Spatial Aggregation-Repulsion-Diffusion model: theory and estimation

Davide Fiaschi — University of Pisa
Angela Parenti — University of Pisa
Cristiano Ricci — University of Pisa

March 15, 2023

To respect the size limit in the submission, figures are available at
https://poisson.phc.dm.unipi.it/~cricci/fiaschi_parenti_ricci_SARD.pdf

Abstract

In this paper, we propose a novel spatial econometric model, denoted as Spatial Aggregation-Repulsion-Diffusion (SARD). The model is derived as the approximation of a class of micro-founded growth models in a diversified space, which encompasses both local accumulation and agglomerative, repulsive and diffusive forces driving spatial factor reallocation. The estimate of the SARD model for the income of Italian municipalities over the period 2014-2019 supports the main predictions derived by the theory and outperforms the most common spatial econometric models used in the literature.

JEL Classification Numbers: C21, C60, R12

Keywords: spatial spillovers, spatial econometrics, aggregation-diffusion equation, partial differential equations, generalized finite-difference method

Contents

1 Introduction

The investigation of the causes of the spatial distribution of regional economic activity is a very debated issue in the literature (Quah, 2002). Several economic geography models have been introduced to explain the emergence of spatial patterns characterized by geographical aggregations of economic activities and well-defined locations of these aggregations relative to one another, i.e. the size distribution of cities and their spatial distribution (Krugman, 1991). Despite an increasing literature, a definitive explanation of how aggregation economies emerge from the behaviour of individual agents is still needed (Rossi-Hansberg, 2019). In particular, the exploration of the micro-foundations of regional economics seems to be the new challenge for a better understanding of the real world, in the light of the increasing availability of accurate data at fine geographical scale (Allen and Arkolakis, 2014; Desmet et al., 2018). Furthermore, the complex nature of the distribution of economic activity across space and of its evolution through time, requires the development of specific quantitative techniques. On the one hand, the standard spatial econometric models have been used in the literature to capture the correlation structure in the equilibrium outcome resulting from social/spatial interaction (e.g., Anselin, 2002; Brueckner, 2003; Ertur and Koch, 2007; Xu and Lee, 2019; Combes and Gobillon, 2015). On the other hand, geostatistical models have focused the on spatial process underlying observations rather than on the interaction among observations (Cressie and Wikle, 2015).

In this paper, we propose a new class of theory-driven spatial econometric models, denoted SARD model, based on a spatial growth model where the collective macroscopic behaviour emerges from the dynamic of interacting and locally optimizing agents, and the macroeconomic dynamic is expressed by a Partial Differential Equation (PDE). The spatial growth model is defined in continuous space and time and belongs to the class of *Aggregation-Diffusion Equations* (ADEs), which over the past 20 years have been employed in several biological applications and stimulated many mathematical works (see Carrillo et al., 2019a for a review). In particular, the competing effect of aggregation, repulsion and diffusion leads to several interesting properties, such as metastability, symmetrization over time, and non-uniqueness in equilibrium solutions. A precise characterization of equilibria is not readily available: in the general case (Carrillo et al., 2019b) proved only that equilibria are radially symmetric in absence of exogenous elements. In the special case where the interaction kernel is of logarithmic type, the authors also proved the uniqueness of the equilibrium as well as the convergence to the unique equilibrium independently on the initial configuration. Moreover, another remarkable feature of ADEs is that they can be obtained as the mean-field limit of a system of discrete interacting agents, as the number of agents becomes large.

To bring the model to the data we have to employ a discretization technique, which allows deriving a discrete space-time counterpart of our continuous space-time model. In particular, the non-uniform spatial structure of the observed economic activities and demographic variables, that is the fact that they are not arranged on a regular cartesian grid but into irregular cells, such as administrative regions, imposes the use of a discretization technique, i.e. *generalised finite difference methods*, to transpose the system of PDEs, which are defined in the continuum, into a set of discrete locations (Jensen, 1972). Given that the theoretical model is developed in continuous space, the derivation of the empirical model imposes some limitations from the point of view of the data to be used since there is an intermediate discretization step from continuous to discrete locations. These limitations are discussed in Section 3.2. This constitutes one of the main differences to the already mentioned literature of spatial econometrics, where locations are assumed to be discrete by the start.

Even though in the end we arrive at a discrete space-time model, the formalism of PDEs in continuous space provides an innovative tool to introduce new elements in the theory of spatial econometrics. In particular, as described in more detail in Section 2, it allows constructing regressors having the remarkable feature of being cross-sectional zero-sum by construction, that is they do not contribute to the overall growth of the variable of interest when summing up across all available locations. This yields an initial disentangle between individual growth and pure reallocation. Moreover, by exploiting the presence of differentials in the level of the variables considered across neighbouring locations, instead of simply measuring the effect of spatial correlation through local averages, our regressors are also able to distinguish between the *aggregating*[1] effect towards more appealing

[1]With the word *aggregation* in this paper we will always refer to the concept of *spatial aggregation*, that is the tendency of economic agents to spatially relocate to be closer to each other.

locations (centripetal force) arising for example from the higher wages in central labour markets, and the *diffusive* effect (centrifugal force) caused for example by the higher cost of living in a congestioned location.

We emphasize that in this paper we are not proposing a methodology to find the best fit for our PDE model but, instead, we aim to introduce a new spatial econometric model following a theory-driven approach from individual agents to macroscopic behaviour. In particular, we derive new expressions for linear regressors which are obtained from the discretization of a spatial theoretical model expressed with the language of PDE. Moreover, the specific form of our model (linear in the parameters to be estimated), also allows us to study directly the process in the transient regimes, i.e. not assuming to be in a neighbour of the equilibrium, therefore without relying on a (log-)linearization around the steady state (Ertur and Koch, 2007). Since the final econometric model can be estimated via OLS, our methodology doesn't suffer from the typical computational issues of spatial econometrics, such as matrix inversion or constrained optimization (Anselin and Rey, 2012). In our case, the computational complexity of the construction of spatial differential matrices, described in detail in Appendix B, still is influenced by the size of the problem. However, it scales only linearly in the number of observations, therefore not constituting a problem even for large datasets.

The paper is organized as follows: Section 2 presents the spatial growth model expressed by a PDE; Section 3 the methodology to derive a linear econometric model by a PDE; Section 4 discuss the empirical application to Italian municipalities; and, finally, Section 5 concludes. Technical details are gathered in Appendix.

2 A prototype model of spatial growth

This section introduces a class of spatial growth models encompassing the main features present in the literature, i.e. the spatial non uniformity caused by geographical and socio-economic factors, the spatial aggregation of economic activity driven by positive spatial spillovers, the centrifugal dynamics driven by congestion, and, finally, the existence of randomness in the individuals' movement (Fujita and Thisse, 2002).

Let $y(t, z)$ be the variable of interest of our model in location z at time t, e.g. municipal total income per square kilometre, where $\Omega \subseteq \mathbb{R}^2$. Each point $z \in \Omega$ is identified by two components $z = (z_1, z_2)$, i.e. latitude and longitude. The dynamics of $y(t, z)$ is assumed to obey the following partial differential equation (PDE):

$$
\begin{aligned}
\partial_t y(t,z) &= \varphi\left(y(t,z)\right) + \\
&+ \gamma_S \mathrm{div}_z \left(y(t,z)\nabla_z S(z)\right) + \\
&+ \gamma_A \mathrm{div}_z \left(y(t,z)\nabla_z \left(K_{h_A} * y\right)(t,z)\right) + \\
&+ \gamma_R \mathrm{div}_z \left(y(t,z)\nabla_z \left(K_{h_R} * y\right)(t,z)\right) + \\
&+ \gamma_D \Delta_z y(t,z). \qquad (1)
\end{aligned}
$$

Appendix A briefly discusses how Eq. (1) can be derived in the limit of an infinity-agent economy as the outcome of the agents' mobility driven by locally maximizing behaviour (see Fiaschi and Ricci, 2023, for a more detailed analysis). Eq. (1) makes a wide use of concepts used in PDE literature, which will be discussed in details below in relation to the different features we are interested to model.

Firstly, the variable on the left-hand side $\partial_t y(t,x)$ represents the total variation of the quantity $y(t,z)$ at time t in location z, expressed through the use of the partial derivative ∂_t with respect to time. The first term on the right-hand side of Eq. (1) $\varphi\left(y(t,z)\right)$ represents the change in $y(t,z)$ due to the *local endogenous process of accumulation*, i.e. the impact of the endogenous variable $y(t,z)$ on the time change of the variable itself, independent of any spatial interaction. Taking as reference the Solovian model, the shape of φ should reflect the *shape of production function*, *saving behaviour*, *depreciation rates of factors* and *employment growth*. In particular, the presence of *increasing returns* in the production function; the presence of a nonlinear relationship between saving rates and income, could make φ' not decreasing and non-monotone (see, e.g., Fiaschi and Lavezzi, 2007).

