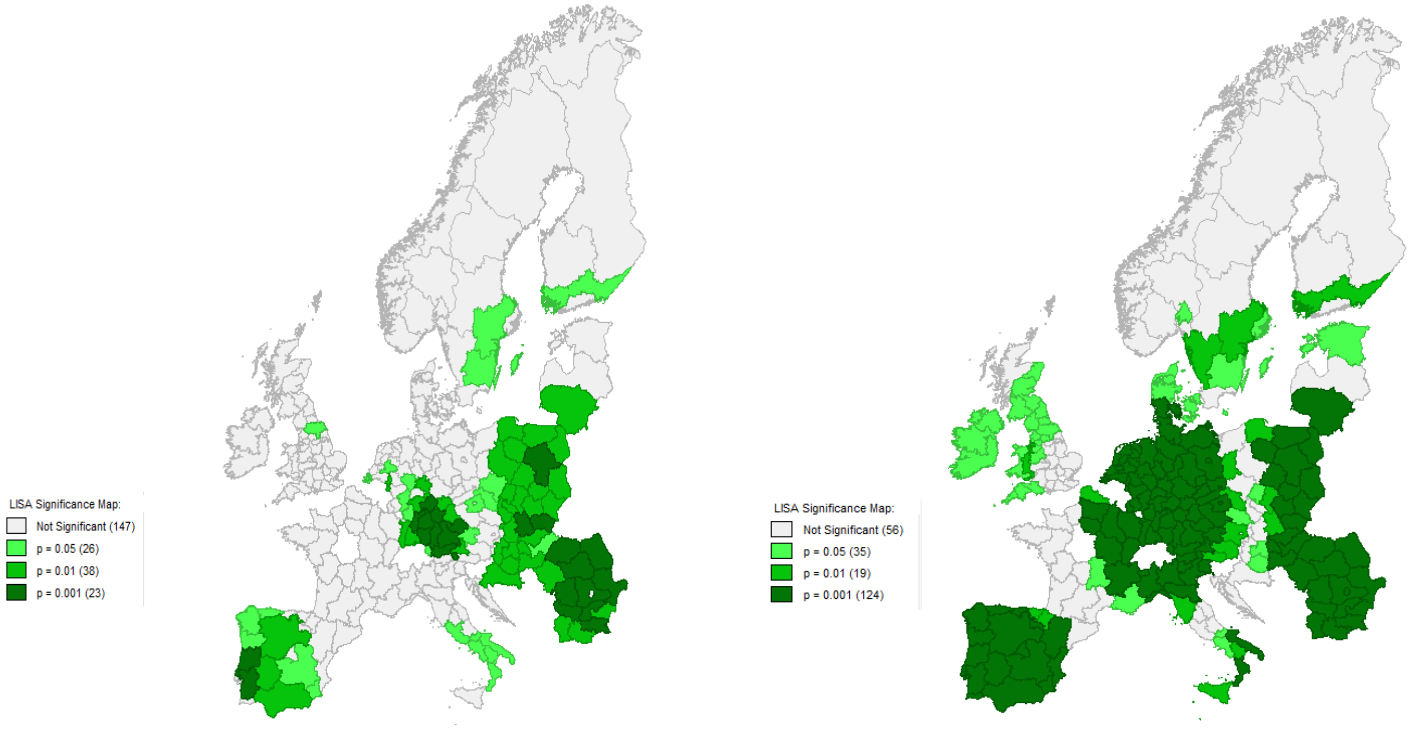
Source: Eurostat, author calculations; Software: GeoDa

Figure B9: Local Indicators of Spatial Association for variable *Pat*, Queen (Left) and EID (Right) matrix- year 2012



