Multi-dimensional Spatial Auto-regressive Models: How do they Perform in an Economic Growth Framework? [1]

Alicja Olejnik[2], Selin Özyurt[3] and Jakub Olejnik[4]

Abstract

This paper presents a conceptual and empirical framework to deal with complex spatial dependence patterns. We introduce the concept of multi-dimensional spatial weight matrix to capture more complex, asymmetric spatial interactions, including spillovers that are not necessarily conditioned by geographic distance. Most spatial studies rely merely on geographic proximity as a result of the parameter interpretation challenges associated with complex spatial structures. This paper fills this gap by suggesting an appropriate instrumental variables (IV) estimation procedure for spatial models incorporating multi-dimensional spatial weight matrices.

We provide an empirical application of the multi-dimensional spatial auto-regressive (MSAR) model. We use a three-dimensional spatial weight matrix including geographical distance as well as socio-economic factors such as economic size and human capital endowment. We find that compared to the traditional Spatial Durbin Model (SDM), the MSAR model performs better in explaining per capita income in the EU.

JEL classification: R11, R12, C21

Key words: Spatial econometrics, multidimensional spatial weight matrix, IV estimation, Solow growth model

Running head: Multi-dimensional Spatial Auto-regressive Models

---

[1] This work was supported by The National Science Centre, Poland under Grant 2011/03/D/HS4/04305.

[2] Department of Spatial Econometrics, Institute of Spatial Economy, University of Lodz, ul. Rewolucji 1905 r. 41, Poland, 90-214 Lodz, e-mail: olejnika@uni.lodz.pl

[3] European Central Bank, DG-Research - Monetary Policy Research Division, email: selin.ozyurt@ecb.int

[4] University of Lodz, Faculty of Mathematics and Computer Science, Department of Applied Computer Science, Banacha 22, 90-238 Lodz, Poland

1. Introduction

Accurate assessment of the dependence structure across cross-sectional units is a crucial step in spatial econometrics. The definition of the underlying spatial structure, which is represented by the spatial weights matrix ($\mathbf{W}$), is key to obtain reliable model parameter estimates and statistical inferences. In most cases, $\mathbf{W}$ is chosen arbitrarily, based on metric distance or contiguity. However, determining interconnectedness between cross-sectional units solely based on geographic dimension maybe quite restrictive (see Deng, 2008). Therefore, the possibility of adding more dimensions to the spatial weights matrix has been frequently suggested in the literature since the early 1970s. To name only a few, Dacey (1968) specified a non-symmetric spatial weight matrix which combined the relative area of spatial units with a binary contiguity factor and the relative length of borders between spatial units. Also, Cliff and Ord (1981) proposed a combination of the distance and boundary measures. Moreover, Bodson and Peeters (1975) introduced a general accessibility weight matrix that combined connectivity (through roads, railway lines, etc.) and geographic distance between spatial units. Besner (2002), on the other hand, presented a spatial weight matrix based on socio-economic similarity, Topa (2002) investigates the spatial patterns of unemployment via a socio-economic distance matrix based on social networks[5]. More importantly, Case et al. (1993) point out that states that are economically similar are more likely to influence each other than states that simply share common border.

However, most of the above-mentioned spatial weight matrices have the drawback of being asymmetric and incorporating unknown parameters in the weight definition. As pointed out by Anselin (1988), these complex weight structures may capture spurious relationships between spatial units and may generate unreliable inferences. Moreover, the concern of spatial weights matrix endogeneity is another factor limiting the use of an appropriate spatial specification. Anselin and Bera (1998) put forward that in the standard spatial approach, the elements of the weighting matrix have to be endogenous to the regressors. Thus, in order to ensure their endogeneity, the elements of the weighting matrix involving socioeconomic indicators should be chosen with great care, unless their endogeneity is considered explicitly in the model and estimation methodology. Kelejian and Piras (2012) argue that this concern of matrix endogeneity …"lead some researchers to purposefully misspecify, and estimate, their model by either ignoring the (obvious) endogeneity of their weighting matrix, or by selecting an inappropriate matrix which can be comfortably viewed as exogenous".

[5] Please see Kelejian and Piras (2014) for an extensive discussion of the studies using non distance-based spatial matrices.

As an alternative to the traditional spatial matrix specification, a growing number of recent papers (see Badinger et al. 2008; Lee 2006; Lee and Liu 2010, Elhorst et al. 2012, Lee and Yu, 2014 Gupta and Robinson, 2015, Han et al., 2017) use high-order spatial autoregressive processes (SAR) to capture the contiguity of units in various degrees of connectivity (e.g. neighbours, neighbours of neighbours and so on). Nevertheless, empirical applications dealing with nonlinear relationships among spatial units remain so far scarce. We believe, this is a natural outcome of the linear line of thinking, which is imposed, in a way, by relying on higher-order SAR models.

Given that the purpose of the spatial weight matrix is to approximate the “true” nature of spatial relations -regardless of their nature-, increased flexibility in the representation of the spatial structure appears rather intuitive. In spatial modelling, the definition of neighbourhood and consecutiveness tends to vary with the focus of research and the underlying cross sectional-structure. In some cases, isolated units with no direct geographical connection to other geographical units (e.g. islands) may need to be included in spatial models. Based on theory or facts adjusting the magnitude of the influence of some neighbours could give a better approximation of ‘true’ spatial effects. As an example, a very likely asymmetrical spatial relationship of commercial activity between a commune and the surrounding rural area can be given. As a result, the underlying spatial structure would be better captured by including not only geographical relations but socio-economic interactions between the core and periphery (such as transportation infrastructure, commuting and trade flows, or even ethnic linkages) as well. In this paper, we propose a conceptual framework to deal with complex spatial interdependence. We introduce the concept of multidimensional spatial weight matrix, which proves to be a more adequate approach to deal with complex spatial structures compared to the traditional weight matrix. In addition, we suggest an appropriate instrumental variables (IV) estimation procedure for spatial models incorporating multi-dimensional spatial weight matrices. The IV procedure enables to tackle, to a certain extent, the endogeneity of the right hand side variables. In addition, we use the threshold method to build instruments; in this method also enables to deal with the endogeneity of the spatial weights matrix.

The remainder of this paper is as follows. Section 2 motivates and discusses the general idea of multidimensional spatial weight matrix. Section 3 proposes the methodology to estimate spatial models with multidimensional weight matrices. An empirical application of the multidimensional spatial model is presented in Section 4. Section 5 provides a summary of the main results and concluding remarks.

2. The motivation behind the multidimensional spatial weight matrices

Let us consider a classic spatial autoregressive (SAR) model with cross-sectional observations having normal disturbances:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{u}, \quad \mathbf{u} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I}), \tag{1}$$

where $\boldsymbol{y}$ $(N \times 1)$ is a spatially-lagged endogenous variable and $\boldsymbol{X}$ is a $(N \times K)$ matrix of observations of K exogenous variables. Typically, matrix $\mathbf{W}$ is a given *a priori* spatial weight matrix with $w_{ij} = 0$ if the unit i and j do not interact spatially.

In this model, spatial parameter ρ is a scalar and is multiplied directly by the spatial weight matrix $\boldsymbol{W}$. Clearly, this approach allows for introducing only a single type of spatial spillover effect at a time. In this paper, we propose the concept of a spatial weight matrix that incorporates more than a single dimension to fit better the various types of spatial interactions amongst cross-sectional units.

As a natural consequence of multidimensionality, the spatial autoregressive coefficient $\boldsymbol{\rho}$ becomes at least a vector of parameters. Typically, the multidimensional matrix consists of a number of two-dimensional layers, each of them describing a certain aspect of spatial interactions. In this way, each element ρ_i of vector $\boldsymbol{\rho}$ corresponds to a certain layer of the multidimensional spatial weight matrix. This idea is a straightforward generalization of the standard spatial weighting concept and is likely to be more suitable for the analysis of complex spatial relations.

First, let us introduce some preliminary remarks on the multidimensional spatial weight matrix $\mathbf{D}$. We assume that the elements of matrix $\mathbf{D}$ are derived from the information on the spatial arrangement of observations in a general sense. All elements of matrix $\mathbf{D}$ must be positive and the matrix itself must be observable and exogenous. The last assumption is especially relevant for the specification of socio-economic interactions, where great attention must be drawn to avoid the issue of endogeneity.

The main advantage of the multidimensional weight matrix is its ability to accommodate and integrate several sources of spatial dependency at a time. More specifically, in some cases, the spatial analysis may require a diversification of the strength of influence on neighbours in order to include additional types of interactions (other than geographical). Thus, establishing the 'true' structure of spatial interactions would require considering explicitly a three or more dimensional space. While we traditionally assume that two dimensions of the spatial weight matrix are based on a geographic grid, the third dimension could be based on some socio-economic conditions. Regarding the nature of the research question, possible alternative dimensions could refer to similarity in the level of development, level of urbanisation and the size of the spatial units (or size of the enterprises). To illustrate, one could expect that large neighbouring enterprises would interact differently with each other than the small ones. Below, we propose some illustrations of multidimensional matrices including various dimensions of connectivity.