The main inspiration behind the terms from the second to the fifth of the right-hand side appearing in Eq. (1) is the so-called *Fokker-Planck Equation*, that describes the time evolution of the probability density of a stochastic process obeying a *Stochastic Differential Equation* (SDE). They contain *differential operators*, which belong to the language of PDEs, and are particularly effective for describing the various sources of change in the spatial distribution of $y(t,z)$. In particular, div_z is the *divergence operator*, i.e. $\mathrm{div}_z f \equiv \partial_{z_1} f_{z_1} + \partial_{z_2} f_{z_2}$, where $f = (f_{z_1}, f_{z_2})$ is any vector field $f : \Omega \to \mathbb{R}^2$ and ∂_{z_i} are the partial derivatives with respect to each of the components of $z = (z_1, z_2)$. All the partial derivatives expressed above are to be intended in the spatial sense, i.e. they measure how the spatial profile is varying along one of the two components, either latitude or longitude. The term $\nabla_z f$ stands for the *gradient operator,* i.e. $\nabla_z f \equiv (\partial_{z_1} f, \partial_{z_2} f)$ while finally the term $\Delta_z f \equiv \partial^2_{z_1 z_1} f + \partial^2_{z_2 z_2} f$ is called the *Laplace operator* and involves the second order pure partial derivatives $\partial^2_{z_1 z_1}$ and $\partial^2_{z_2 z_2}$.

Since partial derivatives express a measure of differentials across different locations, the sign of their coefficients reflects the direction driving the reallocation. For example, as the second term, if the coefficient γ_S is negative, then the term $\gamma_S \mathrm{div}_z \left(y(t,z)\nabla_z S(z)\right)$ expresses the tendency of $y(t,x)$ to increase in those locations where $S(z)$ is higher, and to decrease of the same amount where $S(z)$ is lower.[2] The definition of higher and lower is to be intended as a relative comparison between locations, and is provided by the use of $\nabla_z S(z)$ which measures the *steepness of the transition* moving from one location to the other. This increase-decrease effect is made in such a way that the total amount of mass in the distribution y which is gained where it increases, it is lost in the other locations. Hence, the total variation accounted by the second term, when one takes into account all the locations, is zero (neglecting possible boundary effects). For this reason, we refer to all the terms in Eq. (1) except for the first one, as purely *reallocation* term, since they don't affect the total amount of y.

The second term of Eq. (1) $\gamma_S \mathrm{div}_z \left(y(t,z)\nabla_z S(z)\right)$ is introduced to take into account the *topography* of Ω. In particular, $\nabla_z S(z)$ indicates the possible pure geographical and exogenous advantage to move from location z to neighbouring locations. In this regard, we assume the coefficient $\gamma_S < 0$ so that the reallocation follows the peaks of the exogenous variable $S(z)$.

The third term on the right-hand side of Eq. (1) $\gamma_A \mathrm{div}_z \left(y(t,z)\nabla_z \left(K_{h_A} * y\right)(t,z)\right)$

[2] A simple example in one dimension showing the interpretation of the negative sign is the following: consider the case where the distribution $y(z)$ is constant, equal to some given value $\overline{y}$, and the function $S(z)$ is a sigmoid of the form $S(z) = \frac{1}{1+e^{-z}}$. In this case, the second term of Eq. (1) becomes $\gamma_S \overline{y} S''(z)$. For $z > 0$ the second derivative of S is negative, while for $z < 0$ is positive. Therefore, if $\gamma_S < 0$ this term is positive when $z > 0$, implying that the values of $y(z)$ are increasing, and negative for $z < 0$.

represents the effect of *aggregation* of y, i.e. the tendency of y to concentrate in specific locations. Appendix A shows that this term, and the next two, of Eq. (1) are the outcome of the agents' mobility driven by locally maximizing behaviour. The intensity of this process is measured by $\gamma_A < 0$. K_{h_A} is a kernel function with bandwidth $h \geq 0$, and $*$ is the *convolution operator*, i.e. $(K_{h_A} * y)(t,z) \equiv \int_\Omega K_{h_A}(k-z)\, y(t,k)dk$. In other words, $(K_{h_A} * y)(t,z)$ is the weighted sum of all y around location z at period t, where the weights are defined by kernel K_{h_A}; in particular, the shape of K_{h_A} and the value of h_A decide how these weights change with the distance from location z. Since $\nabla_z (K_{h_A} * y)(t,z)$ points to the directions where the level of $y(t,z)$ is higher when averaged among neighbours, it plays a similar role of $\nabla_z S(z)$ as the term on the second line of Eq. (1). An important difference is that instead of being purely exogenous as for $S(z)$, the direction which drives the reallocation is now endogenous, since it depends on the spatial distribution of $y(t,z)$. In particular, we assume $\gamma_A < 0$, so that the reallocation is driven towards those areas where the local average of y, $(K_{h_A} * y)(t,z)$ is higher, providing the intuition on why this term will tend to concentrate y over space. A standard explanation in economics is the observed process of aggregation of workers, i.e. the emergence of cities based on the positive externalities generated by working in places where other activities and/or skilled workers are already present (Fujita and Thisse, 1996; Krugman, 1998; Moretti, 2004).

The fourth term on the right-hand side of Eq. (1) $\gamma_R \operatorname{div}_z (y(t,z)\nabla_z (K_{h_R} * y)(t,z))$ represents the effect of *repulsion* of y across different locations, that is the tendency of y to flow away from locations with higher levels. The intensity of this process is measured by $\gamma_R > 0$. This term is exactly analogous to the one related to aggregation, except that the Kernel function K_{h_R} can be different from K_{h_A} (for example, in the speed of decay as a function of the distance), and that the sign γ_R is assumed to be positive, so the effect instead of being centripetal is centrifugal.

Finally, the last term on the right-hand side of Eq. (1) $\gamma_D \Delta_z y(t,z)$, represents the effect of *diffusion* across different locations of y, which tends to uniformly spread y over space. The intensity of this diffusion process in location z at time t depends on the parameter $\gamma_D > 0$, and on the sign and magnitude of second derivatives of $y(t,z)$.[3] In economics,

[3]The use of the second derivative can be understood intuitively if one thinks of the one-dimensional example of a bell-shaped distribution of $y(t,z)$ (e.g. a gaussian distribution). In this case, the spatial second derivative measure the convexity/concavity of the distribution, being negative in the centre of the

the diffusion process can be justified by the random component, generally determined by hidden characteristics/preferences of agents, in the agents' choice (Wozniak, 2010).

While is it true that both *repulsion* and *diffusion* represent centrifugal forces, they are very different in nature. While the repulsion effect only takes place in presence of overcrowding, the diffusion instead is always affecting the dynamics. In particular, diffusion always tends to equalize all the levels of $y(t, z)$ across locations, while repulsion only lowers the level of y where density is too high. In other words, in the same way that aggregation expresses the tendency of individuals to relocate to be closer to each other, repulsion can be seen as an effect due to *congestion*, where individuals aim to avoid overcrowded locations. In economics, a possible source of the observed outflows of individuals from very crowded locations can arise from the higher housing prices and, in general, from the higher cost of living in locations with high population density (Krugman, 1998). Moreover, this tendency is generally justified by the factor-return equalization across different locations in presence of decreasing marginal returns to factors.[4]

Summarizing, the coefficients of the Model (1) should respect the following constraints, which we will check in the estimation:

1. $\varphi' > 0$ and $\varphi'' < 0$, to reflect the shape of production functions and saving behaviour;

2. $\gamma_S < 0, \gamma_A < 0$, to reflect the reallocation due to topography and the aggregation effect respectively; and

3. $\gamma_R > 0, \gamma_D > 0$ to reflect the repulsive effect and the presence of diffusion.

3 From theory to empirics

The estimate of Eq. (1) is not trivial. The typical approach used in growth empirics, i.e. a log-linear approximation around the steady state/long-run equilibrium, derives a

peak (convex) and positive in the area outside the peaks (concave). Therefore, the effect when $\gamma_D > 0$ is to decrease where the distribution is convex (i.e. on the local peaks) and to increase in all other regions. Farlow (1993, p. 12) provides an intuition of why the second derivative is crucial for describing a diffusion process, which tends to uniformly spread the variable of interest over space.

[4]Suppose to consider two locations, 1 and 2, with a different endowment of capital $k_1 > k_2$ but with the same production function; then $f'(k_1) < f'(k_2)$ under the hypothesis of $f''(\cdot) < 0$; with free movement of capital we should observe a flow of capital from location 1 to location 2. The second derivative with respect to the distribution over space of capital is a proxy for the difference in the level of k_1 and k_2, which, in turn, is reflected in the difference in factor returns and, hence, in the intensity of reallocation.

reduced-form model that is linear with respect to the parameters. Hence, any interesitng out-of-equilibirum dynamics is excluded by the analysis. In our case, we show that the econometric model derived by Eq. (1) is already linear by construction and, therefore, our estimate encompasses all possible dynamics.

Geographical space can be modelled either as a continuous plane in two dimensions (longitude and latitude) or as a set of discrete locations connected by arcs (Mendes and Mendes, 2015). The continuous specification does not easily fit the actual data which is a discrete representation. For this reason, standard spatial econometrics uses a discrete representation whose core is the specification of the weights matrix which defines the strength of the interaction among spatial units. In particular, the weight matrix is used to represent which spatial units are neighbours to each other and might be specified in several ways, for example using geographical distance, contiguity, economic or cultural distance (see LeSage and Pace, 2009). The choice of the matrix is crucial in the spatial analysis and should be driven by the "problem being modelled, and perhaps particular additional non-sample information which may be available" (Mendes and Mendes, 2015).

As opposed to the spatial econometric literature that uses only the weight matrix to exploit the geographical structure, we propose to use such structure also to compute an approximation of partial derivatives, therefore measuring differentials of the variable of interest across neighbouring locations. The field of numerical methods for PDEs is very broad and provides a plethora of techniques to approximate partial derivatives numerically to compute approximate solutions. These techniques range from the simpler ones based on *finite difference method* on a regular Cartesian grid (Ames, 2014), to those developed to work with irregular grids, as in our case. For the case of irregular grids, two main approaches can be used: the so-called *finite element methods*, broadly applied in physics and engineering, where partial derivatives are approximated by using a discrete meshing of the domain (e.g. triangulations) and exploiting the weak formulation of the PDE; and, the *generalized finite difference methods* that use a Taylor approximation to compute partial derivatives even on irregular grids (Jensen, 1972). We follow the second approach (see Appendix B for a self-contained overview).