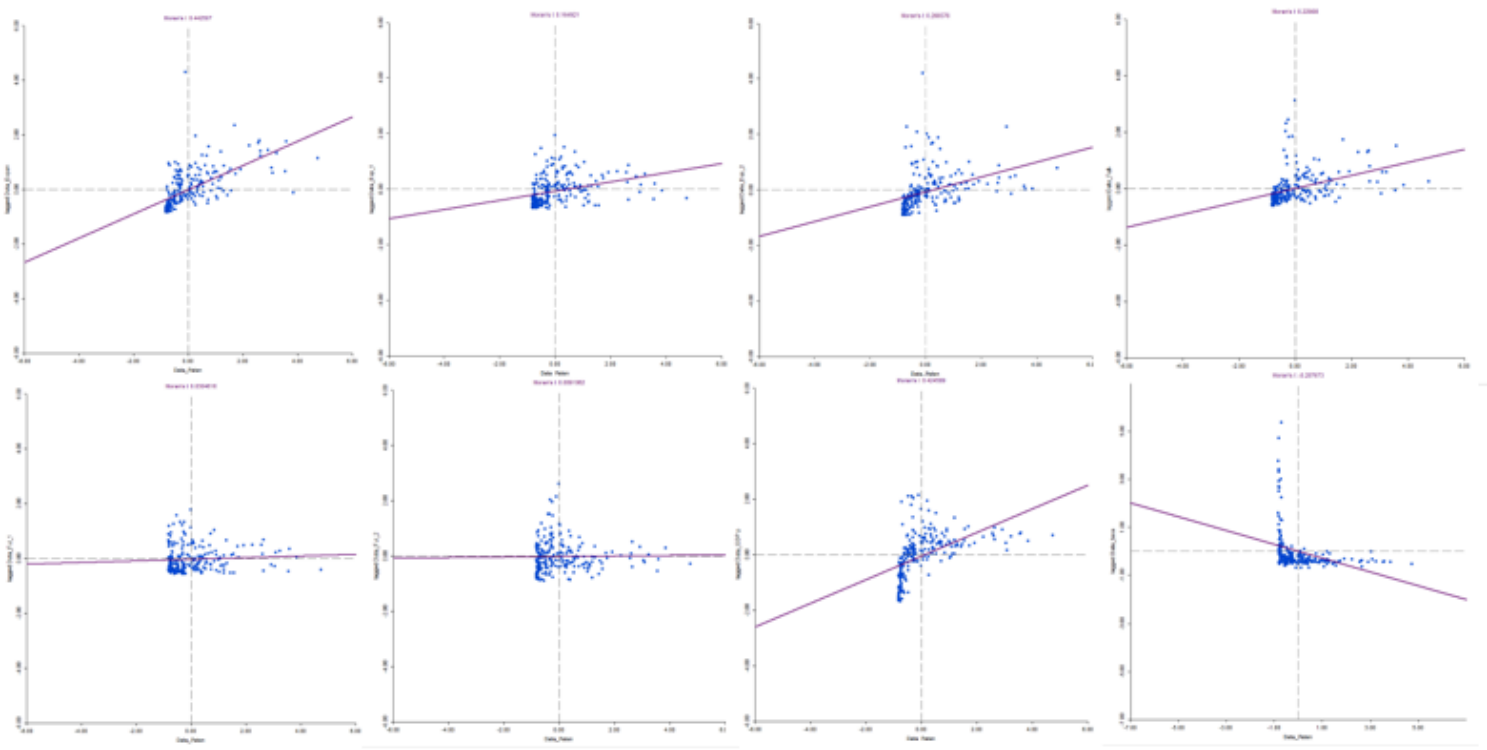
Source: Eurostat, author calculations; Software: GeoDa

Figure B10: Local Indicators of Spatial Association Significance Map for variable *Pat*, Queen (Left) and EID (Right) matrix- year 2012