Case 1

In light of the New Economic Geography literature, a three dimensional spatial weight matrix makes it possible to identify the impact of core to peripheral area (see Fujita, 1999). Let us assume a geographical entity with N different spatial units. From a set of all N regions, we can select a subset of areas that are local centres, characterised by a high level of technology and innovation, employment, human capital and so on (e.g. metropolitan regions). Intuitively, we could assume that core regions not only exert a greater influence on their neighbours but also interact within larger radius than the peripheral ones[6].

This example is derived from Olejnik (2013). Here, the three-dimensional spatial weight matrix will take the form of two layers (two dimensional matrices). The first layer of the matrix reveals the pure distance and represents geographical proximity, while the second layer incorporates only core regions and their wider range of neighbours. Let $S(j)$ denote a set of neighbours of region j (i.e. a set of nearest neighbours affected by region j or regions located within a certain radius r from the centre of region j). Moreover, let C be a set of core regions selected from all N regions and SC be a set of neighbours of core regions[7]. Then, we can define $N \times 2 \times N$ matrix $\mathbf{D} = \left[D_{ikj}\right]$ of elements as follows

$$D_{i1j} = \begin{cases} 1 & for\ \ i \in S(j) \\ 0 & otherwise \end{cases} \qquad D_{i2j} = \begin{cases} 1 & for\ \ (i \in SC(j) \wedge j \in C) \\ 0 & otherwise \end{cases}. \tag{2}$$

In this matrix each layer is associated with its parameter. The parameter connected with the first layer, $\rho_{1,}$ reflects the spatial dependence resulting from pure geographical effects. In turn, the parameter associated with the second layer, ρ_2, refers to the conditional spatial effect, and therefore, reflects the extent to which regions 'surrounded' by core regions benefit from their location. Let us however notice, that the interpretation of this parameter is not straightforward, as a change by ϕ in neighbouring, but not core regions, by χ in the neighbouring core regions and by ψ in

[6] Compared to small regions, large regions would have a bigger impact on distant regions as a result of extensive trade and financial flows as well as greater labour demand and commuting flows they generate.
[7] For example, regions located within a greater radius: $r_c > r$, or larger number of nearest neighbours.

core, but distant regions (distant with radius greater than r and less than r_c), *ceteris paribus*, change the dependent variable $\boldsymbol{y}$ by

$$\rho_1 \left(\alpha \sum_{j:\, i \in (S(j) \setminus SC(j))} D_{i1j} + \beta \sum_{j:\, i \in (SC(j) \cap S(j))} D_{i1j} \right) + \rho_2 \left(\beta \sum_{j:\, i \in (SC(j) \cap S(j))} D_{i2j} + \gamma \sum_{j:\, i \in (SC(j) \setminus S(j))} D_{i2j} \right). \tag{3}$$

[8]

Case 2

Let us assume the case where the focus of interest is the influence of the neighbouring regions on the endogenous variable and it is investigated with respect to different levels of urbanization. Intuitively, more urbanised regions are could be expected to be more connected to each other. Such a research question would require the discretization of the urbanization variable. For instance, to obtain, say, three levels of urbanisation, we will need to allocate all N regions into three classes: rural regions, middle and highly-urbanised regions using the function $u: N \rightarrow \{1,2,3\}$. We will, then, obtain a three-dimensional matrix $N \times 3 \times N$ with two dimensions reserved for geographical relations and the third for the level of urbanization. In this case, Matrix $\mathbf{D}$ will have nonzero elements whenever the i-th region is a neighbour of the j-th region and these non-zero weights will relate to different levels of urbanization. This matrix has the elements

$$D_{ikj} = \begin{cases} 1 & for\ \left(i \in S(j) \wedge k = u(j)\right) \\ 0 & otherwise \end{cases}. \tag{4}$$

The spatial effects would be, then, captured by the estimated coefficients ρ_i which would vary across different levels of urbanisation but remain constant across the regions in the same class. The parameters ρ_1, ρ_2, ρ_3 are separate for the rural neighbours, middle and highly urbanised neighbours and, therefore, control for the significance of weight of the urbanisation level. Specifically, spatial effects captured by ρ_1 refer to the change in the dependent variable $\boldsymbol{y}$ with respect to a unit change in the average $\boldsymbol{y}$ in the neighbouring rural regions, *ceteris paribus*.

Case 3

Following up on the core-periphery example presented in Case 1, now let us consider that we need to account for the area size of the spatial unit to get a 'true' representation of the underlying spatial dependence structure. More specifically, some research topics may require distinguishing the strength of spatial interactions between small and large neighbouring units with large units influencing the neighbouring small ones. Thus, N spatial units will be re-grouped with respect to their

[8] Note that in this case **D** does not meet the 'row-wise orthogonality condition': $(D_{ijk} \neq 0) \Rightarrow (\forall_{l \neq k} D_{ijl} = 0)$ (see Anselin and Smirnov, 1996; Anselin and Bera, 1998). We will discuss the endogeneity issue more in detail furtherer.

size into small and large ones using the size-level-function $s: \mathbb{N} \to \{1,2\}$. Then, in our matrix $\mathbf{D}$ of dimension $N \times 3 \times N$ the first layer represents the spatial interactions among small geographically neighbouring regions, the second layer interactions among large neighbouring regions and the last one the influence of large regions on small ones. The elements of $\mathbf{D}$ are:

$$D_{ikj} = \begin{cases} 1 & for\ (i \in S(j)\ \wedge s(j) = s(i) = k) \\ 0 & otherwise \end{cases}, \qquad \text{for } k = 1,2,$$

$$D_{i3j} = \begin{cases} 1 & for\ \left(i \in S(j)\ \wedge s(j) \wedge s(j) \neq s(i)\right) \\ 0 & otherwise \end{cases}. \tag{5}$$

The degree of spatial spillover effects would be then captured by the estimated coefficients ρ_i, varying with respect to small and large neighbours, ρ_1, $\rho_{2,}$ respectively. In addition, the parameter ρ_3 will refer to the degree of spatial dependence between the regions which are significantly different in size. Therefore, ρ_3 is the change in the dependent variable resulting from a unit change in small (large) surrounding areas.

The size of a spatial unit could also be approximated by the size of its output rather than its area. As pointed out by Fingleton (2001), in some cases, assuming that large economies have greater influence on their neighbours than the smaller ones may be plausible. The size of the economy can be approximated by the level of output Q_j in region j. Therefore, given the level of output in region i, the interaction with region j would be stronger if region j has higher output. Drawing on Fingleton (2001) we suggest incorporating both size and proximity effects into the spatial weight matrix. Moreover, we suggest to control for the relative distance to the economic centre. To do so, we split the set of N regions into three subsets ($K = 3$) of regions with low, medium and high output. Let us define the function $L(k, j)$ by the following formula

$$L(k,j) = \begin{cases} 1 & \text{if } Q_j \text{ is at level } k \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

with $1 \leq j \leq N$ and $1 \leq k \leq K$. Moreover, denote the Euclidean distance between regions i and j by δ_{ij}. In the spatial weight matrix we employ, the distance decay function $\left(\delta_{ij}\right)^{-2}$, which assumes that the strength of interactions diminish with distance. Additionally, if the set of N regions constitutes an integrated economy, we can identify an economic centre of gravity for the considered economic entity (e.g. Luxemburg for the European Union as suggested by Fingleton, 2001). Assume that the distance between the centroid of each region and the exogenously assumed economic centre of gravity is denoted with δ_{jC}. Using a cut-off distance δ^* one can define the set of spatial units which are within the distance δ^* from the economic centre. That is, define $C(j)$, for $1 \leq j \leq N$, by

$$C(j) = \begin{cases} 1 & if \ \ \delta_{jC} < \delta^* \\ 0 & otherwise \end{cases} . \quad (7)$$

Finally, one can establish the fourth layer of the spatial weights matrix in the following manner.[9] The $N \times 4 \times N$ matrix has the cell entries for regions i and j and layer k:

$$D_{ikj} = \begin{cases} Q_j\left(\delta_{ij}\right)^{-2} L(k,j) & for \ \ k = 1{,}2{,}3 \\ \left(\delta_{ij}\right)^{-2} C(j) & for \ \ k = 4 \end{cases} . \quad (8)$$

The spatial contagion or clustering effects would be then captured by the estimated coefficients ρ_k, $k \leq K$ which would vary across the three levels of output but remain constant across regions. Moreover, a separate parameter ρ_4 is considered for regions in the vicinity of the economic centre. The parameters ρ_1, ρ_2, ρ_3 correspond to the weights for low, medium and high output regions, respectively. The level of spatial effects ρ_1, ρ_2, ρ_3 is reflected by the change in the dependent variable $\boldsymbol{y}$ with respect to the unit change in the average $\boldsymbol{y}$ in the surrounding regions at the comparable level of output (1, 2 or 3). The parameter ρ_4 indicates the strength of interactions in the surrounding regions located close (by the means of δ^* distance) to the economic centre of gravity.[10]

We can also impose additional assumption on the spatial effect that rural regions interact differently within the country than with regions from neighbouring countries. On the other hand, the spatial effect for the specific *core regions* can be assumed to be the same for both the level of high urbanised and high output regions.