3.1 The SARD econometric model

To understand the derivation of the econometric model we have to approximate the spatial distribution $y(t,z)$ into a set of finite locations. We denote $y_{ti} \equiv y(t,z_i)$ with z_i for $i = 1, \dots N$ a discrete set of locations. Moreover in the special case where $t = 0$, we denote $y_i \equiv y(0,z_i)$. Let us start by discretizing time. Therefore we approximate the time partial derivative ∂_t by its discrete counterpart $\Delta_t y_i \equiv y_{ti} - y_i$ (assuming a time step of unitary size). Secondly, let us focus on the much bigger task of discretizing space, focusing on the second term of Eq. (1), $\gamma_S \text{div}_z (y(t,z)\nabla_z S(z))$. Expanding the expressions for div_z and ∇_z as defined in Section 2 by making partial derivatives with respect to (z_1, z_2) explicit, we have:

$$\gamma_S \text{div}_z (y(t,z)\nabla_z S(z)) = \gamma_S \left[\partial_{z_1} (y(t,z)\partial_{z_1} S(z)) + \partial_{z_2} (y(t,z)\partial_{z_2} S(z))\right]. \tag{2}$$

Hence, we first approximate partial derivatives across space with respect to (z_1, z_2) as described in Appendix B. Therefore we introduce the matrices $M_{z_1}, M_{z_2}, M_{z_1 z_1}, M_{z_1 z_2}$ to approximate respectively the space partial derivatives $\partial_{z_1}, \partial_{z_2}, \partial^2_{z_1 z_1}, \partial^2_{z_2 z_2}$, so that, if one considers locations z_i one has:

$$[\partial_{z_1} y(t,z)]_{z=z_i} \approx (M_{z_1}\mathbf{y}_t)_i, \quad [\partial_{z_2} y(t,z)]_{z=z_i} \approx (M_{z_2}\mathbf{y}_t)_i, \tag{3}$$

$$[\partial^2_{z_1 z_1} y(t,z)]_{z=z_i} \approx (M_{z_1 z_1}\mathbf{y}_t)_i, \quad [\partial^2_{z_2 z_2} y(t,z)]_{z=z_i} \approx (M_{z_2 z_2}\mathbf{y}_t)_i. \tag{4}$$

Putting together Eqq. (2),(3) and (4) we derive the discrete counterpart of the second term in Eq. (1):

$$\gamma_S \text{div}_z (y(t,z)\nabla_z S(z)) \approx \gamma_S \left[M_{z_1} (\mathbf{y}_t \odot M_{z_1}\mathbf{s}) + M_{z_2} (\mathbf{y}_t \odot M_{z_2}\mathbf{s})\right]_i,$$

with $s_i = S(z_i)$ and where $\odot$ denotes the element-wise product between vectors.

Therefore, applying the procedure just described to all the expressions in Eq. (1) we get the *PDE-derived econometric model* in the finite-state space and in finite time, which we label as *SARD* - Spatial Aggregation Diffusion (Repulsion) - and which is given by:

$$
\begin{aligned}
\Delta_t y_i &= \varphi_1 y_i + \varphi_2 y_i^2 + \\
&+ \gamma_S \underbrace{\left[M_{z_1}\left(\mathbf{y} \odot M_{z_1}\mathbf{s}\right) + M_{z_2}\left(\mathbf{y} \odot M_{z_2}\mathbf{s}\right)\right]_i}_{x_{Si}} + \\
&+ \gamma_A \underbrace{\left[M_{z_1}\left(\mathbf{y} \odot M_{z_1} W_{K_A}\mathbf{y}\right) + M_{z_2}\left(\mathbf{y} \odot M_{z_2} W_{K_A}\mathbf{y}\right)\right]_i}_{x_{Ai}} + \\
&+ \gamma_R \underbrace{\left[M_{z_1}\left(\mathbf{y} \odot M_{z_1} W_{K_R}\mathbf{y}\right) + M_{z_2}\left(\mathbf{y} \odot M_{z_2} W_{K_R}\mathbf{y}\right)\right]_i}_{x_{Ri}} + \\
&+ \gamma_D \underbrace{\left[(M_{z_1 z_1} + M_{z_2 z_2})\mathbf{y}\right]_i}_{x_{Di}} + \\
&+ \epsilon_i,
\end{aligned}
\tag{5}
$$

where i is the index of the unit, with $i = 1, \cdots, N$, $\varphi(\cdot)$ is defined as $\varphi(y_i) = \varphi_1 y_i + \varphi_2 y_i^2$ and $\mathbf{y}$ is the vector of the variable of interest in the first period of observation. M_{z_1}, M_{z_2}, $M_{z_1 z_1}$ and $M_{z_2 z_2}$ are matrices $N \times N$ used to calculate the approximation of the first and second derivatives of distribution of the variable of interest in space respectively (see Appendix B for a complete discussion), W_{K_A} and W_{K_R} are the matrices $N \times N$ representing the kernel K_{h_A} and K_{h_R} respectively, and $\mathbf{s}$ the vector of exogenous geographical components. Finally, ϵ_i is the stochastic component for unit i.

In matrix form, the empirical model can be rewritten as:

$$\Delta_t \mathbf{y} = \varphi_1 \mathbf{y} + \varphi_2 \mathbf{y}^2 + \gamma_S \mathbf{x}_S + \gamma_A \mathbf{x}_A + \gamma_R \mathbf{x}_R + \gamma_D \mathbf{x}_D + \epsilon \tag{6}$$

Given that the econometric model is derived from Eq. (1), we expect the following signs for the estimated parameter: i) $\gamma_1 > 0$ and $\gamma_2 < 0$ which should reflect the positive and marginally decreasing impact of initial conditions on the variation of the variable of interest over time; ii) $\gamma_S < 0$ representing the spatial effect of reallocating towards more appealing locations iii) $\gamma_A < 0$ indicating evidence of aggregation effect over space for the variable of interest; iv) $\gamma_D > 0$, highlighting a diffusion effect over space of the variable of interest y; finally, v) $\gamma_R > 0$ showing evidence of repulsion effect.

3.2 Some caveats to our approach

The methodology proposed in the previous section, i.e. taking the discrete form of a continuous space model, has some intrinsic limitations that we now proceed to explain.

Firstly, we need to highlight that the discretization procedure inherently introduces a bias in the model. As one approximates a measure of differentials from a discrete set of values, a discretization error arises. However, discretization errors are proportional to the distance between the evaluation points, so if the discretization is fine enough the bias is limited and can be overlooked. If the geographical resolution, i.e. the distance between georeferenced spatial locations, is too rough, then the model cannot be applied profitably. To make a concrete example, the current methodology would most likely fail or produce meaningless results if one tries to apply it to the GDP distribution at the national level since countries constitute a too-large unit of aggregation. Also, this framework is currently too general to provide a precise rule of whether the resolution is high enough to employ the current model successfully. However, one possible indicator that one can look at to discern if the discretization bias is affecting the result is the sign of the estimated coefficients. Since the theoretical model in Eq. (1) prescribes the sign of each parameter, an estimated coefficient whose sign matches the expected one confirms the goodness of the approximation. Another important element to keep into account is the magnitude of the spatial correlation among the units of observation. If they are highly correlated, that is neighbouring locations don't differ too much in relative terms, then the bias induced from the discretization is limited even if locations are spread far apart. On the contrary, if the opposite holds the approximation may be very poor even when the units of observations are of small size.

Secondly, the present model also imposes limitations on the possible choice of the spatial matrix W_{K_A} (W_{K_R} analogously). The matrix W_{K_A} derives from the space discretization of the function K_{h_A}, appearing in Eq. (1). Accordingly, it requires to be defined in terms of *relative geographical distance* between locations. Therefore, in the current framework, we are not allowed to use spatial matrices which are based on contiguity between areas of observations like N-order contiguity or even K-nearest neighbourhood.

Thirdly, the theoretical framework also imposes the necessity to work with variables which are density-like, i.e. rescaled by the area of the unit of observation. This is again due to the fact that the econometric model is obtained by the discretization of a continuous space model.

These are currently drawbacks with respect to other spatial econometrics models, which do not require any relations between the relative size of observations between contiguous locations and their geographical distance. For this reason, the model proposed in this paper would be more suitable to be employed in all those cases where geographical data not affected by administrative limitations are available, for example for the Night-Lights VIIRS[5], or the more recent DataForGood[6], which are both defined in a uniformly spaced high-resolution grid. This kind of dataset has also the remarkable feature of having units of observations with the same geographical size so that there would be no need to rescale each variable.

4 Empirical application

In this section, we propose an application of the SARD model to the estimate of spatial distribution dynamics of economic activity in Italy. For this purpose, we decided to use the income of Italian municipalities, which represents a good proxy for economic activity with a detailed enough geographical scale.