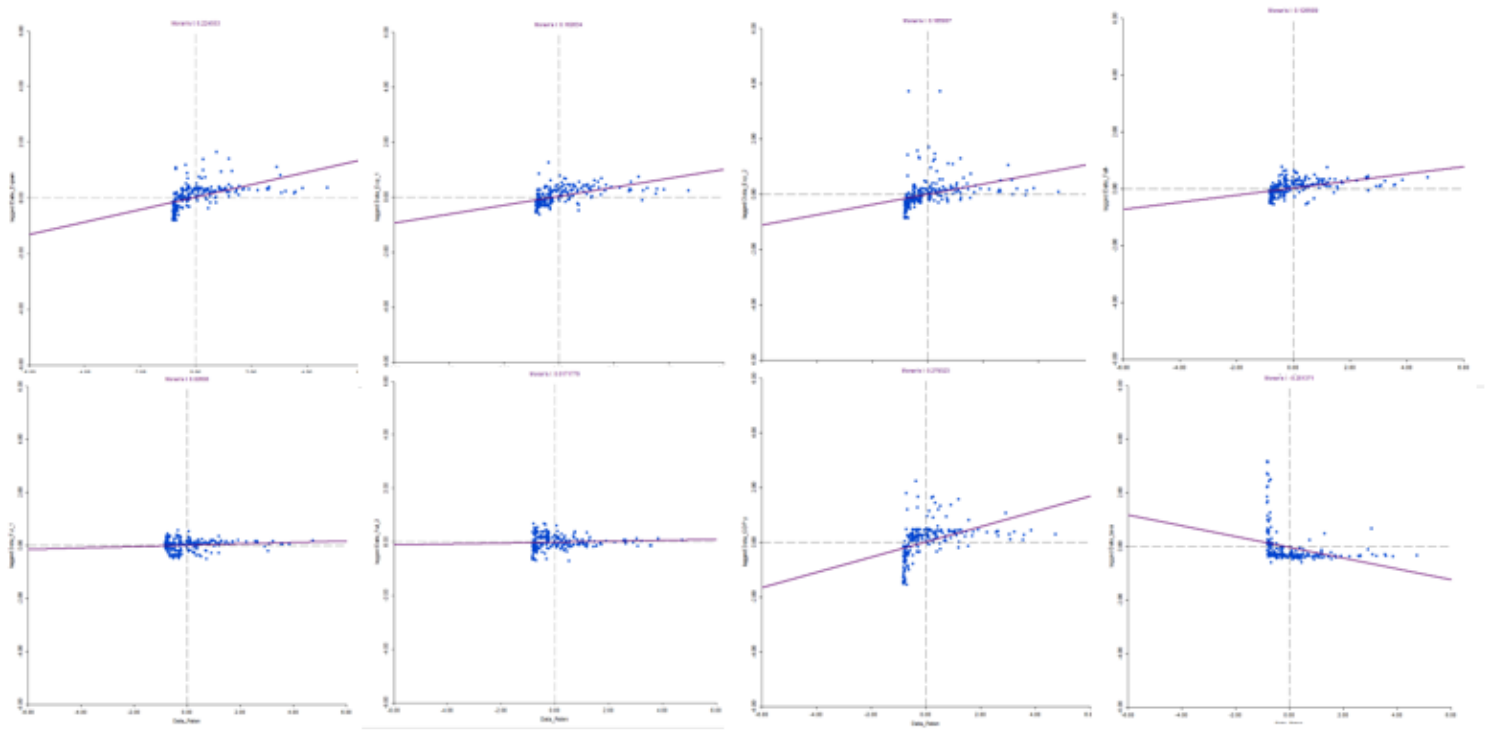
Source: Eurostat, author calculations; Software: GeoDa

Figure B11: Moran Bivariate Global Statistics I- matrix Queen



Notes:
abscissa axis – Variable *Pat*
ordinate axis - The spatially lagged covariates (from right to left and top to bottom*: R&D_B; R&D_G; R&D_U; Pers_B; Pers_G; Pers_U; GDP; Mort*)
Source: Eurostat, author calculations; Software: GeoDa

Figure B12: Moran Bivariate Global Statistics I- matrix EID



Notes:
abscissa axis – Variable *Pat*
ordinate axis - The spatially lagged covariates (from right to left and top to bottom*: R&D_B; R&D_G; R&D_U; Pers_B; Pers_G; Pers_U; GDP; Mort*)
Source: Eurostat, author calculations; Software: GeoDa