Case 4

In some cases, the "true" spatial interactions between cross-sectional units could be better approximated by assuming a four-dimensional spatial weight matrix. To illustrate, let us assume that we investigate economic spillovers within the European countries using a four-dimensional spatial matrix. In this case, the first two dimensions could be reserved for geography, the second dimension may refer to the EU and non EU countries – two layers respectively, where the third dimension might introduce the core- periphery relation. Therefore, we would obtain a matrix of dimensions $N \times 2 \times 2 \times N$:

$$D_{iklj} = \begin{cases} 1 & for \ \ (i \in \mathrm{EU}(j) \ \wedge \ k = 1) \ or \ (i \in \mathrm{Non\ EU}(j) \ \wedge \ k = 2) \\ 0 & otherwise \end{cases}, l = 1{,}2, \quad (9)$$

$$D_{iklj} = \begin{cases} 1 & for \ \ (i \in \mathrm{Core}(j) \ \wedge \ l = 1) \ or \ (i \in Periphery\,(j) \ \wedge \ l = 2) \\ 0 & otherwise \end{cases}, \ k = 1{,}2.$$

[9] It is possible that the third and the fourth slices of the matrix bring redundant information. In this case, it would be necessary to reduce the matrix **D** by eliminating one of those slices.
[10] This is a simplified interpretation, while the exact one would require a formula similar to that described in the first example.

Clearly, the illustrations above are not exhaustive and additional cases of multidimensional spatial weight matrices can be derived. We included these cases to facilitate the understanding of the general idea and motivate further research for alternative applications.

3. Multidimensional spatial models: General description and estimation methodology

3.1. General Case of the Multidimensional Spatial Autoregressive Model

Let us consider the following spatial econometric model with a spatially-lagged dependent variable:

$$\mathbf{y} = \boldsymbol{\rho}\mathbf{D}\mathbf{y} + \mathbf{X}\beta + \mathbf{u} \quad \mathbf{u} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I}), \tag{10}$$

Above, the spatial structure of the process is represented by the general expression $\boldsymbol{\rho}\mathbf{D}\mathbf{y}$. The matrix **D** denotes a multidimensional spatial weight matrix of dimensions $N \times P_1 \times P_2 \times \ldots \times P_R \times N$, while **ρ** is a multidimensional spatial coefficient matrix. The definition of $\mathbf{D}$ and $\boldsymbol{\rho}$ depends on the focus of interest of the empirical study. In the expression $\boldsymbol{\rho}\mathbf{D}\mathbf{y}$ we use a multidimensional multiplication (as introduced in Section 2) which is determined by the dimensions of $\mathbf{D}$ and $\boldsymbol{\rho}$. Henceforth, we will call Multidimensional Spatial Autoregressive Model (MSAM) the model incorporating a multidimensional spatial weight matrix $\mathbf{D}$ as in Equation (10).

In theory, one can consider highly complex multilevel interactions within the multidimensional framework, however the obvious constraint of degrees of freedom should be borne in mind.[11]

A natural geometric interpretation of a multidimensional spatial weight matrix $\mathbf{D}$ is to view each of the square matrices $\left[D_{i,k_1,\ldots,k_R,j}\right]_{i,j\leq N}$, for every $k_1 \leq P_1, \ldots, k_R \leq P_R$, as a kind of conventional weight matrix, representing a *slice* of spatial interactions described by $\mathbf{D}$. Since a matrix of any dimension can be split into two-dimensional matrices, by applying appropriate summing operations we can rewrite the Multidimensional Spatial Autoregressive Model in the similar algebraic form to higher-order SAR model (see Brandsma and Ketellapper 1979; Blommestein 1985):

$$\mathbf{y} = \rho_1\Upsilon_1\mathbf{y} + \rho_2\Upsilon_2\mathbf{y} + \ \ldots + \rho_Q\Upsilon_Q\mathbf{y} + \mathbf{X}\beta + \mathbf{u}, (**), \tag{11}$$

where each Υ_i is a particular type of a spatial weight matrix. It is derived from multidimensional spatial weight matrix **D** as a function of its elements. Unlike the typical spatial weight matrix, it is often a sparse matrix with possibly empty rows and in most cases, not row standardised. Hence,

[11] Therefore, in spite the wide possibilities that the multi-dimensionality offers, ρ-s and D-s must be chosen in such way that the model could be identified.

because of its highly sparse and asymmetric nature, it may be sometimes difficult to give Υ_i a straightforward economic interpretation.

In the case of higher-order SAR model, matrices $\mathbf{W}_i$ correspond to various orders of contiguity, however a few more general approaches can also be found in the literature (e.g. Tao, 2005 with two spatial weight matrices):

$$\mathbf{y} = \rho_1 \mathbf{W}_1 \mathbf{y} + \rho_2 \mathbf{W}_2 \mathbf{y} + \cdots + \rho_Q \mathbf{W}_Q \mathbf{y} + \mathbf{X}\beta + \mathbf{u}. \tag{12}$$

Furthermore, in traditional papers using high-order SAR models, it is usually assumed that a given unit cannot be both a neighbour of order p and a neighbour of order r to the certain i-th unit (Anselin and Smirnov, 1996; Anselin and Bera, 1998). This prevents circularity and redundancy in the definition of contiguity, as well as easing the interpretation and identification of the spatial parameters. Nevertheless, some recent papers on the matter tend to overlook these restrictions (e.g. Lee and Liu 2010).

However, it should be borne in mind that this 'row-wise orthogonality condition' is neither necessary nor sufficient for the identification of the model (11). There is a wide range of spatial weight matrices introducing additional, relevant and irredundant information which do not satisfy the above condition and, at the same time, fulfil the condition that $[\mathbf{W}_1\mathbf{y}, \mathbf{W}_2\mathbf{y}, \dots, \mathbf{W}_Q\mathbf{y}, \mathbf{X}]$ is a full rank matrix. This statement becomes even more intuitive if we consider a multidimensional spatial weight matrix (cf. Cases 1 and 4) with each index of the matrix representing a separate 'dimension' of the spatial interactions. Obviously, if the slices are similar or introduce redundant information, they must be removed or replaced in the model.

All in all, if the spatial structure has a multidimensional character, technically, it is not systematically imperative to describe the relations in terms of multidimensional matrices. Nevertheless, the 'row-wise orthogonality stipulation' is quite binding and restricts considerably the set of all possible (identifiable) econometric models. Moreover, the multidimensional approach can appear more intuitive and eventually more accommodative when the actual spatial structure has a more compound nature. In such cases, the employment of multidimensional matrix requires from the researcher the multidimensional line of thought, which is key to establishing the 'true' structure of spatial interactions. Furthermore, incorporating the multidimensional algebra allows us to better justify the choice of the set of instruments for IV estimation procedure which will be described here.

3.2. Estimation Methodology

Previously, we pointed out that the MSAR model can be rewritten as a SAR model of higher order and, thus, can be estimated with the Maximum Likelihood (ML) procedure. Importantly, in the MSAR model, the introduction of the multidimensional spatial weight matrix allows for a multidimensional train of thought. In addition, this allows for using a multidimensional Instrumental Variables procedure involving a wider range of instruments.

a. Some notation of the multidimensional matrix algebra

Let us first introduce some elementary properties of the multidimensional matrix algebra that we use in the presentation of our MSAR model. Below, we give two definitions which are essential to our reasoning. Namely, we define the idea of order M matrix multiplication, as well as the notion of an *f-column*. Let $[a_{ij}]_{i,j}$ denote the matrix with elements 'a_{ij}' and multi-indexes 'ij' as in Saint Raymond (1991) and Suijlekom, W.D. Van. ."Schwartz Function"[12] . Moreover, for any matrix $\boldsymbol{A}$, let $(\boldsymbol{A})_{ij}$ denote the i,j-th element of $\boldsymbol{A}$, namely $(\boldsymbol{A})_{ij} = a_{ij}$ if $\boldsymbol{A} = [a_{mk}]_{m,k}$.

Definition 1

Let $\mathbf{A}$ *be a matrix with* $G+M$ *number of dimensions and* $\mathbf{A} = [a_{ij}]_{i,j}$ *where* $i = (i_1, \dots, i_G)$, $j = (j_1, \dots, j_M)$ *with* $1 \le i_1 \le g_1, \dots, 1 \le i_G \le g_G$ *and* $1 \le j_1 \le m_1, \ \dots, 1 \le j_M \le m_M$. *Similarly, let* **B** *be a matrix of* $M+R$ *dimensions such as* $\boldsymbol{B} = [b_{jl}]_{j,l}$ *where* $j = (j_1, \dots, j_M)$, $l = (l_1, \dots, l_R)$ *satisfy* $1 \le j_1 \le m_1, \dots, 1 \le j_M \le m_M$ *and* $1 \le l_1 \le r_1, \dots, 1 \le l_R \le r_R$. *We define the multidimensional matrix product of order M (the symbol* $\cdot_M$*) by the following:*

$$\boldsymbol{A} \ \cdot_{\mathrm{M}} \boldsymbol{B} = \left[c_{i_1 i_2 \dots i_G l_1 l_2 \dots l_R}\right]^{(G+R)}_{(i.\le g.),(l.\le r.)},$$

where

$$c_{i_1 i_2 \dots i_G l_1 l_2 \dots l_R} = \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \dots \sum_{j_M=1}^{m_M} \left[a_{i_1 i_2 \dots i_G j_1 j_2 \dots j_M} \cdot b_{j_1 j_2 \dots j_M l_1 l_2 \dots l_R}\right]. \tag{13}$$

In practical terms, $j_1, \dots, j_M$ correspond to the dimensions of the weighting matrix resulting from the multiplication of multidimensional spatial weight matrix by the dependent variable. Along these dimensions, we find the different layers of interactions through which the neighbourhood defined along the remaining dimensions $i_1, \dots, i_G, l_1, \dots, l_R$ exerts an influence.