4.1 Data on Italian municipalities

Italian Ministry of Economy and Finance (Agenzia delle Entrate) releases information at the municipal level (about 8000 in Italy) on the *nominal personal income* declared for tax purposes (IRPEF) for the period 2008-2020 from resident households.[7] During this period, however, Italy was hit by several shocks: the subprime mortgage crisis coming from the US in 2007-2011, the sovereign debt crisis started in 2011-2013, and the COVID pandemic in 2020. While the former was very asymmetric over the territory, the other two

[5]Visible Infrared Imaging Radiometer Suite (VIIRS), from the Suomi satellite launched in 2011 by NASA and the National Oceanic and Atmospheric Administration (NOAA), `https://eogdata.mines.edu/products/vnl/`.

[6]Data For Good, High-Resolution Population Density Maps from META `https://dataforgood.facebook.com/dfg/tools`.

[7]`https://www1.finanze.gov.it/finanze/pagina_dichiarazioni/public/dichiarazioni.php`

were more symmetric although very profound and leading to a period of high instability. We, therefore, decide to restrict our analysis to the period 2014-2019. As robustness, in Appendix D, we also include the 2020 year.[8]

For each municipality, the total declared incomes are divided by the size of the available land of the municipality, measured in Km^2, to make comparable the incomes of municipalities with very different sizes and different types of territory.[9] This choice of considering income density instead of income per capita is driven by the theoretical framework we are working on.

Figure 1 reports the map of total income per Km^2 of about 8000 Italian municipalities for the years 2014 and 2019 and its variation over the period. The general impression is of a high level of aggregation in the spatial distribution of income which seems to be not decreasing over time. We also observe a remarkable difference between the North and South part of Italy, as well as between inner areas and coastal areas.

In the estimate, we also use the data on the mean altimetry of the municipality coming from ISTAT as an exogenous variable.

4.2 Estimation results

Similarly to the definition of a weight matrix in the spatial model, for the empirical application, we need to define the kernels K_{h_A} and K_{h_R}, i.e. to set the bandwidth (specifying within which distance units can be considered as neighbours) and the weight of each neighbouring unit. In the estimation, we try different specifications of the kernels. In particular, we use an exponential distance-decay function, that is:

$$K_{h_A}(z) = \begin{cases} e^{-||z||\beta_A} & \text{if } ||z|| \leq h_A \\ 0 & \text{otherwise,} \end{cases} \quad w_{K_A ij} = \begin{cases} e^{-d_{ij}\beta_A} & \text{if } d_{ij} \leq h_A \ \forall i,j \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

$$K_{h_R}(z) = \begin{cases} e^{-||z||\beta_R} & \text{if } ||z|| \leq h_R \\ 0 & \text{otherwise,} \end{cases} \quad w_{K_R ij} = \begin{cases} e^{-d_{ij}\beta_R} & \text{if } d_{ij} \leq h_R \ \forall i,j \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

[8]Indeed, differently from the GDP in the personal income we cannot observe a high drop in 2020 for Italian municipalities. This discrepancy can be explained by the fact that many measures were introduced by the Italian government to support income over 2020.

[9]We, therefore, consider the income density of each municipality.

Figure 1: The map of total income per Km^2 of the available land of Italian Municipalities in 2014 and 2019.

pics/map_incomeKm2_2014.pdf

(a) 2014

pics/map_incomeKm2_2019.pdf

(b) 2019

pics/map_diff_incomeKm2_2014_2019.pdf

(c) 2014-2019

Source: Italian Ministry of Economy and Finance (Agenzia delle Entrate) and ISTAT (Italian National Institute of Statistics).

Notice that the discrete counterpart of the kernels K_{h_A} and K_{h_R}, i.e. W_{K_A} and W_{K_R}, are non-zero-diagonal differently from the standard spatial econometric literature.[10]

With this definition of the kernels K_{h_A} and K_{h_R} even if the two kernels are not the same, the resulting regressors turn out to be highly correlated. To avoid this issue we define two completely separated spatial kernels $K_{short} \equiv K_{h_R}$ and $K_{long} \equiv K_{h_A} - K_{h_R}$. This reformulation allows us to estimate the parameters of the original model in Eq. (5) by OLS, but separating the long from the short-range spatial effects, without the problem of high correlation between regressors. This implies the estimation of the following model:

$$\begin{aligned}
\Delta_t y_i &= \varphi_1 y_i + \varphi_2 y_i^2 + \\
&+ \gamma_S [M_{z_1} (\mathbf{y} \odot M_{z_1} \mathbf{s}) + M_{z_2} (\mathbf{y} \odot M_{z_2} \mathbf{s})]_i + \\
&+ \gamma_{long} \left[M_{z_1} \left(\mathbf{y} \odot M_{z_1} W_{K_{long}} \mathbf{y}\right) + M_{z_2} \left(\mathbf{y} \odot M_{z_2} W_{K_{long}} \mathbf{y}\right)\right]_i + \\
&+ \gamma_{short} [M_{z_1} (\mathbf{y} \odot M_{z_1} W_{K_{short}} \mathbf{y}) + M_{z_2} (\mathbf{y} \odot M_{z_2} W_{K_{short}} \mathbf{y})]_i + \\
&+ \gamma_D [(M_{z_1 z_1} + M_{z_2 z_2}) \mathbf{y}]_i + \\
&+ \epsilon_i, \qquad (9)
\end{aligned}$$

where $\gamma_{long} = \gamma_A$ and $\gamma_{short} = \gamma_R + \gamma_A$. Therefore, in Eq. (9) while the parameter of the aggregation γ_A will be directly given by γ_{long}, the one of the repulsion γ_R will be given by the difference $\gamma_{short} - \gamma_{long}$.

In the empirical application, we try different values both for h_A, h_R and β_A, β_R. In particular, we consider all the combinations of $\beta_A, \beta_R \in (0.1, 0.2, 0.3, 0.4, 0.5)$ and $h_A, h_R \in (20, 25, 30, 35, 40, 45, 50)$ (kilometers) such that $h_R \le h_A$, which amount at estimating 315 models. In Table 1 we only report the estimated result of the best model, that is the model with the lowest AICc.

In particular, in Table 1 we show: in column (1) the OLS estimation of a model that only includes the initial level of income and its square; in column (2) the OLS estimation also including a variable measuring the altimetry of the municipality (OLS altim); in column (3) the OLS estimation of a model which includes the initial level

[10] The matrices W_{K_A} and W_{K_R} are used to approximate an integral in continuous space. Therefore one has to consider also the value of the function in the centre of the neighbour where integration takes place, i.e. the diagonal elements.

of income and its square, as well as our variable proxying for the topography of the municipality based on the altimetry (S-OLS); in column (4) the OLS estimation of a model which includes the initial level of income and its square and the regressors relative to the aggregation and diffusion (AD); in column (5) the OLS estimation of the AD model which also includes the proxy for the topography (SAD); in column (6) the OLS estimation of a model which includes the initial level of income and its square and the regressors relative to the aggregation, repulsion and diffusion (ARD); in column (7) the ARD model is added with the regressor of the topography (SARD). Finally, columns (8)-(10) report the estimate of three commonly used spatial models, that is the SLX, LAG and Durbin respectively.

The models SLX, LAG and Durbin are estimated using a spatial weight matrix W whose elements are defined as:

$$w_{ij} = \begin{cases} e^{-d_{ij}\beta} & \text{if } d_{ij} \leq D \ \forall i \neq j \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

where $\beta \in (0.1, 0.2, 0.3, 0.4, 0.5)$ and $D \in (20, 25, 30, 35, 40, 45, 50)$ which amount to estimating 35 models for each type of spatial model considered. In Table 1 we only report the estimated result over the period 2014-2019 of the best SLX, LAG and Durbin models, that is those with the lowest AICc. Table 2 in Appendix D is analogous to the period 2014-2020.

Considering the results reported in Table 1, we find evidence of a significant positive and concave effect of the initial level of income on its variation. The estimated parameters are very stable across all the specifications, although lower in magnitude and significant only in the linear term in the SLX, LAG and Durbin models. The inclusion of the altimetry variable is negative and significant only in the OLS altim and LAG models. Differently, when the topography is directly considered through our regressor (in S-OLS, SAD, and SARD models), it turns out to be always positive and highly significant: the higher the geographical barriers (as the presence of the mountains), the lower the variation of income towards this place.[11] The inclusion of our regressor of topography also leads to a deep

[11]In the theoretical model of Eq. (1) we assumed the coefficient γ_S to be negative, to have an effect that drives the reallocation towards maxima of the exogenous function $S(z)$. Since here S is given by the

reduction in the AICc, showing its informative content. This highlights the potential of our methodology: the exogenous variable given by altimetry has a big informative content that to be extracted successfully needs to be properly manipulated. The geographic manipulation of the SARD model (or even S-OLS) appears to better extract such informative content than the classical spatial econometrics model SLX, LAG or Durbin.

The coefficients relative to the aggregation, diffusion and repulsion effects ($\gamma_A = \gamma_{long}$, γ_D and $\gamma_R = \gamma_{short} - \gamma_{long}$) are always statistically significant at the usual level of significance and exhibit the expected sign as prescribed by the theoretical model in Eq. (1). This confirms on the geographical scale used (i.e. the municipal level) the discretization implies a relatively small bias. In terms of AICc, the model with the best performance among all possible variants of SARD models is the one where all the components (φ/S/A/R/D) are present. Although the relative estimated coefficients are very low in magnitude[12] their effect on the spatial dynamic is meaningful (see Section 4.3 for details).