Table B13: Restricted SAR-Poisson coefficients and APE estimations

SAR-Poisson 1ºStep-ML											
			Average Partial Effects						Average Partial Effects		
Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout	Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout
ρ	6,94E-01 ***	0,06011				ρ	8,87E-01 ***	0,05916			
R&D_B	7,14E-04 **	0,00035	0,0787	0,1392	0,1332	R&D_B	9,02E-04 ***	0,00022	0,1182	1,0102	0,9131
R&D_G	-1,57E-03 **	0,00080	-0,1731	-0,3061	-0,2930	R&D_G	-1,47E-03 **	0,00078	-0,2139	-1,8285	-1,6527
R&D_U	1,41E-04	0,00073	0,0155	0,0274	0,0262	R&D_U	3,06E-04	0,00050	0,0072	0,0618	0,0559
Educ	-4,62E-06	0,00983	-0,0005	-0,0009	-0,0009	Educ	-4,78E-06	0,00718	-0,0003	-0,0028	-0,0026
Pop	8,96E-08 **	3,61E-08	9,88E-06	1,75E-05	1,67E-05	Pop	1,13E-08	3,28E-08	4,67E-06	3,99E-05	3,60E-05
GDP	3,77E-02 ***	0,00871	4,1596	7,3552	7,0398	GDP	2,15E-02 ***	0,00700	2,9851	25,5180	23,0639
Mort	-1,89E-01 **	0,09749	-20,8236	-36,8216	-35,2424	Mort	-4,54E-01 ***	0,10339	-46,0337	-393,5196	-355,6745
Log Likelihood	-6682,112					Log Likelihood	-5362.831				
W	Queen					W	Inverse distance				
c						c					
LR test	249,9154					LR test	175.831				
P-Value	0,0000					P-Value	0,0000				
N	234					N	234				

SAR-Poisson 1ºStep-OLS ad hoc constant c=1											
			Average Partial Effects						Average Partial Effects		
Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout	Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout
ρ	5,94E-01 ***	0,07224				ρ	9,05E-01 ***	0,07893			
R&D_B	9,13E-04 ***	0,00030	0,10097	0,12284	0,11744	R&D_B	1,08E-03 ***	0,00028	0,11818	0,90327	0,86201
R&D_G	-1,88E-03	0,00159	-0,20744	-0,25239	-0,24129	R&D_G	-1,95E-03 *	0,00116	-0,21390	-1,63494	-1,56027
R&D_U	-5,56E-04	0,00092	-0,06145	-0,07476	-0,07147	R&D_U	6,58E-05	0,00078	0,00723	0,05527	0,05274
Educ	-1,18E-05	0,01128	-0,00130	-0,00159	-0,00152	Educ	-3,01E-06	0,00954	-0,00033	-0,00253	-0,00242
Pop	1,42E-07 **	0,00000	0,00002	0,00002	0,00002	Pop	4,24E-08 *	0,00000	0,00000	0,00004	0,00003
GDP	4,92E-02 ***	0,01274	5,43799	6,61639	6,32538	GDP	2,72E-02 ***	0,01050	2,98509	22,81627	21,77419
Mort	-3,60E-02	0,09447	-3,98193	-4,84480	-4,63172	Mort	-4,19E-01 ***	0,10744	-46,03373	-351,85510	-335,78490
Log Likelihood	-8770.868					Log Likelihood	-6517.696				
W	Queen					W	Inverse distance				
c	1					c	1				
LR test	2133.639					LR test	1051.16				
P-Value	0,0000					P-Value	0,0000				
N	234					N	234				

Notes:

1) Standart errors were computed using Bootstrap method.
2) Significance levels: *10%, **5%, ***1%.
3) SAR-Poisson 1ºStep-ML is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients are estimated using a poisson regression.
4) SAR-Poisson 1ºStep-OLS is estimated using a two-step process. In the first step, the unobservable variable $W\log(\mu_j)$ is estimated using na OLS regression, adding a ad hoc constant (c=1) when $y_j=0$, and in the second step, the coefficients are estimated using a poisson regression.
5) All estimations were computed using the software R.
6) Both SAR-Poisson LR test p-value is based on a $\chi 2(3)$.