[12] Available from the link http://mathworld.wolfram.com/SchwartzFunction.html

Let us notice that this is a generalization of the classic matrix multiplication. The 'traditional' matrix product is a special case of the operation $\cdot_1$ for the rectangular matrices. Indeed, we have $\boldsymbol{AB} = \boldsymbol{A} \cdot_1 \boldsymbol{B} = [\sum_{j=1}^{m} a_{ij} b_{jl}]_{i\leq n, l\leq r}$. Moreover, as suggested in the proof of Lemma 1, the multidimensional matrix multiplication is associative if the two consecutive operations in the chain do not involve overlapping matrix dimensions. For example, if number of dimensions of $\boldsymbol{B}$ is at least $k_1 + k_2$, then the un-parenthesized statement $\boldsymbol{A} \cdot_{k_1} \boldsymbol{B} \cdot_{k_2} \boldsymbol{C}$ can be defined unambiguously.

Definition 2

Let $\boldsymbol{A} = [a_{i,i_1,\dots,i_n}]$ *be a matrix with* $ndim(\boldsymbol{A}) \geq 1$, $(n \geq 0)$. *We say that vector* **x** *is an* f-column *of* **A** *(or alternatively – a column along the first dimension) if* $\boldsymbol{x} = [x_i]_i = [(\boldsymbol{A})_{i\, i_1 \dots i_n}]_i$ *for some fixed* $i_1, \dots, i_n$.

Thence, an *f-column* is a vector extracted from a matrix $\boldsymbol{A}$ by fixing all except the first index.

b. MSAR Model specification

Above, we introduced a general definition of multidimensional spatial weight matrix $\mathbf{D}$. Overall, a matrix $\mathbf{D}$ can be considered as matrix of any number of dimensions. However, without loss of generality, we can restrict our attention to the case where $\mathbf{D}$ is a three dimensional matrix of dimensions $N \times P \times N$ with zero diagonal whenever all but the first and the last index is fixed. Moreover, we assume that it is an ordinary sub-normalized spatial weight matrix whenever the second index is fixed. More precisely, we assume that the sum of the row elements in each slice is not more than one. Furthermore, we assume that one of the slices is a 'classic' row standardized **W** matrix[13] representing purely geographical interactions. For further simplicity of notation, we assume that the dimension P equals 3. Therefore, $\mathbf{D}$ consists of three rectangular matrices concatenated with respect to the third dimension: $\mathbf{D}_{[.1.]} = [\mathbf{W}], \mathbf{D}_{[.2.]} = [\mathbf{D}_2], \mathbf{D}_{[.3.]} = [\mathbf{D}_3]$.

Consider a Multidimensional Spatial Autoregressive Model

$$\mathbf{y} = \boldsymbol{\rho}\mathbf{D}\mathbf{y} + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\theta + \mathbf{u} = \rho\mathbf{D}\mathbf{y} + \Xi\mathbf{B} + \mathbf{u},\ \ \mathbf{u} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I}), \tag{14}$$

with $\Xi = [\mathbf{X}, \mathbf{WX}]$.[14] In the above model, the multidimensional lag operator $\mathbf{D}\boldsymbol{y}$ is an $N \times 3$ matrix, where $\mathbf{D}\boldsymbol{y} \coloneqq \mathbf{D} \cdot_1 \boldsymbol{y}$. The parameter $\boldsymbol{\rho}$ is a composition of three diagonal matrices with elements ρ_1, ρ_2, ρ_3, respectively. Then, $\boldsymbol{\rho}\mathbf{D} \coloneqq \boldsymbol{\rho} \cdot_2 \mathbf{D}$, where $\boldsymbol{\rho}_{(N\times N\times 3)} = [\rho_k (\mathbf{I}_N)_{i,j}]_{i\leq N, j\leq N, k\leq 3} = [\rho_{ijk}]_{i\leq N, j\leq N, k\leq 3}$.

[13] Where, the elements w_{ij} in each row sum to 1.

[14] Let us notice that we do not put any additional restrictions on the vector of parameters $\boldsymbol{\theta}$ as it is done in Spatial Durbin model, where each element θ_i of the vector $\boldsymbol{\theta}$ would have to satisfy the condition $\theta_i = \rho_i \beta_i$.

An alternative way to give meaning to the term $\boldsymbol{\rho}\mathbf{D}$ in the case of this particular $\boldsymbol{\rho}$ is by $\boldsymbol{\rho}\mathbf{D}\mathbf{y} \coloneqq (\mathbf{D} \cdot_1 \boldsymbol{y}) \cdot_1 \overline{\boldsymbol{\rho}}$, with simply $\overline{\boldsymbol{\rho}} = (\rho_1, \rho_2, \rho_3)$. The matrix $\mathbf{X}$ is an $N\times K$-matrix of explanatory variables, while $\mathbf{y}$ is a dependent variable. $\mathbf{WX}$ is $N \times K$-matrix while $\mathbf{Z}$ is a matrix of $N \times (3 + 2K)$ dimensions. Thus, matrix $[\mathbf{Dy}, \Xi]$ must be, obviously, a full rank matrix.

c. *The IV Estimation procedure*

The Instrumental Variables procedure is a suitable approach for the estimation of the MSAR model (Kelejian and Prucha, 1998). Below, we introduce reasoning for establishing the matrix of instruments appropriate for the multidimensional spatial model. Let us first define and highlight the key elements of the estimation methodology. For any multidimensional matrix $\boldsymbol{A}$ with $\mathrm{ndim}(A) \geq 1$, assuming the following notation:

$$\boldsymbol{A}^{n[1]} = \underbrace{\boldsymbol{A} \cdot_1 \boldsymbol{A} \cdot_1 \dots \cdot_1 \boldsymbol{A}}_{n\ times} \tag{15}$$

For introducing our reasoning, it is necessary to formulate a Lemma. The proof of the Lemma 1 is given in Appendix.

Lemma 1

We have

$$(\boldsymbol{\rho}\ \cdot_2 \mathbf{D})^{n[1]}\Xi\mathbf{B} = (\overbrace{\boldsymbol{\rho}\ \cdot_2 (\boldsymbol{\rho}\ \cdot_2 \dots}^{n} \underbrace{(\mathbf{D} \cdot_1 \dots \cdot_1 \mathbf{D})))}_{n}\ \Xi\mathbf{B} = \boldsymbol{\rho}\ \cdot_2 (\boldsymbol{\rho}\ \cdot_2 \dots (\mathbf{D} \cdot_1 \dots \cdot_1 \mathbf{D}\ \cdot_1 (\Xi\mathbf{B}))). \tag{16}$$

With ρ and $\boldsymbol{B}$ being fixed, the left-hand side of (3) *is a linear combination of the f-columns of $\boldsymbol{D}^{n[1]}\ \cdot_1 \Xi$. Moreover each column of $\boldsymbol{D}\ \cdot_1 (\boldsymbol{\rho}\ \cdot_2 \boldsymbol{D})^{n[1]}\Xi\boldsymbol{B}$ is a linear combination of the f-columns of $\boldsymbol{D}^{(n+1)(1)} \cdot_1 \Xi$.*

An important issue which must resolved before we suggest a possible method of estimation is the construction of parameter space $\mathcal{P}$. This can be conveniently done in terms of members of (vectorised) matrix $\boldsymbol{\rho}$. Since it is natural to include $\boldsymbol{\rho}_0 = \mathbf{0}$ in the set $\mathcal{P}$, as the parameter value indicating the absence of autocorrelation, any pathwise-connected neighbourhood $\mathcal{N}$ of $\boldsymbol{\rho}_0$ satisfying $\mathcal{N} \subset \{\boldsymbol{\rho} : (\boldsymbol{I} - \boldsymbol{\rho}\mathbf{D}) \text{ is invertible}\}$ could serve as a reasonable candidate for $\mathcal{P}$. Obviously, one is interested in the parameter space $\mathcal{N}$ to be reasonable large. However, in case of the estimation procedure used in this paper we also need the inverse $(\mathbf{I} - \boldsymbol{\rho}\mathbf{D})^{-1}$ to be a sum of the geometric series $\sum_{k=0}^{\infty}(\boldsymbol{\rho}\mathbf{D})^k$, thus the set $\mathcal{N}$ has to be suitably trimmed. Some available methods of construction of the set $\mathcal{N}$ are given in the Appendix.