The only model that performs better in terms of AICc is the Durbin model, which also shows a positive, very high and significant parameter of the spatial lag. This implies evidence of a strong spatial dependence on the variation of the municipalities' income. However, some issues have to be emphasized: firstly we lose the expected concavity of the initial income terms since the estimated φ_2 is not significant. Secondly, the altimetry of the municipality loses all significance, pointing out that all that effect has been absorbed by some other variables. Thirdly, the estimated ρ corresponding to the spatial lag is very big and therefore is accountable for the lack of significance of the other estimated coefficients. It is worth reminding that although the Durbin model has a lower value of AICc it doesn't allow to capture the different sources of the spatial dependence.

average altimetry level for each municipality, the rationale is that reallocation is not driven towards the maxima of S, but on the contrary, the reallocation towards maxima should be discouraged. Therefore the expected sign is now the opposite, i.e. $\gamma_S > 0$.

[12]The magnitude of the coefficients reflect the scale of the regressor which in our case is not easy to measure.

	Dependent variable: ΔY (2014-2019)									
	OLS	OLS altim	S-OLS	AD	SAD	ARD	SARD	SLX	LAG	Durbin
φ_1	0.0952***	0.0945***	0.0959***	0.0956***	0.0963***	0.09574***	0.09563***	0.07579***	0.073473***	0.076649***
	(0.001)	(0.001)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.001)	(0.001)	(0.001)
φ_2	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−1.38e-05	5.97e-06	−2.11e-05
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(1.43e-05)	(0.00001)	(0.00001)
γ_{altim}		−0.00003**						8.63e-08	−3.7285e-05***	−1.55e-05
		(0.00001)						(0.00002)	(0.00001)	(0.00001)
γ_S			6.37e-08***		5.68e-08***		5.78e-08***			
			(2.97e-09)		(3.73e-09)		(3.71e-09)			
γ_{long}				−2.20e-07***	−8.37e-08***	−0.2138***	−8.48e-07***			
$(=\gamma_A)$				(1.31e-08)	(1.57e-08)	(0.0241)	(9.46e-08)			
γ_{short}						−1.74e-08***	−8.57e-08***			
$(=\gamma_A+\gamma_R)$						(1.4e-08)	(1.56e-08)			
γ_D				2.92e-06***	3.74e-06***	3.39e-06***	3.88e-06***			
				(5.88e-07)	(5.82e-07)	(5.88e-07)	(5.79e-07)			
θ_1								0.01218***		−0.028833***
								(0.001)		(0.002)
θ_2								−4.31e-05*		6.45e-05*
								(0.00001)		(0.00002)
θ_{altim}								−0.0001***		−4.88e-05
								(2.36e-05)		(0.00003)
ρ									0.095707***	0.36713***
									(0.00101)	(0.02)
β_R						0.5	0.5			
h_R						30	20			
β_A				0.5	0.5	0.5	0.1			
h_A				20	20	50	25			
β								0.3	0.3	0.5
D								20	20	50
AICc	8675.9	8672.2	8228.3	8401.1	8174.6	8324.8	8096.5	8225.2	8121.3	7859.6

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1: 2014-2019

4.2.1 Analysis of residuals

Figure 2 reports the spatial distributions of residuals for models SARD and spatial Durbin, for the transition from 2014-2019. We see that the SARD model can better explain the dynamic of the innermost regions, like the Apennine mountains and the centre of Sicily and Sardinia islands. This in particular highlights how our model is better able to exploit the information coming from the altimetry than the classical spatial econometrics models. Moreover, we also appreciate how the spatial pattern of residuals appears more spotted between regions of red and blue colours (corresponding to negative and positive residuals) in the case of SARD models with respect to spatial Durbin, even if the latter performs better in terms of AICc (see Table 1).

pics/residGARD19.pdf

(a) SARD

pics/residDurbin19.pdf

(b) Durbin

Figure 2: The map of residuals for Italian Municipalities for the period 2014-2019 for SARD and spatial Durbin.

When looking at the spatial correlograms for Moran's I and Geary's C indices (Figure 3), we find weak evidence of additional spatial dependence unaccounted for in both the SARD and Durbin models, although higher for the SARD. However, the magnitude of Moran's I index has a maximum of around 0.03, while Geary's C index is never statistically significant.

pics/correlogramGSARD_I19.pdf

(a) Moran's I: SARD

pics/correlogramDurbin_I19.pdf

(b) Moran's I: Durbin

pics/correlogramGSARD_G19.pdf

(c) Geary's C: SARD

pics/correlogramDurbin_G19.pdf

(d) Geary's C: Durbin

Figure 3: Spatial correlograms of residuals for Moran's I and Geary's C for the period 2014-2019.

4.3 Counterfactual analysis

To understand if the effect of each identified component of the SARD is relevant to explain the spatial dynamic of the Italian municipal income, this section presents a counterfactual analysis. Given our short time-span and the small magnitude of the estimated coefficients, we perform the counterfactual analysis in forecast for 50 years. First, we compute the forecasted distribution in 2069, $\hat{Y}^{2069}$, starting from the observed 2019 using the estimated SARD model. To measure the contribution of each reallocation components of the SARD model, we set to zero each estimated coefficient $\hat{\gamma}_S$, $\hat{\gamma}_A$, $\hat{\gamma}_R$ and $\hat{\gamma}_D$ one by one and we compute a counterfactual forecasted distribution in 2069, $\hat{Y}^{2069}_{\hat{\gamma}_S=0}$, $\hat{Y}^{2069}_{\hat{\gamma}_A=0}$, $\hat{Y}^{2069}_{\hat{\gamma}_R=0}$, $\hat{Y}^{2069}_{\hat{\gamma}_D=0}$. Figure 4a shows the estimated cross-section densities of income (log) in 2069 of the full model and the counterfactuals. The densities look similar except around the peak of the distribution, and around 4-4.5 where $\hat{Y}^{2069}_{\hat{\gamma}_A=0}$ and $\hat{Y}^{2069}_{\hat{\gamma}_R=0}$ are the only one that appear different from $\hat{Y}^{2069}$. To measure quantitatively the differences in the forecasted distributions in Figure 4b we report the normalized spatial Gini indices of each models computed with a first-order contiguity spatial matrix (Rey and Smith, 2013). The spatial Gini of $\hat{Y}^{2069}_{\hat{\gamma}_S=0}$ and $\hat{Y}^{2069}_{\hat{\gamma}_D=0}$ are not statistically different at the usual levels of significance from that of $\hat{Y}^{2069}$. Differently, the spatial Gini of $\hat{Y}^{2069}_{\hat{\gamma}_A=0}$ is lower while that of $\hat{Y}^{2069}_{\hat{\gamma}_R=0}$ is higher: in absence of agglomeration the spatial inequality would be lower while in absence of repulsion it would be higher.

Moreover, we also consider the differences between the forecasted final income and the initial one in 2019, denoted by $\Delta SARD = \hat{Y}^{2069} - Y^{2019}$. In Figure 5a we show the distribution of $\Delta SARD$ centred around the median to focus only on the reallocation effects: blue (red) colour implies that the total variation of municipal income is higher (lower) than the median variation. We immediately notice that the metropolitan municipalities show a variation in their income higher than the median, while remote municipalities a variation lower than the median. Then, we consider the differences between the forecasted distribution in 2069 obtained with all coefficients and the forecasted counterfactuals, labelled by $\Delta S = \hat{Y}^{2069} - \hat{Y}^{2069}_{\hat{\gamma}_S=0}$, $\Delta A = \hat{Y}^{2069} - \hat{Y}^{2069}_{\hat{\gamma}_A=0}$, $\Delta R = \hat{Y}^{2069} - \hat{Y}^{2069}_{\hat{\gamma}_R=0}$ and $\Delta D = \hat{Y}^{2069} - \hat{Y}^{2069}_{\hat{\gamma}_D=0}$. Figures 5b-5e show the results of the counterfactual analysis: blue (red) colour indicates that the omitted component would have a positive (negative) effect on the income of the municipality. Figure 5b shows that the presence of mountains causes an increase in income in the surrounding valleys and a corresponding decrease in the mountains themselves. Moreover, the coefficient has a negligible effect in all flat areas as expected. In Figure 5c we see on average a positive effect in metropolitan areas and a corresponding negative effect in the surrounding municipalities confirming the presence of agglomerative effects. It is worth to notice a strong positive effect around the main highway and railways (see e.g., the Po Valley). The impact of the centrifugal components is showed in Figures 5d and 5e: the effect of the two is similar, i.e. negative around metropolitan cities, even though the repulsive effect appears more spatially correlated due to the presence of the averaging spatial kernel W_{K_R}.

pics/density_forecastedDistr-eps-converted-to.pdf

(a) Income density (log) in 2069

pics/spatialGini-eps-converted-to.pdf

(b) Spatial Gini in 2069

Figure 4: Counterfactual distributions

5 Concluding remarks

In this paper, we propose a new family of spatial econometric models, which are inspired by continuous space-time models defined with the formalism of PDEs. The family of models proposed represents an alternative description of the space-time dynamics to other more classical spatial econometrics models, which exhibits some advantages but also some drawbacks, as discussed deeply in the paper. This new class of models allows the disentangling of the mechanism of *accumulation* from that of *reallocation*, i.e. the pics/deltaSARDnorm.pdf

(a) Income variation 2019-2069

pics/deltaS.pdf

(b) Impact of $\hat{\gamma}_S$

pics/deltaA.pdf

(c) Impact of $\hat{\gamma}_A$

pics/deltaR.pdf

(d) Impact of $\hat{\gamma}_R$

pics/deltaD.pdf

(e) Impact of $\hat{\gamma}_D$

Figure 5: Counterfactual analysis 2019-2069

construction of exogenous regressors that are cross-sectional zero-sum. In particular, it makes it possible to further subdivide spatial correlation into different classes of spatial effects, i.e. *aggregation* and *diffusion*. The class of spatial matrices that can be used to specify these terms are more limited with respect to other spatial econometrics models, as discussed in Section 3.2. The methodology employed however allows identifying the effect of aggregation (centripetal) from diffusion (centrifugal) by looking at the sign of the estimated coefficients. Furthermore, this paper also introduces the possibility to construct additional exogenous variables by exploiting the topography of the territory not only in terms of correlation but also in terms of relative differences. Finally, the construction of differential matrices does not suffer from classical computational issues as in other spatial econometric models.