Table B14: Restricted SAR-LogLinear & Aspatial Poisson ML coefficients and APE estimations

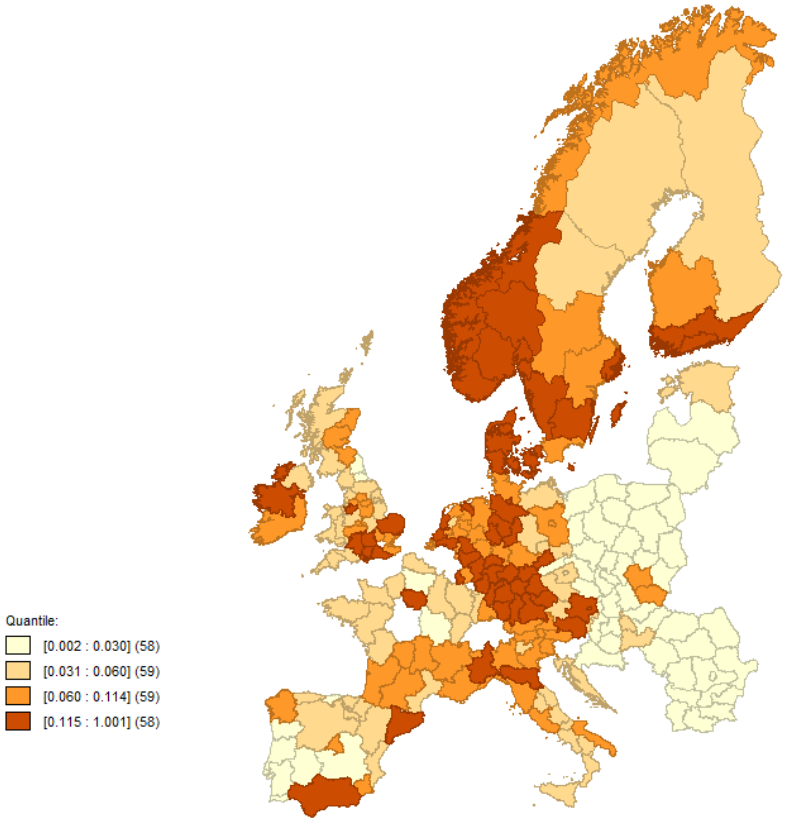
SAR-LogLinear ad hoc constant c=0.5											
			Average Partial Effects						Average Partial Effects		
Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout	Variable	Coefficients	Bootstrap SE	Direct	Spillin	Spillout
ρ	4,03E-01 ***	0,09214				ρ	4,08E-01 ***	0,09269			
R&D_B	1,05E-03 ***	0,00034	0,3514	0,2184	0,2097	R&D_B	1,15E-03 ***	0,00032	0,3689	0,2500	0,2486
R&D_G	-1,73E-03	0,00163	-0,1084	-0,0669	-0,0647	R&D_G	-1,85E-03	0,00132	-0,1115	-0,0780	-0,0753
R&D_U	-1,97E-04	0,00083	-0,0281	-0,0174	-0,0167	R&D_U	-1,77E-04	0,00080	-0,0243	-0,0164	-0,0164
Educ	1,41E-02	0,01270	21,6945	0,2423	0,2405	Educ	1,43E-02	0,01190	21,2256	0,2671	0,2648
Pop	8,27E-08	6,54E-08	0,1725	0,1025	0,1020	Pop	8,09E-08	5,77E-08	0,1622	0,1073	0,1085
GDP	5,58E-02 ***	0,01158	1,5758	0,9557	0,9384	GDP	5,47E-02 ***	0,01072	1,4860	1,0099	0,9998
Mort	-2,69E-01 ***	0,06512	-0,2833	-0,1616	-0,1703	Mort	-3,07E-01 ***	0,06777	-0,3131	-0,2025	-0,2103
Log Likelihood	-328.2144					Log Likelihood	-328.0643				
W	Queen					W	Inverse distance				
c	0.5					c	0.5				
LR test	12.482					LR test	6.378				
P-Value	0.0019					P-Value	0.0412				
N	234					N	234				