We use an Instrumental Variables procedure as the OLS estimator would be biased and inconsistent in Model 2 where the lagged dependent variable is included to the right hand side. In view of the complexity of the matrix **D,** the choice of the matrix of instruments is not obvious. Following Kelejian and Prucha (1998), the most suitable instrument for $\mathbf{Dy}$ would be the expectation $\mathbf{Dy}$. Let $\boldsymbol{\rho} \in \mathcal{N}$, applying a standard reasoning to the multidimensional model we obtain:

$$\begin{aligned}\mathcal{E}\mathbf{Dy} &= \mathbf{D}(\mathbf{I}-\rho\mathbf{D})^{-1}\Xi\mathbf{B} = \mathbf{D}[\mathbf{I}+\rho\ \cdot_2 \mathbf{D} + (\rho\ \cdot_2 \mathbf{D})\ \cdot_1 (\rho\ \cdot_2 \mathbf{D}) + \ldots]\Xi\mathbf{B} = \\ &= \mathbf{D}[\mathbf{I}+\rho\ \cdot_2 \mathbf{D} + (\rho\ \cdot_2 \mathbf{D})^{2(1)} + \cdots]\Xi\mathbf{B} = \mathbf{D}\Xi\mathbf{B} + \mathbf{D}(\rho\ \cdot_2 \mathbf{D})\Xi\mathbf{B} + \mathbf{D}(\rho\ \cdot_2 \mathbf{D})^{2(1)}\Xi\mathbf{B} + \cdots .\end{aligned} \tag{17}$$

The above and Lemma 1 imply that $\mathbb{E}\mathbf{Dy}$ is a (infinite) linear combination of the *f-columns* of the matrices $\mathbf{D}^{\boldsymbol{n}(\mathbf{1})}\ \cdot_{\mathbf{1}} \Xi,\ \textbf{for}\ \boldsymbol{n} = \mathbf{1}, \mathbf{2}, \ldots$.

Applying Kelejian and Prucha (1998) reasoning to the multidimensional model, we choose as instrument, a matrix **H** of linearly independent *f-columns* of matrices Ξ, $\mathbf{D}\Xi$, $\mathbf{D}^{\mathbf{2}(1)}\Xi$. In our case, since the operation **D** 'includes' **W**, it is sufficient to consider a set of linear independent *f-columns* of matrices: $\mathbf{X}$, $\mathbf{DX}$, $\mathbf{D}^{2(1)}\mathbf{X}$, $\mathbf{D}^{2(1)}\mathbf{WX}$.[15]

The projection matrix is defined as $\mathbf{P}_{\mathrm{H}}^{(N\times N)} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$. As a result, the projected matrix **Z** takes form $\hat{\mathbf{Z}} = \mathbf{P_H Z} = \mathbf{P_H}[\mathbf{Dy}, \mathbf{X}, \mathbf{WX}] = [\mathbf{P_H Dy}, \mathbf{X}, \mathbf{WX}]_{(N\times(3+2K))}$. Moreover, by the Definition 1 we have $\mathbf{DX} = \mathbf{D}\ \ \cdot_1 \mathbf{X} = [\sum_{k=1}^{N} D_{ijk}X_{kl}]_{i\le N, j\le 3, l\le K}$.

Therefore, we obtain the IV estimator of the form

$$\hat{\delta} = (\hat{\mathbf{Z}}^T\hat{\mathbf{Z}})^{-1}\hat{\mathbf{Z}}^T\mathbf{y}, \tag{18}$$

where $\hat{\delta} = [[\rho_1, \rho_2, \rho_3], \beta^T, \theta^T]^T$ is of dimension $(3+2\,K)\ \times 1$.

The asymptotic properties of the IV estimator are well known. Thus, similarly to the standard reasoning, it can be shown that the estimator is consistent and its variance is given by the formula $Var(\hat{\delta}) = (\sum_i e_i^2/N)(\hat{\mathbf{Z}}^T\hat{\mathbf{Z}})^{-1}$with e_i denoting residuals. Besides, it should be emphasised that although we consider here a three dimensional spatial weight matrix, the analogical reasoning can be applied to the matrices of higher dimensions.

4. Empirical illustration of the MSAR Model

In this section, we illustrate the multidimensional spatial dependence structure applied to economic data. Our research focus is to identify the determinants of regional per-capita income growth in the EU-25. The main contribution of our analysis is the use of the three-dimensional spatial weight matrix

[15] We shall assume that **H** contains at least f-columns of $[\mathbf{X}, \mathbf{WX}]$.Similar estimation procedure could be conducted without this assumption since at least **X** alone is a full rank matrix.

that enables the incorporation of some additional economical interactions into a Spatial Autoregressive Model (SAR). We employ a refined version of the idea of directional spatial dependence described previously in Deng (2008).

4.1 Theoretical background and data

The theoretical background of our empirical application relates to the growth literature and the specification chosen can be seen as a traditional Solow (1953) model augmented with human capital and technology level (as in Mankiw et al., 1992). In the standard neoclassical Solow model, national output is determined by the level of capital, technology and labour. The rationale for including human capital into the model specification has been stressed in subsequent studies (Lucas 1988, Mankiw et al., 1992).

The Mankiw et al. (1992) model can be written as:

$$Y = K^{\alpha} H^{\beta} [AL]^{1-\alpha-\beta}, \tag{19}$$

where Y denotes the level of output, K - physical capital, H - human capital, A - technology, L - labour input, α and β are the physical and human capital elasticities of output, $(0 < \alpha < 1,\ 0 < \beta < 1)$ and $1 - \alpha - \beta > 0$ is labour input elasticity of output. L_t and A_t are assumed to grow at the exogenous rates n and g, respectively. The dynamics of the economy is determined by:

$$\dot{k}_t = s_k \frac{Y_t}{A_t L_t} - (n + g + \delta) \frac{K_t}{A_t L_t} \tag{20}$$

$$\dot{h}_t = s_h \frac{Y_t}{A_t L_t} - (n + g + \delta) \frac{H_t}{A_t L_t} \tag{21}$$

where s_k and s_h are the investment rates in physical and human capital and δ is the depreciation rate (assumed to be the same for the two types of capital). Assuming decreasing returns to physical and human capital $(\alpha + \beta < 1)$, Eq. (5) and (6) imply that the economy converges to a steady state (denoted by superscript $*$) defined by:

$$k^* = \left(\frac{s_k^{1-beta} s_h^{beta}}{n+g+\delta} \right)^{\frac{1}{1-\alpha-\beta}} \tag{22}$$

$$h^* = \left(\frac{s_k^{\alpha} s_h^{1-\alpha}}{n+g+\delta} \right)^{\frac{1}{1-\alpha-\beta}} \tag{23}$$

Substituting the two steady-state forms above into one and taking logs gives the equation for output per capita, which will be the theoretical basis of our empirical specification:

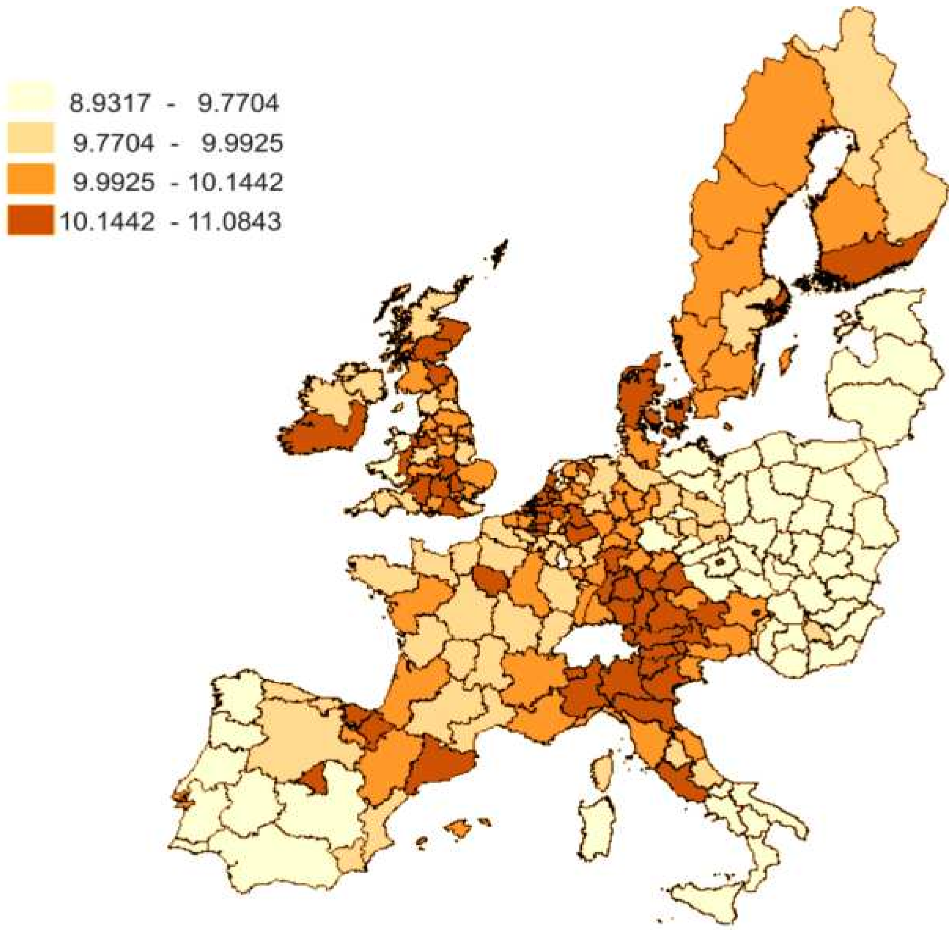
$$\ln y = \alpha_0 - \frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) + \frac{\alpha}{1-\alpha-\beta}\ln s_k + \frac{\beta}{1-\alpha-\beta}\ln s_h + \xi \quad (24)$$

Equation above shows that output per capita depends on initial technology level (A_0), technological progress (g), demographic changes (n), investment in physical capital (s_k) and the level of human capital (s_h). These variables are included in our empirical specification conditioned on data availability.