As discussed in Section 3.2 this continuous to discrete approach may suffer from the space granularity of the observations. However, it has been more and more common in modern times, with the advent of high-resolution satellite images or the participation of big companies in academic research, to be able to dispose of highly detailed geo-referenced data, which are not constrained by administrative boundaries. Therefore this paper aims to make an advancement in the field, proposing a model which is more tailored for the advent of the resources available to date.

References

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.

Ames, W. F. (2014). *Numerical methods for partial differential equations*. Academic press.

Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, 27(3):247–267.

Anselin, L. and Rey, S. J. (2012). Spatial econometrics in an age of cybergiscience. *International Journal of Geographical Information Science*, 26(12):2211–2226.

Barthelemy, M. (2016). *The structure and dynamics of cities*. Cambridge University Press.

Benito, J., Urena, F., and Gavete, L. (2001). Influence of several factors in the generalized finite difference method. *Applied Mathematical Modelling*, 25(12):1039–1053.

Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International regional science review*, 26(2):175–188.

Carrillo, J. A., Craig, K., and Yao, Y. (2019a). Aggregation-diffusion equations: dynamics, asymptotics, and singular limits. In *Active Particles, Volume 2*, pages 65–108. Springer.

Carrillo, J. A., Hittmeir, S., Volzone, B., and Yao, Y. (2019b). Nonlinear aggregation-diffusion equations: radial symmetry and long time asymptotics. *Inventiones mathematicae*, 218(3):889–977.

Catellier, R., D'angelo, Y., and Ricci, C. (2021). A mean-field approach to self-interacting networks, convergence and regularity. *Mathematical Models and Methods in Applied Sciences*, 31(13):2597–2641.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics*, volume 5, pages 247–348. Elsevier.

Cressie, N. and Wikle, C. K. (2015). *Statistics for spatio-temporal data*. John Wiley & Sons.

Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). The geography of development. *Journal of Political Economy*, 126(3):903–983.

Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of applied econometrics*, 22(6):1033–1062.

Farlow, S. J. (1993). *Partial differential equations for scientists and engineers*. Courier Corporation.

Fiaschi, D. and Lavezzi, A. M. (2007). Nonlinear economic growth: Some theory and cross-country evidence. *Journal of Development Economics*, 84(1):271–290.

Fiaschi, D. and Ricci, C. (2023). The spatial dynamics of population: an agent-based approach. *Mimeo*.

Fujita, M. and Thisse, J.-F. (1996). Economics of agglomeration. *Journal of the Japanese and international economies*, 10(4):339–378.

Fujita, M. and Thisse, J.-F. (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge University Press.

Jensen, P. S. (1972). Finite difference techniques for variable grids. *Computers & Structures*, 2(1-2):17–29.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3):483–499.

Krugman, P. (1994). Complex landscapes in economic geography. *The American Economic Review*, 84(2):412–416.

Krugman, P. (1998). Space: the final frontier. *Journal of Economic perspectives*, 12(2):161–174.

LeSage, J. and Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Mendes, D. A. and Mendes, V. M. (2015). Parametric models in spatial econometrics: a survey. In *Complexity and Geographical Economics*, pages 51–71. Springer.

Moretti, E. (2004). Human capital externalities in cities. *Handbook of Regional and Urban Economics*, 4:2243–2291.

Quah, D. (2002). Spatial agglomeration dynamics. *American Economic Review*, 92(2):247–252.

Rey, S. J. and Smith, R. J. (2013). A spatial decomposition of the gini coefficient. *Letters in Spatial and Resource Sciences*, 6(2):55–70.

Rossi-Hansberg, E. (2019). Geography of growth and development.

Sznitman, A.-S. (1991). Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer.

Wozniak, A. (2010). Are college graduates more responsive to distant labor market opportunities? *The Journal of Human Resources*, 45(4):944–970.

Xu, X. and Lee, L.-f. (2019). Theoretical foundations for spatial econometric research. *Regional Science and Urban Economics*, 76:2–12.

Appendix

A The micro-foundation of the spatial growth model

This appendix summarizes the main results of Fiaschi and Ricci (2023). Consider a set of N_a rational agents. For the sake of simplicity, we briefly describe here the case where $\varphi \equiv 0$, i.e. there is no change in the total number of agents. The case $\varphi \neq 0$ can be treated similarly, see for example Catellier et al. (2021). Each agent is identified by index i with $i = 1, \ldots, N_a$, and is characterised by its location in the domain $\Omega \subseteq \mathbb{R}^2$, labelled by X_t^{i,N_a}. At $t = 0$ agents are independently distributed at random in the domain Ω following a common probability density distribution on Ω denoted by $y_0(z)$. For $t > 0$ each agent's location evolves by obeying the following Stochastic Differential Equation (SDE):

$$
\begin{aligned}
dX_t^{i,N_a} = & - \gamma_S \nabla_z S\left(X_t^{i,N_a}\right) dt + \\
& - \gamma_A \frac{1}{N_a} \sum_{j=1}^{N_a} \nabla_z K_{h_A}\left(X_t^{i,N_a} - X_t^{j,N_a}\right) dt + \\
& - \gamma_R \frac{1}{N_a} \sum_{j=1}^{N_a} \nabla_z K_{h_R}\left(X_t^{i,N_a} - X_t^{j,N_a}\right) dt + \\
& + \sqrt{2\gamma_D}\, dB_t^i,
\end{aligned} \tag{11}
$$

where $(B_t^i)_{i\in\mathbb{N}}$ is a sequence of independent Brownian motions, and γ_S and $\gamma_A < 0$, while γ_R and $\gamma_D < 0$, to respect the coherence with the phenomena (spatial non uniformity, aggregation, repulsion, diffusion) we are interested to model.

According to Eq. (11) the spatial distribution of agents is evolving by keeping into account its relative position with respect to all other agents. In particular, the first term on the right-hand side of Eq. (11) expresses the tendency of agents to move where the function S is higher, which can decided by the agents' spatial distribution and/or the particular characteristics of different locations. The second and third terms reflect the interactions among agents. In particular, by the linearity of the derivative:

$$\begin{aligned}
\frac{1}{N_a}\sum_{j=1}^{N_a}\nabla_z K_{h_A}\left(X_t^{i,N_a}-X_t^{j,N_a}\right) &= \nabla_z\left(\frac{1}{N_a}\sum_{j=1}^{N_a}K_{h_A}\left(\cdot - X_t^{j,N_a}\right)\right)(X_t^{i,N_a});\ \text{and}\\
\frac{1}{N_a}\sum_{j=1}^{N_a}\nabla_z K_{h_R}\left(X_t^{i,N_a}-X_t^{j,N_a}\right) &= \nabla_z\left(\frac{1}{N_a}\sum_{j=1}^{N_a}K_{h_R}\left(\cdot - X_t^{j,N_a}\right)\right)(X_t^{i,N_a}),
\end{aligned}$$

that is, each agent is moving along the direction of the gradient of a local average (the concept of locality is defined by the functions K_{h_A} and K_{h_R}) of the other agents' location. Finally, the fourth term of Eq. (11) represents the agents' idiosyncratic movements and is expressed through the independent additive noise B_t^i.

The *empirical distribution* of agent's location is:

$$E_t^{N_a} := \frac{1}{N_a}\sum_{i=1}^{N_a}\delta_{X_t^{i,N_a}}, \tag{12}$$

where δ_z is the random variable on $\mathbb{R}^2$ with unitary mass in the point z. $E_t^{N_a}$ is a continuous set of random variables on $\mathbb{R}^2$ depending on time. For any given $N_a \in \mathbb{N}$ and $t > 0$, this distribution is singular, in the sense that is a distribution over $\mathbb{R}^2$ which does not admit a probability density function, since it has a positive probability only on a finite set (corresponding to the location of the N_a agents). However, when N_a goes to infinity the family of random variable $E_t^{N_a}$ becomes diffuse and converges (in distribution) to a continuous family of random variables over $\mathbb{R}^2$, labelled by E_t for any $t > 0$. The distribution of E_t is now regular and admits a probability density function for every t, called $y(t, z)$. An explicit expression for $y(t, z)$ for every t is not available. However one can prove that the probability density function $y(t, z)$ is the solution to Eq. (1) (see Sznitman, 1991, Section 1.1).

B The computation of the partial derivative matrices for the estimate

Below we will report a summary of the *Generalized Finite Difference Method* for calculating partial derivatives when observations are not equally distributed across the space, as originally introduced in (Jensen, 1972). For a self-contained overview of the method see for example (Benito et al., 2001).