Aspatial Poisson ML			
Variable	Coefficients	Bootstrap SE	Average Partial Effects
R&D_B	6,65E-04 **	0,00028	0,04945
R&D_G	-3,63E-03 *	0,00211	-0,26953
R&D_U	-1,18E-03	0,00103	-0,08744
Educ	6,66E-02 ***	0,01268	4,95503
Pop	1,11E-07	1,04E-07	0,00001
GDP	5,76E-02 ***	0,01481	4,27970
Mort	-1,64E-01 **	0,07559	-12,19905
Log Likelihood	-13798.34		
LR test	-3607.42		
P-Value	0,0000		
W			
c			
N	234		

Notes:

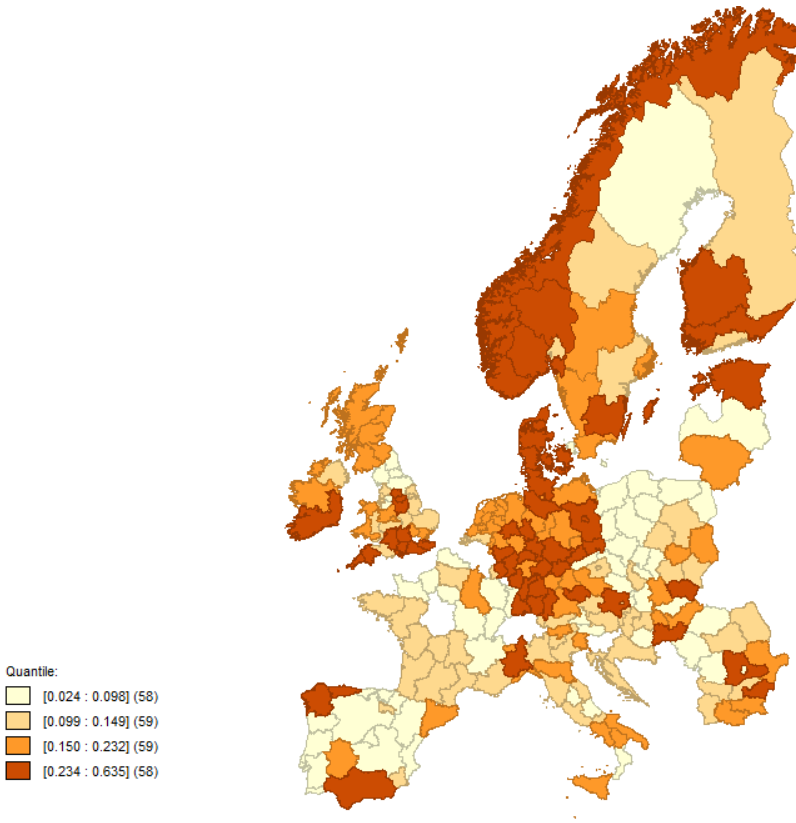
1) Standart errors were computed using Bootstrap method.
2) Significance levels: *10%, **5%, ***1%.
3) SAR-LogLinear is estimated using a two-step process. In the first step, the unobservable variable μ_j is estimated using a Poisson regression, and in the second step, the coefficients β_1,β_2 and ρ are estimated using a loglinear regression. A constant (c=0.5) is added when the dependente variable in the secondo step is zero.
4) All estimations were computed using the software R.
5) Both SAR-LogLinear and Aspatial ML Poisson LR test p-value is based on a $\chi 2(3)$.