The data used originate from the *Regio* - regional database of the *Eurostat*[16] and covers the year 2004. Table 1 reports the summary statistics of the variables. The income (LY) is expressed in Purchasing Power Parity in constant prices. Figure 1 illustrates the distribution of the per-capita income in the EU regions considered. It can be observed that there is a clear tendency for regions to cluster based on their per-capita income levels (i.e. positive autocorrelation). Low-income regions are mostly concentrated in the Central, Eastern and South-eastern Europe (CESEE) countries, in Portugal as well as in Southern Italy and Spain. By contrast, we can identify a concentration of high-income regions in a band going from the London area to Northern Italy, including South-Western Germany, Austria and the South-East of France. The largest European cities are also among the regions with highest income levels, although more dispersed geographically (e.g. Paris, Madrid, Brussels, Hamburg, Manchester, Edinburg). Moreover, strong income disparities in the Southern Europe, Italy, Portugal, Spain and to a lesser extent France, have already existed in 2004. The investment variable refers to the investment rate in the manufacturing sector (per person employed, € 1,000) and it is assumed as a predetermined variable (INV). Human capital is approximated by the Human Resources in Science and Technology, defined as the percentage of active population in research and development (HR). Furthermore, by adopting a standard approach, we assume that both depreciation and productivity grow at a constant rate of 5% (*NGD*). Moreover, we have also brought in a new-member-country dummy variable (NEW) in order to capture other economic factors that are not considered in the model specification. This variable takes the value of one for the countries accessing the EU in the Fifth Enlargement in 2004 and zero for all remaining states. All variables are included in the log form with the exception of the dummy variable. The underlying spatial structure is specified using a three dimensional spatial weight matrix $\mathbf{D}$ as described previously.

[16] http:// epp.eurostat.ec.europa.eu

Figure 1. The per-capita income in the 228 NUTS 2 European Union Regions (EUR), 2004



Source: *Regio – Eurostat*

Table 1. Description of the variables

Variable	Definition	Mean	σ	Min	Max
Y	natural logarithm of gross domestic product per-capita converted into a common scale using *PPP*	9.93	0.36	8.93	11.08
NGD	$\ln(n + g + \delta)$, where: n - rate of labour growth, g - technology level rate of growth, δ - the depreciation rate	2.31	0.17	1.8	2.81
INV	natural logarithm of investment in manufacturing (per person employed (€ 1000)) $\approx s_k$ - Investment Rate	1.89	0.43	0.64	3.07
HR	natural logarithm of Human Resources in Science and Technology (percentage of active population) $\approx s_h$ - Investment in Human Capital	3.53	0.22	2.69	4.04
NEW	new member countries dummy variable	0.17	0.37	0.0	1.0
$\mathbf{D}$	three dimensional spatial weight matrix $(228 \times 3 \times 228)$				
$\boldsymbol{W}$	row standardised weight matrix (228×228) of three nearest neighbours				

The data used refers to 228 NUTS 2 European regions, covering 23 Member States of the European Union except Malta, Cyprus and some regions of France and Portugal because of their geographically isolated positions. Greece, Slovenia and Luxemburg had to be dropped from the sample due to the absence of comparable data.

4.2 Model specification and results

Here, we introduce a three-dimensional spatial weight matrix which we employ in the empirical illustration. Our goal is to account for those socio-economic factors which may not necessary result in symmetric spatial interactions between regions. Such interactions may be modelled, as an asymmetric matrix being a function of an observable variable.

For $\alpha \in \mathbb{R}_+$ and a variable X, we define the auxiliary matrix ($\boldsymbol{D}_X^\alpha$) that indicates whether value of the economic variable X is higher in the j-th spatial unit than α times the value in the i-th. Thus, the asymmetric economic effect of a given variable is not generated by its mere level in a unit *per se* but by the relation to its level in the unit's spatial neighbours'[17]. The threshold parameter α enables to introduce a quantitive definition of the relation between levels. In particular, α defines statements

[17] For further details see Deng (2008).

such as ‘considerably larger than’, for $\alpha > 1$, and vice versa. An important advantage of using the threshold parameter $\alpha \neq 1$ is the substantial elimination of endogeneity in $\boldsymbol{D}_X^\alpha$, should the variable X be itself endogenous. Namely, choosing α, in a way that satisfies $|\alpha - 1| > \sigma(X_i)/\mathbb{E}X_i$ - spatial variability, for each location i, lets the matrix $\boldsymbol{D}_X^\alpha$ reflect profound structural differences between spatial units instead of being sensitive to local variation in X.

In particular, in our model, the auxiliary matrices: $\mathbf{D}_{\mathrm{GDP}}^{0.95}$ and $\mathbf{D}_{\mathrm{HR}}^{0.95}$ are defined as follows:

$$\mathrm{D}_{\mathrm{GDP}}^{0.95} = (\mathrm{d}_{\mathrm{GDP}}^{\mathrm{ij}})_{\mathrm{n}\times\mathrm{n}}, \text{where } \mathrm{d}_{\mathrm{INV}}^{\mathrm{ij}} = 1_{\{\mathrm{i,j:\ GDP_j > 95\% \cdot GDP_i}\}}(\mathrm{i},\mathrm{j}), \tag{25}$$

$$\mathbf{D}_{\mathrm{HR}}^{0.95} = (d_{\mathrm{HR}}^{ij})_{n\times n}, \text{where } d_{\mathrm{HR}}^{ij} = \mathbf{1}_{\{i,j:\ \mathrm{HR}_j > 95\% \cdot \mathrm{HR}_i\}}(i,j), \tag{25}$$

where (only in the above two formulas) the variables are not expressed in logarithms. Then, to define the three-dimensional matrix $\mathbf{D}$, we first consider $\widetilde{\mathbf{D}}$ which takes the following form:

$$\widetilde{\mathbf{D}}_{(N\times 3\times N)} = [(\widetilde{\mathbf{D}}_k)_{i,j}]_{i\le N, k\le 3, j\le N} = [\widetilde{D}_{ikj}]_{i\le N, k\le 3, j\le N}, \tag{26}$$

where $\mathbf{D}_1 = \mathbf{W}, \widetilde{\mathbf{D}}_2 = \mathbf{D}_{\mathrm{GDP}}^{0.95} \circ \mathbf{W}, \widetilde{\mathbf{D}}_3 = \mathbf{D}_{\mathrm{HR}}^{0.95} \circ \mathbf{W}$**,** with $\circ$ denoting the Hadamard, product, matrix $\mathbf{W}$ is an $N \times N$ geographical spatial weight matrix of three nearest neighbours. In order to facilitate the spatial interpretation of the $\boldsymbol{\rho}$ parameter, we also choose to introduce an additional correction for $\mathbf{D}_2$ and $\mathbf{D}_3$. Namely, for $k = 2,3$ let us denote $M^k = \left\{1 \le i \le N\text{: } \sum_{j=1}^N D_{ikj} = 0\right\}$ and let $\left|M^k\right|$ be cardinality of the set M^k. Then, we stipulate that

$$\widetilde{D}_{ikj} = \left(N - \left|M^k\right|\right)^{-1} \cdot \sum_{j' \notin M^k} \widetilde{D}_{j'ki}\,, \text{for } i \in M^k, 1 \le j \le N \text{ and } k = 2,3. \tag{27}$$

Finally, all $\mathbf{D}_{\mathrm{k}}$, for $k = 1,2,3$ are obtained from $\widetilde{\mathbf{D}}_k$ by standardising row sums. Below, we employ the multidimensional matrix to investigate the determinants of per capita income in the EU. To this end, we consider the following specification.

$$\begin{aligned} \mathrm{Y} = \alpha + \boldsymbol{\rho} \cdot_2 \mathbf{D} \cdot_1 \mathrm{Y} + \beta_1 \mathrm{NGD} + \beta_2 \mathrm{INV} + \beta_3 \mathrm{HR} + \\ + \theta_1 \mathbf{W} \cdot \mathrm{NGD} + \theta_2 \mathbf{W} \cdot \mathrm{INV} + \theta_3 \mathbf{W} \cdot \mathrm{HR} + \varsigma \mathrm{NEW} + \mathrm{u}, \end{aligned} \tag{28}$$

where $\boldsymbol{\rho}$ has been defined in Section 3.1 and $\mathbf{D}$ described in the previous paragraph. Although the specification above appears similar to the conventional Spatial Durbin Model (SDM), the multi-dimensional nature of Matrix $\mathbf{D}$ enables to capture more complex, asymmetric spatial interactions, including spillovers that are not necessarily conditioned by geographic distance. As a result, the coefficient associated with the spatially lagged endogenous variable provides a better approximation of the power and the nature of inter-regional interactions.