Let $y(z)$ be the value of a function at location $z \in \mathbb{R}^2$ and define a set of neighbouring areas z^j whose function values are indicated by $y(z^j)$, for $j = 1, \cdots, n_s$, where n_s is the number of neighbouring locations. Define the following function:

$$\begin{aligned} B(z) &= \sum_{j=1}^{n_s} \left\{ \left[y(z) - y(z^j) + h_j \frac{\partial y(z)}{\partial z_1} + k_j \frac{\partial y(z)}{\partial z_2} + \right.\right. \\ &+ \left.\left. \frac{1}{2} \left(h_j^2 \frac{\partial^2 y(z)}{\partial z_1^2} + k_j^2 \frac{\partial^2 y(z)}{\partial z_2^2} + 2 h_j k_j \frac{\partial^2 y(z)}{\partial z_1 \partial z_2} \right) \right] w(h_j, k_j) \right\}^2 \end{aligned} \quad (13)$$

where $h_j = z_1 - z_1^j$, $k_j = z_2 - z_2^j$ and w is an appropriate weighting function decreasing in both arguments and always not negative. The function $B(z)$ is a weighted linear combination of squares of the error that one has by approximating the function $y(z^j)$ with its second-order linear approximation in the point $y(z)$ for every $j = 1, \ldots, n_s$. Therefore $B(z)$ is close to zero if the second-order Taylor approximation of $y(z^j)$ is approximately correct for all j. Hence, given $y(z)$, $y(z^j)$, h_j, k_j and w_j, for $j = 1, 2, \cdots, n_s$, Eq. (13) can be use to calculate the values of $\frac{\partial y(z)}{\partial z_1}$, $\frac{\partial y(z)}{\partial z_2}$, $\frac{\partial^2 y(z)}{\partial z_1^2}$, $\frac{\partial^2 y(z)}{\partial z_2^2}$ and $\frac{\partial^2 y(z)}{\partial z_1 \partial z_2}$ under the condition that $B(z)$ is minimized. In particular, due to the quadratic shape of the function B, this minimization amount to solve the following system of linear equation (see Benito et al., 2001, p. 6):

$$D_z = \begin{bmatrix} \frac{\partial y(z)}{\partial z_1} \\ \frac{\partial y(z)}{\partial z_2} \\ \frac{\partial^2 y(z)}{\partial z_1^2} \\ \frac{\partial^2 y(z)}{\partial z_2^2} \\ \frac{\partial^2 y(z)}{\partial z_1 \partial z_2} \end{bmatrix} = \mathbf{D} \begin{bmatrix} y(z) \\ y(z^1) \\ y(z^2) \\ \cdots \\ y(z^{n_s}) \end{bmatrix}, \quad (14)$$

with $\mathbf{D} \equiv \mathbf{A}^{-1}\mathbf{B}$ is a $(5 \times (n_s + 1))$, $w_j = w(h_j, k_j)$ and where $\mathbf{A}$ is a symmetric (5×5)

matrix defined as:

$$A = \begin{bmatrix} \sum_{j=1}^{n_s} h_j^2 w_j^2 & \sum_{j=1}^{n_s} h_j k_j w_j^2 & \frac{1}{2}\sum_{j=1}^{n_s} h_j^3 w_j^2 & \frac{1}{2}\sum_{j=1}^{n_s} h_j k_j^2 w_j^2 & \sum_{j=1}^{n_s} h_j^2 k_j w_j^2 \\ & \sum_{j=1}^{n_s} k_j^2 w_j^2 & \frac{1}{2}\sum_{j=1}^{n_s} h_j^2 k_j w_j^2 & \frac{1}{2}\sum_{j=1}^{n_s} k_j^3 w_j^2 & \sum_{j=1}^{n_s} h_j k_j^2 w_j^2 \\ & & \frac{1}{4}\sum_{j=1}^{n_s} h_j^4 w_j^2 & \frac{1}{4}\sum_{j=1}^{n_s} h_j^2 k_j^2 w_j^2 & \frac{1}{2}\sum_{j=1}^{n_s} h_j^3 k_j w_j^2 \\ & & & \frac{1}{4}\sum_{j=1}^{n_s} k_j^4 w_j^2 & \frac{1}{2}\sum_{j=1}^{n_s} h_j k_j^3 w_j^2 \\ & & & & \sum_{j=1}^{n_s} h_j^2 k_j^2 w_j^2 \end{bmatrix} \quad (15)$$

and **B** is a $(5 \times (n_s + 1))$ matrix defined as:

$$B = \begin{bmatrix} -\sum_{j=1}^{n_s} h_j w_j^2 & h_1 w_1^2 & h_2 w_2^2 & \cdots & h_{n_s} w_{n_s}^2 \\ -\sum_{j=1}^{n_s} k_j w_j^2 & k_1 w_1^2 & k_2 w_2^2 & \cdots & k_{n_s} w_{n_s}^2 \\ -\frac{1}{2}\sum_{j=1}^{n_s} h_j^2 w_j^2 & \frac{1}{2} h_1^2 2 w_1^2 & \frac{1}{2} k_1^2 w_2^2 & \cdots & \frac{1}{2} h_{n_s}^2 w_{n_s}^2 \\ -\frac{1}{2}\sum_{j=1}^{n_s} k_j^2 w_j^2 & \frac{1}{2} h_2^2 w_1^2 & \frac{1}{2} k_2^2 w_2^2 & \cdots & \frac{1}{2} k_{n_s}^2 w_{n_s}^2 \\ -\sum_{j=1}^{n_s} h_j k_j w_j^2 & h_1 k_1 w_1^2 & h_2 k_2 w_2^2 & \cdots & h_{n_s} k_{n_s} w_{n_s}^2 \end{bmatrix} \quad (16)$$

Benito et al. (2001) suggests using the following as a weighting function:

$$w(d_j) = 1 - 6\left(\frac{d_j}{dm_j}\right)^2 + 8\left(\frac{d_j}{dm_j}\right)^3 - 3\left(\frac{d_j}{dm_j}\right)^4, \quad (17)$$

where $d_j^2 = h_j^2 + k_j^2$ is the squared distance between z and z_j and dm_j is the maximum distance on all possible neighbouring locations.

The set of all neighbouring locations used to approximate the partial derivatives in location z is called the *star* of the location; the choice of the elements of the *star* is an important factor to ensure the accuracy of the method. In our framework we will adopt the *closest neighbourhood criterion* that is, fixed the number of elements of the *star* n_s, we select the n_s closest locations to each given location.

It is now clear that for every given location z, it is possible to construct the matrix D and all the partial derivatives in location z up to the second order. We now explain how to carry out these operations in a compact manner in the case of a finite number of locations. Assume we have $(z^i)_{i=1,\dots,N}$ a finite number of locations. For each of these locations denote by $\mathbf{J^i} = (\mathbf{J^i}_k)_{k=1,\dots,n_s}$ the vector of indices of locations in the *star* of location z^i (by definition of star i is not an element of $\mathbf{J^i}$), and by $\mathbf{D_i}$ the matrix $\mathbf{D}$ associated to the location z^i. It is now clear that we have

$$\begin{bmatrix} \dfrac{\partial y(z^i)}{\partial z_1} \\ \dfrac{\partial y(z^i)}{\partial z_2} \\ \dfrac{\partial^2 y(z^i)}{\partial z_1^2} \\ \dfrac{\partial^2 y(z^i)}{\partial z_2^2} \\ \dfrac{\partial^2 y(z^i)}{\partial z_1 \partial z_2} \end{bmatrix} = \mathbf{D_i} \underbrace{\begin{bmatrix} y(z^i) \\ y(z^{\mathbf{J^i}_1}) \\ y(z^{\mathbf{J^i}_2}) \\ \vdots \\ y(z^{\mathbf{J^i}_{n_s}}) \end{bmatrix}}_{\overline{\mathbf{y}}_i} = \mathbf{D_i}\overline{\mathbf{y}}_i. \tag{18}$$

Therefore, by denoting with $(D_i)_{k,}$ the k-th row of the matrix D_i, one has

$$\begin{aligned} \frac{\partial y(z^i)}{\partial z_1} &= \langle (\mathbf{D_i})_{1,}, \overline{\mathbf{y}}_i \rangle, \\ \frac{\partial y(z^i)}{\partial z_2} &= \langle (\mathbf{D_i})_{2,}, \overline{\mathbf{y}}_i \rangle, \\ \frac{\partial^2 y(z^i)}{\partial z_1^2} &= \langle (\mathbf{D_i})_{3,}, \overline{\mathbf{y}}_i \rangle, \\ \frac{\partial^2 y(z^i)}{\partial z_2^2} &= \langle (\mathbf{D_i})_{4,}, \overline{\mathbf{y}}_i \rangle, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product of $\mathbb{R}^{n_s+1}$.