Figure B15: Spatial Quartil Distribution map of SAR-Poisson 1ºStep-ML Queen Contiguity matrix DPE - Variable *R&D_B*, year 2012



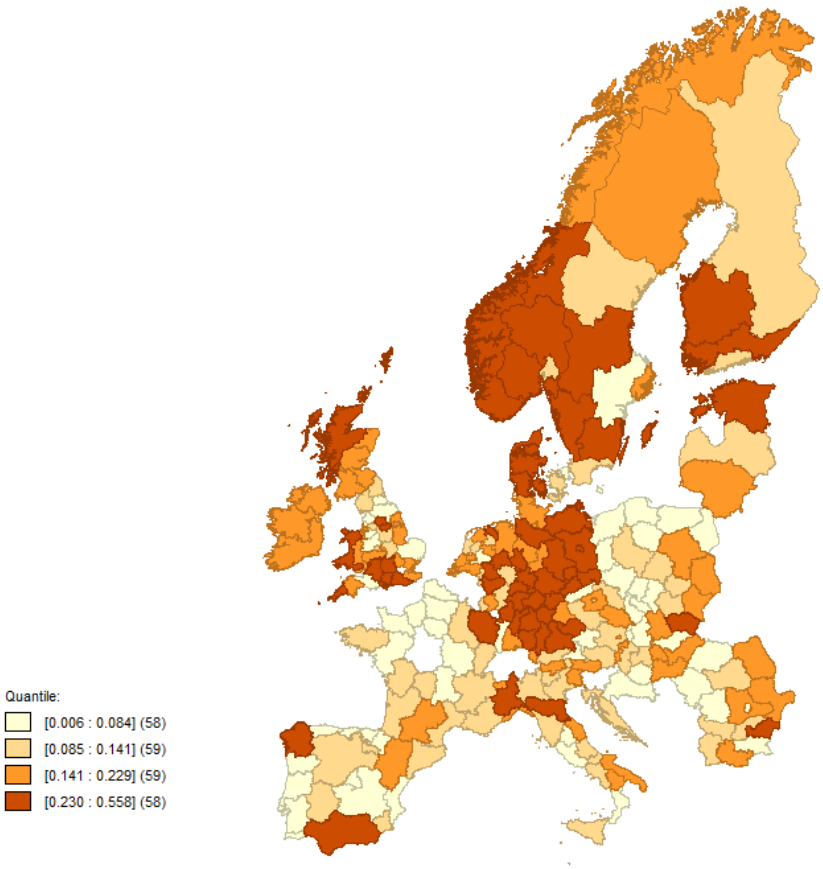
Source: Eurostat, author calculations; Software: QGIS

Figure B16: Spatial Quartil Distribution map of SAR-Poisson 1ºStep-ML Queen Contiguity matrix Spillin - Variable *R&D_B*, year 2012



Source: Eurostat, author calculations; Software: QGIS

Figure B17: Spatial Quartil Distribution map of SAR-Poisson 1ºStep-ML Queen Contiguity matrix Spillout- Variable *R&D_B*, year 2012



Source: Eurostat, author calculations; Software: QGIS

Notes for sub-chapter ***5.1.1 Exploratory Data Analysis:***

Notes: The database contains data of 234 NUTS II regions, split between 24 European Countries: Bulgaria; Czech Republic; Denmark; Germany; Estonia; Ireland; Spain; France; Croatia; Italy; Latvia; Lithuania; Hungary; Netherlands; Austria; Poland; Portugal; Romania; Slovakia; Finland; Sweden; United Kingdom; Norway. Some countries such as Belgium, Switzerland and Greece were initially considered, however in the elaboration of the final database, given the considerable lack of data in several Nuts II, these countries were excluded from the final application. Subsequently, from the set of NUTS II of the 24 selected countries, all regions with zero neighbors were excluded, and therefore, all regions consisting only of islands were bleached. Finally, of the 24 selected countries, NUTS II London (UK) and Centre (France), were excluded by incongruity of data.