Table 2. The OLS Model

$\mathbf{Y} = \boldsymbol{\alpha_0} + \boldsymbol{\beta_1}\mathbf{NGD} + \boldsymbol{\beta_2}\mathbf{INV} + \boldsymbol{\beta_3}\mathbf{LHR} + \boldsymbol{\beta_4}\mathbf{NEW} + \boldsymbol{u}$				
Variable	Coefficient	Std.	t-Stat	p-value
Const.	6.1684	22.04	24.02	$<$ 0.0001
INV	0.0345	0.84	-0.50	0.3996
NGD	0.3514	4.47	4.44	$<$ 0.0001
HR	0.8332	12.46	13.07	$<$ 0.0001
NEW	-0.4104	8.39	-10.45	$<$ 0.0001
$\bar{R}^2 = 0.72$				

As a starting point, Table 2 reports the OLS estimation results. However, in the specification above, the inclusion of the spatially lagged dependent variable into the right-hand side is likely to create endogeneity as the spatially lagged dependent variable $\mathbf{W}\mathrm{Y}$ would be correlated with the error term u. As a consequence, the estimation of the spatial Durbin model with the OLS estimator may generate biased and inconsistent parameters and statistical inferences. This also implies that the OLS results may not be directly comparable to the IV estimation results.

From the results, we notice that the parameter associated with human capital is highly significant although the coefficient of investment intensity is not. This result may be due to imperfections of the investment indicator which corresponds to the investment rate in the manufacturing sector only. The high level of significance of the new member countries dummy variable confirms that such control variable is relevant for explaining the differences in per capita income dynamics, most probably reflecting the catch up effect of the 'new' member countries. Table 3 presents the results from the Instrumental Variable estimation procedure for the chosen SDM specification[18].

Table 3. The estimation results of SDM

$\mathbf{Y} = \boldsymbol{\alpha_0} + \boldsymbol{\rho}\mathbf{WY} + \boldsymbol{\beta_1}\mathbf{NGD} + \boldsymbol{\beta_2}\mathbf{HR} + \boldsymbol{\beta_3}\mathbf{NEW} + \boldsymbol{\theta_1}\mathbf{WNGD} + \boldsymbol{\theta_2}\mathbf{WHR} + \boldsymbol{u}$				
Variable	Coefficient	Std.	t-Stat	p-value
WY	0.8355	0.1630	5.1264	< 0.0001
Const.	1.0298	1.0561	0.9751	0.3306
NGD	0.4497	0.1017	4.4236	< 0.0001
HR	0.8892	0.0699	12.7135	< 0.0001
NEW	-0.1183	0.0714	-1.6562	0.0991
WNGD	-0.3403	0.1333	-2.5532	0.0113
WHR	-0.7887	0.1763	-4.4739	< 0.0001
$\bar{R}^2 = 0.79$				

[18] Variables INV and **W**INV were excluded from the model as they did not appear significantly different from zero.

Table 4. IV estimation of MSAR Model

$$\mathbf{Y} = \boldsymbol{\alpha_0} + \boldsymbol{\rho} \cdot_2 \mathbf{D} \cdot_1 \mathbf{Y} + \boldsymbol{\beta_1}\mathbf{NGD} + \boldsymbol{\beta_2}\mathbf{HR} + \boldsymbol{\beta_3}\mathbf{NEW} + \boldsymbol{\theta_1}\mathbf{WNGD} + \boldsymbol{\theta_2}\mathbf{WHR} + \boldsymbol{u}$$

Variable	Coefficient	Std.	t-Stat	p-value
ρ_1	0.7762	0.1372	5.6561	< 0.0001
ρ_2	0.3324	0.0922	3.6054	0.0004
ρ_3	-0.3294	0.1084	-3.0372	0.0027
Const.	1.4373	0.7085	2.0287	0.0437
NGD	0.4592	0.0979	4.6879	< 0.0001
HR	0.9076	0.0743	12.2178	< 0.0001
NEW	-0.1375	0.0537	-2.5636	0.011
WNGD	-0.3629	0.1237	-2.9352	0.0037
WHR	-0.7575	0.1332	-5.6871	< 0.0001
$\boldsymbol{F_{relative\ to\ SDM}}$	6.7912			0.0014
S-H	64.6751			0.1746
$\overline{\boldsymbol{R}}^2 = \mathbf{0.81}$				

Table 4 provides the estimation results of the Multidimensional Spatial Autoregressive Model (MSAR) using IV estimation method described in Section 3[19]. First, all coefficients associated with the explanatory variables of the model appear significant at the 5% confidence level. This suggests that per capita income in the EU regions is significantly explained by the explanatory variables included and their spatial lags. Most importantly, the significance of the spatial coefficients validates the assumed complex structure of the interregional interactions. Furthermore, let us notice that the goodness of fit improves from 0.79 in the Spatial Durbin model to 0.81 in the MSAR model. The improvement of the model is also confirmed by the significant result of the F-test (6.79) in which SDM is used as the nested alternative. Additionally, the Sargan–Hansen test (64.68) confirms the validity of the instruments applied for the IV procedure.

The empirical outcomes show that human capital in a given region has a positive influence on regional per-capita income. At the same time, its spatial lag is found to have a negative impact on the surrounding regions. This suggests that, the level of human capital in neighbouring locations would affect negatively the level of per-capita income in a given region. One possible explanation for this negative impact may relate to the finite quantity of human resources available in neighbouring regions. Accordingly, an increase in the level of available human resources in one region would, to a

[19] It is should be emphasised that the matrix $(\boldsymbol{I} - \widehat{\boldsymbol{\rho}}\mathbf{D})$ is invertible and its inverse is the sum of appropriate geometric series in $\widehat{\boldsymbol{\rho}}\mathbf{D}$ as it can be verified that the obtained value for $\widehat{\boldsymbol{\rho}}$ is a member of the parameter space, which has been identified by the method described in the Appendix.

greater extent, be explained by inter-regional migration, draining human capital from the neighbours.

The results reveal that the overall contagion or clustering effects, captured by the three spatial coefficients ρ_i, exist and the use of matrix $\mathbf{D}$ enables to establish a more complex form of a spatial dependence across regions. The spatial coefficient ρ_1 is positive (0.78) and significant, confirming the existence of pure geographical dependence. Therefore, the level of *per-capita* income in the neighbouring regions is found to have a positive impact on the level of *per-capita* income in a given location. This result also concurs with positive per-capita income spillovers found in Özyurt and Dées (2015), using a similar theoretical framework and data for a more recent period. Moreover, the positive value of the second spatial parameter ρ_2 (0.33) associated with the variable Y suggests an important novelty to the interpretation of the magnitude of the autoregressive effect. In fact, provided that a region i has a neighbour or neighbours with considerably larger (or at least equally large) economy, then the spillover effect from those 'larger' neighbours turns out to be even greater (as for the effective value of the coefficient) than from the rest of the neighbours of i. Although this finding might have been anticipated intuitively, it would not be possible to capture it numerically through the classical one-matrix approach. By construction, standard spatial matrices gauge the relation between the value of a variable in a spatial unit and a mean value over all its neighbours. All in all, the particular use of higher-dimensional weight matrix enables to quantify the power of (conditional) spatial effect between regions of different 'economic size', and in a particular direction. Most importantly, in the MSAR model, the magnitude of the basic spatial coefficient ρ_1 (0.78) is lower than what is found in the SDM specification (0.84 for $\mathbf{W}Y$). Apparently, the inclusion of the ρ_2 coefficient (and the corresponding matrix) enables to capture the full impact of the alternative types of spatial interactions, which would be, otherwise, only partly captured by the traditional spatial auto-regressive coefficient. In a sense, the spatial coefficient in the SDM specification overestimates the 'true' magnitude of spatial interactions by overlooking the multidimensional aspect of spatial interactions.

In light of the significant parameter ρ_3, we conclude that the spatial lag with respect to regions with richer human capital ($\mathbf{D}_3\mathrm{Y}$), provides meaningful explanation of the remainder (unexplained by other terms) variance of the dependent variable. Moreover the relations turn out to be negative ($\widehat{\rho_3} = -0.33$). This implies that there is a statistically significant effect of negative difference in the strength of the spatial spillover (in terms of spatial dependence of Y) from regions with higher to regions with lower level of human capital, *ceteris paribus*, compared to the dependency in the opposite direction. In other words, once a great part of usual spatial interactions is explained with ρ_1 and ρ_2, the value of $\widehat{\rho_3}$ tells us that having a human capital richer neighbour has a detrimental effect on the value of Y in the given region. This might be attributed to a similar effect (e.g. competition for skilled employees) as in case of negative coefficient corresponding to spatially lagged human capital. An important difference, however, is that the inclusion of the $\mathbf{W}\mathrm{HR}$ variable captures the role of the mean value of HR across neighbours, while $\mathbf{D}_3$ considers the explanatory role of the dependant variable (Y) itself in the HR richer neighbours. It should also be noted that the construction of the multidimensional weight matrix implies that the interpretation of the parameter ρ_3 should be conditional. Namely, the inference regarding ρ_3 for a given region is valid if the region has a human capital richer neighbour, otherwise the value of $\widehat{\rho_3}$ should be ignored.