We will now show how to compute all the partial derivatives in all the locations via single matrix multiplications. We start by setting

$$\mathbf{y} = (y(z^1), y(z^2), \ldots, y(z^N))^t$$

and the vectors of partial derivatives in all the locations as

$$\frac{\partial y}{\partial z_1} = \begin{bmatrix} \dfrac{\partial y(z^1)}{\partial z_1} \\ \dfrac{\partial y(z^2)}{\partial z_1} \\ \vdots \\ \dfrac{\partial y(z^N)}{\partial z_1} \end{bmatrix}, \frac{\partial y}{\partial z_2} = \begin{bmatrix} \dfrac{\partial y(z^1)}{\partial z_2} \\ \dfrac{\partial y(z^2)}{\partial z_2} \\ \vdots \\ \dfrac{\partial y(z^N)}{\partial z_2} \end{bmatrix}, \frac{\partial^2 y}{\partial z_1^2} = \begin{bmatrix} \dfrac{\partial^2 y(z^1)}{\partial z_1^2} \\ \dfrac{\partial^2 y(z^2)}{\partial z_1^2} \\ \vdots \\ \dfrac{\partial^2 y(z^N)}{\partial z_1^2} \end{bmatrix}, \frac{\partial^2 y}{\partial z_2^2} = \begin{bmatrix} \dfrac{\partial^2 y(z^1)}{\partial z_2^2} \\ \dfrac{\partial^2 y(z^2)}{\partial z_2^2} \\ \vdots \\ \dfrac{\partial^2 y(z^N)}{\partial z_2^2} \end{bmatrix}.$$

If we introduce the matrices of size $N \times N$, M_{z_1}, M_{z_2}, $M_{z_1 z_1}$ and $M_{z_2 z_2}$ defined by

$$(M_{z_1})_{i,j} = \begin{cases} (\mathbf{D_i})_{1,1} & \text{if } j = I, \\ (\mathbf{D_i})_{1,k+1} & \text{if } j = \mathbf{J^i}_k \text{ for some } k \in \{1, \dots, n_s\}, \\ 0 & \text{otherwise,} \end{cases}$$

$$(M_{z_2})_{i,j} = \begin{cases} (\mathbf{D_i})_{2,2} & \text{if } j = I, \\ (\mathbf{D_i})_{2,k+1} & \text{if } j = \mathbf{J^i}_k \text{ for some } k \in \{1, \dots, n_s\}, \\ 0 & \text{otherwise,} \end{cases}$$

$$(M_{z_1 z_1})_{i,j} = \begin{cases} (\mathbf{D_i})_{3,3} & \text{if } j = I, \\ (\mathbf{D_i})_{3,k+1} & \text{if } j = \mathbf{J^i}_k \text{ for some } k \in \{1, \dots, n_s\}, \\ 0 & \text{otherwise,} \end{cases}$$

$$(M_{z_2 z_2})_{i,j} = \begin{cases} (\mathbf{D_i})_{4,4} & \text{if } j = I, \\ (\mathbf{D_i})_{4,k+1} & \text{if } j = \mathbf{J^i}_k \text{ for some } k \in \{1, \dots, n_s\}, \\ 0 & \text{otherwise,} \end{cases}$$

then one has

$$\frac{\partial y(z^i)}{\partial z_1} \approx (M_{z_1}\mathbf{y})_i \,, \quad \frac{\partial y(z^i)}{\partial z_2} \approx (M_{z_2}\mathbf{y})_i \,, \quad \frac{\partial^2 y(z^i)}{\partial z_1^2} \approx (M_{z_1 z_1}\mathbf{y})_i \,, \quad \frac{\partial^2 y(z^i)}{\partial z_2^2} \approx (M_{z_2 z_2}\mathbf{y})_i \,.$$

C A numerical investigation of the properties of the spatial growth model

An explicit solution for Eq. (1) is not available in general. We will then proceed to discuss some of the basic properties of Eq. (1) by numerical simulations. In all examples, we make some simplifications to better discern the peculiar features of the model from the exogenous effects related to the particular case considered in Section 4. In particular, we take the total cross-sectional amount of the variable y of the system constant and equal to 1, i.e. we take $\varphi \equiv 0$. Moreover, we take the exogenous function $S(z)$ to be identically zero, so that no exogenous variable is present. We also neglect the repulsive effect, i.e. set $\gamma_R = 0$. Summarizing we consider only the *aggregation* (γ_A) and *diffusion* (γ_D) effects. We study the problem on a squared domain $\Omega = [0, 4] \times [0, 4]$ with a discrete set of locations uniformly spaced with a distance between contiguous locations of $\Delta x = 10^{-2}$ (i.e. $160'000$ total points).

In figures 6 and 7 we set an initial condition that is flat at the centre of the domain and zero close to the boundary, with some intermediate values in between so that the initial profile is a continuous function. We then set $\gamma_A = -0.01, \gamma_D = 0.005$ and explore the impacts of changes in h_A since the distance at which aggregation takes place is one of the main characteristics of the system which determines the evolution of the spatial pattern. Figure 6 reports the distribution dynamics of the baseline model which highlights the dynamics of aggregation of income, i.e. the emergence of a city, in a central location at around $t = 20$. We notice that there is an initial temporary formation of four smaller clusters around $t = 5$, which then agglomerates together to contribute to the formation of the single large cluster in the centre of the domain. In Figure 7 we kept the initial condition unchanged and only reduced the distance of aggregation from 0.4 to 0.3. The spatial pattern in the first few periods is roughly the same, exhibiting the formation of the four temporary clusters around period $t = 5$. However, since the distance at which aggregation takes place is smaller, the four clusters are now so far apart that they are not able to merge anymore. Therefore we observe at the final period $t = 20$ a configuration which is made of four separate clusters instead of a single larger one. The spatial pattern strongly remembers the ones discussed in Krugman (1994) and in Barthelemy (2016, cap. 8).

D The estimation over the period 2014-2020

When looking at the estimates for the period 2014-2020 in Table 2, all the results found for the period 2014-2019 are confirmed. However, over this period the SARD model is also the best in terms of AICc.

pics/t=0h=0.4-eps-converted-to.pdf | pics/t=0.5h=0.4-eps-converted-to.pdf | pics/t=1h=0.4-eps-converted-to.pdf

(a) $t = 0$ (b) $t = 0.5$ (c) $t = 1$

pics/t=5h=0.4-eps-converted-to.pdf | pics/t=10h=0.4-eps-converted-to.pdf | pics/t=20h=0.4-eps-converted-to.pdf

(d) $t = 5$ (e) $t = 10$ (f) $t = 20$

Figure 6: The distribution dynamics of $y(t, z)$ over space and time for the baseline case with $\gamma_S = 0, \gamma_A = -0.01, \gamma_R = 0$, $\gamma_D = 0.005$, $h_A = 0.4$, $\Omega = [0, 4] \times [0, 4]$.

pics/t=0h=0.3-eps-converted-to.pdf | pics/t=0.5h=0.3-eps-converted-to.pdf | pics/t=1h=0.3-eps-converted-to.pdf

(a) $t = 0$ (b) $t = 0.5$ (c) $t = 1$

pics/t=5h=0.3-eps-converted-to.pdf | pics/t=10h=0.3-eps-converted-to.pdf | pics/t=20h=0.3-eps-converted-to.pdf

(d) $t = 5$ (e) $t = 10$ (f) $t = 20$

Figure 7: The distribution dynamics of $y(t, z)$ over space and time for the baseline case with $\gamma_S = 0, \gamma_A = -0.01, \gamma_R = 0$, $\gamma_D = 0.005$, $h_A = 0.3$, $\Omega = [0, 4] \times [0, 4]$.

	Dependent variable: ΔY (2014-2020)									
	OLS	OLS altim	S-OLS	AD	SAD	ARD	SARD	SLX	LAG	Durbin
φ_1	0.0715***	0.0704***	0.0738***	0.07238***	0.07428***	0.0729***	0.07441***	0.04858***	0.041681***	0.047491***
	(0.001)	(0.001)	(0.0009)	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.001)	(0.01)	(0.001)
φ_2	−0.0002***	−0.0002***	−0.0002***	−0.0002***	−0.0002***	−0.0001***	−0.0001***	−6.94e-05***	−2.68e-06	−4.81e-05
	(0.00001)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(1.80e-05)	(0.00001)	(0.00002)
γ_{altim}		−0.00004***						6.84e-06	−5.22e-05***	−2.75e-05
		(0.00001)						(0.00002)	(0.00001)	(0.00002)
γ_S			1.54e-07***		1.48e-07***		1.38e-07***			
			(3.43e-09)		(4.30e-09)		(4.51e-09)			
γ_{long} $(= \gamma_A)$				−4.32e-07***	−7.63e-08***	−7.00e-05***	−2.56e-06***			
				(1.60e-08)	(1.81e-08)	(3.66e-06)	(3.35e-07)			
γ_{short} $(= \gamma_A + \gamma_R)$						−3.21e-07***	−6.95e-08***			
						(1.67e-08)	(1.81e-08)			
γ_D				2.04e-06***	4.18e-06***	3.01e-06***	4.52e-06***			
				(7.18e-07)	(1.81e-08)	(7.03e-07)	(6.70e-07)			
θ_1								0.01209***		−0.014777***
								(0.001)		(0.002)
θ_2								−6.02e-06		5.14e-05**
								(0.00002)		(0.00002)
θ_{altim}								−0.0001***		−5.08e-05*
								(0.00003)		(0.00003)
ρ									0.016081***	0.32599***
									(0.00554)	(0.01)
β_R						0.5	0.5			
h_R						30	30			
β_A				0.5	0.5	0.2	0.1			
h_A				20	20	40	35			
β								0.3	0.4	0.3
D								20	20	35
AICc	12242.4	12236.8	10449.2	11512.1	10405.7	11155.6	10349.7	11805.3	11448.5	11225.4

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: 2014-2020

pics/residGARD20.pdf

(a) SARD

pics/residDurbin20.pdf

(b) Durbin

Figure 8: The map of residuals for Italian Municipalities for the period 2014-2020 for SARD and spatial Durbin.

pics/correlogramGSARD_I20.pdf

(a) Moran's I: SARD

pics/correlogramDurbin_I20.pdf

(b) Moran's I: Durbin

pics/correlogramGSARD_G20.pdf

(c) Geary's C: SARD

pics/correlogramDurbin_G20.pdf

(d) Geary's C: Durbin

Figure 9: Spatial correlograms of residuals for Moran's I and Geary's C for the period 2014-2020.