As a last point, we would like to discuss how our model specification addresses the endogeneity of the spatially lagged dependant variable and the potential endogeneity of the spatial weight matrix. The first problem is handled by using the multidimensional version of a fundamental IV approach, developed by Kelejian and Prucha (1998). As regards the second issue, we build the spatial weights matrix and the instruments using the threshold model. The intuition behind this technique has been already explained earlier, in Section 4.2. Briefly speaking, we follow the idea put forward by Kelejian and Piras (2014) where they state that if the spatial weight matrix W is non-exogenous, one can approximately represent it as a function of other variables p and q.

i.e.: $\mathbf{W} \approx f(p, q) + \zeta$, where p is believed to be exogenous and ζ corresponds to some, in a sense small, disturbance term. In line with this, Kelejian and Piras suggest building the set of instruments for IV estimation based on variable(s) p and the recognised form of the function f (e.g. its linear coefficients). Though we do not employ this approach, our technique of thresholding can be roughly described in the same setting, as we adjust the form of the nonlinear function f (the threshold parameter) in such a way that possible endogeneity is removed. Let us point out, that the result of the Sargan–Hansen test also confirms the validity (sufficient non-endogeneity) of instruments applied in the IV procedure.

Concluding Remarks

This paper presents a novel approach to deal with complex spatial dependence patterns. It introduces -a rather intuitive- concept of multi-dimensional spatial weight matrix to capture more complex, asymmetric spatial interactions, including spillovers that are not necessarily conditioned by geographic distance.

Empirical results from our proposed multi-dimensional spatial auto-regressive (MSAR) model are encouraging. As opposed to the traditional distance-based weighting scheme, the MSAR model includes a three-dimensional spatial weight matrix based on geographical distance as well as on socio-economic factors (e.g. economic size and human capital endowment). Judging by the result of the relative F-test (Table 4.), the MSAR model performs significantly better in explaining per capita income compared to the traditional Spatial Durbin Model (SDM). Most importantly, the significance of the spatial coefficients validates the assumed complex structure of the interregional interactions. The associated parameters with the different layers of the weight matrix suggest a correction of the magnitude of the autoregressive effect. More precisely, the additional dimensions of the weight matrix enable to capture the full impact of the alternative types of spatial interactions, which would be, otherwise, only partly captured by the traditional spatial auto-regressive coefficient. Moreover, our findings suggest that the spatial coefficient in the SDM specification may overestimate the 'true' magnitude of spatial interactions, by overlooking the multidimensionality of spatial interactions.

All in all, the multi-dimensional approach proves to be more adequate to deal with complex spatial patterns. Going forward, we believe further investigation of the spatial autoregressive effect following a multi-dimensional line of thinking can make a significant contribution to the spatial literature. Importantly, there is for a clear need for the development of the additional methodological tools to extend the range of available procedures of estimation and validation of multidimensional models.

References

Anselin, L. (1988). "Spatial Econometrics: Methods and Models". Kluwer, Dordrecht.

Anselin, L., Bera, A.K. (1998). "Spatial dependence in linear regression models with an introduction to spatial econometrics." In: Ullah, A., Giles, D. (Eds.). *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, 237–289.

Anselin, L., Smirnov, O. (1996). "Efficient algorithms for constructing proper higher order spatial lag operators." *Journal of Regional Science* 36, 67–89.

Badinger H., Egger P. (2008). „GM Estimation of Higher-Order Spatial Autoregressive Processes in Cross-Section Models with Heteroskedastic Disturbances." *CESifo Working Paper Series* No. 2356. Available at SSRN: http://ssrn.com/abstract=1173662

Besner, C. (2002). "A Spatial Autoregressive Specification with a Comparable Sales Weighting Scheme." *Journal of Real Estate Research* 24, 193–212.

Blommestein, H. (1983). "Specification and estimation of spatial econometric models: a discussion of alternative strategies for spatial economic modelling." *Regional Science and Urban Economics* 13, 250–271.

Blommestein, H. (1985). "Elimination of circular routes in spatial dynamic regression equations." *Regional Science and Urban Economics* 15, 121–130.

Bodson, P., Peeters, D. (1975). "Estimations of the Coefficients of a Linear Regression in the Presence of Spatial Autocorrelation: An Application to a Belgian Labour-Demand Function." *Environment and Planning* 7(4), 455–72.

Boschma, R. (2005). „Proximity and Innovation: A Critical Assessment.” *Regional Studies* 39, 61–74.

Brandsma, A.S., Ketellapper, R.H. (1979). “A biparametric approach to spatial autocorrelation.” *Environment and Planning* 11(1), 51–58.

Cliff, A., Ord, J.K. (1981). *Spatial Processes, Models and Applications*. London: Pion.

Dacey, M.F. (1968). “A Review on Measures of Contiguity for Two and k-Color Maps”. *Technical Report No.2, Spatial Diffusion Study*, Department of Geography, Evanston: Northwestern University.

Deng, M. (2008). “An anisotropic Model For Spatial Processes.” *Geographical Analysis* 40(1), 26-51.

Elhorst, J. (2008). „A spatiotemporal analysis of aggregate labour force behaviour by sex and age across the European Union.” *Journal of Geographical Systems* 10, 167–190.

Elhorst J., Lacombe D. J., Piras G. (2012), “On model specification and parameter space definitions in higher order spatial econometric models” *Regional Science and Urban Economics* 42 (2012), 211 – 220

EUROSTAT (2002). European regional Statistics. Reference guide, European Communities, Luxembourg.

Fingleton, B. (2001). “Theoretical economic geography and spatial econometrics: dynamic perspectives.” *Journal of Economic Geography* 1(2), 201-225.

Fujita, M., Krugman, P., Venables, A. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge MA: MIT Press.

Gelfand, A.E., Ecker, M.D., Knight, J.R., Sirmans, C.F. (2004). “The Dynamics of Location in Home Price.” *Journal of Real Estate Finance and Economics* 29(2), 149-166.

Getis, A., Aldstadt, J. (2004). “Constructing the Spatial Weights Matrix Using A Local Statistic.” *Geographical Analysis*, 36, 90–114.

Giacomini, R., Granger, C.W.J. (2004). “Aggregation of space-time processes.” *Journal of Econometrics* 118, 7–26.

Gupta, A., Robinson, P.M. (2015). “Inference on higher-order spatial autoregressive models with increasingly many parameters." Journal of Econometrics 186, 19-31.

Han, X., Hsiehc, C.S., Lee, L.F. (2017). “Estimation and model selection of higher-order spatial autoregressive model: An efficient Bayesian approach.” *Regional Science and Urban Economics* 63, 97-120.

Horn R. A., Johnson C. R., Matrix Analysis. Cambridge etc., Cambridge University Press (1985)

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

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

Lauridsen, J. (2006). “Spatial autoregressively distributed lag models: equivalent forms, estimation and an illustrative commuting model.” *The Annals of Regional Science* 40, 297-311.

Lee L. F., and Liu X. (2010). “Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances.” *Econometric Theory* 26, 187-230.

Lee L. F., and Liu X. (2014). ”Efficient GMM estimation of spatial dynamic panel data models with fixed effects," *Journal of Econometrics*, Elsevier, vol. 180(2), pages 174-197.

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

Mankiw, G. N., D.Romer and D.Weil (1992), “A contribution to the empirics of economic growth”, *Quarterly Journal of Economics*, 107, No. 2.

Olejnik, A. (2008). “Using the spatial autoregressively distributed lag model in assessing the regional convergence of per-capita income in the EU25.” *Papers in Regional Science* 87/3, 371-384.

Olejnik, A. (2013). “Assessing the space-time structure with a multidimensional perspective.” *Acta Universitatis Lodziensis*, *Folia Oeconomica,* 292, 2013.

Özyurt S. and Dées S. (2015), Regional dynamics of economic performance in the EU: To what extent spatial spillovers matter?, *ECB Working Paper Series*, no 1870 / December 2015

Ord, J.K., Getis, A. (1995). “Local Spatial Autocorrelation Statistics: Distributional Issues and an Application.” *Geographical Analysis* 27, 286-306.

Saint Raymond, X. (1991). “Elementary Introduction to the Theory of Pseudodifferential Operators”. CRC Press.

Suijlekom, W.D. Van. "Multi-Index Notation." From *MathWorld*--A Wolfram Web Resource, created by Eric W. Weisstein. http://mathworld.wolfram.com/Multi-IndexNotation.html

Takesaki M., “Theory of Operator Algebras I”, Berlin-Heidelberg-New York, Springer-Verlag (1979)