WORKING PAPERS

N° 1357

February 2023

"Generalizing impact computations for the autoregressive spatial interaction model"

Thibault Laurent, Paula Margaretic and Christine Thomas-Agnan

Toulouse School of Economics

Generalizing impact computations for the autoregressive spatial interaction model

Thibault Laurent,* Paula Margaretic,† and Christine Thomas-Agnan‡

February 10, 2023

Abstract

We extend the impact decomposition proposed by LeSage and Thomas-Agnan (2015) in the spatial interaction model to a more general framework, where the sets of origins and destinations can be different, and where the relevant attributes characterizing the origins do not coincide with those of the destinations. These extensions result in three flow data configurations which we study extensively: the square, the rectangular, and the non-cartesian cases. We propose numerical simplifications to compute the impacts, avoiding the inversion of a large filter matrix. These simplifications considerably reduce computation time; they can also be useful for prediction. Furthermore, we define local measures for the intra, origin, destination and network effects. Interestingly, these local measures can be aggregated at different levels of analysis. Finally, we illustrate our methodology in a case study using remittance flows all over the world.

***Keywords*—** Impact decomposition, local effects, spatial interaction autoregressive models, non-cartesian flow data

1 Introduction

Spatial interaction models (SIMs) are a class of models that are used to characterize origin-destination (OD) flow data. OD flow data represent movements of tangible entities (such as persons, goods) and intangible ones (such as capital, investment, knowledge) from origins to destinations (which can be cities, regions, countries or positions). Spatial interaction models explain the interaction between origin and destination locations using: (i) origin-specific attributes characterizing the ability of the origins to generate flows; (ii) destination-specific characteristics representing the attractiveness of destinations; (iii) variables that characterize the way spatial separation of origins from destinations constrain the interaction.[1] Previous literature has shown that using

*Toulouse School of Economics, CNRS
†University Adolfo Ibañez, Chile
‡Toulouse School of Economics, Université Toulouse Capitole

[1]In the initial SIMs, the spatial flows between an origin and a destination were conceived to be proportional to their gravitational forces and inversely related to their spatial separation. Characteristics of origins and destinations, hereafter U_i and V_j, respectively, hence represent the gravitational forces pushing and/or pulling the tangible or intangible entities

spatial separation variables, such as distance between origin and destination locations, is generally not enough to eradicate the spatial dependence and/or the spatial autocorrelation among the sample of OD flows (Curry et al., 1975; Curry, 1972).

Over the last 50 years, SIMs have been applied in a wide variety of contexts, such as retail, transport, housing, public policy, land use, urban and population modelling and planning. SIMs have been used for two main purposes: prediction (of the size and direction of the spatial flows) and inference (about the factors shaping the spatial interactions in a network of flows; Rowe and Dennett, 2022). Furthermore, the literature has proposed five methodological approaches to account for the spatial structure: the competing destination model (which includes an accessibility measure by using a Box–Cox functional form of distance, see, e.g., Fotheringham, 1983); the Box–Cox transformation (by using a Box–Cox functional form of distance, see, e.g., Tiefelsdorf, 2003); spatial choice modelling by accounting for destination alternative substitution patterns (Hunt et al., 2004); spatial econometric modelling (LeSage and Pace, 2008) and eigenvector spatial filtering (Chun and Griffith, 2011).

Starting with Griffith and Jones (1980), the SIM literature distinguishes among: (i) spatial autocorrelation in the residuals, which arises from missing origin and destination spatially autocorrelated explanatory variables and/or a model term that captures location accessibility (Tiefelsdorf and Griffith, 2007; Fischer and Griffith, 2008; Fotheringham, 1981);[2] (ii) spatial dependence arising from the direct interaction between the dependent flows themselves. Moreover, in relation to (ii), the literature identifies two different types of spatial dependence in flow magnitudes: endogenous and exogenous interaction effects.

Endogenous interaction effects reflect situations where flows from regions (or sites) neighboring the origin i or destination j, as well as flows between regions neighboring the origin and neighboring the destination may exert an impact (a feedback effect) on the magnitude of flows between regions i and j (LeSage and Fischer, 2016). We can model this type of interaction relying on spatial autoregressive specifications as in LeSage and Pace (2009), who include spatial lags of the dependent variable to quantify the magnitude and extent of these feedback effects, hence the term endogenous interaction. These specifications imply a simultaneous or endogenous response relationship between the variation in the dependent variable of a given dyad and flows between other regions in the observed network of flows. One example would be international trade flows, where tariff or non-tariff barriers limit trade, which in turn leads to a long-run response that changes the structure of trade flows across the network of trading countries. We call these global spillovers, as they imply diffusion over space (thus impacting neighbors, neighbors of neighbors, and so on).

In contrast, exogenous interaction represents a situation where the spillovers are local and arise from changes in the characteristics of nearby sites. Exogenous effects do not generate reaction to feedback impacts or global diffusion. We account for exogenous effects including spatial lags of the explanatory variables (which represent characteristics of connected sites). One example would be congestion in neighboring regions in an application where we are interested in modelling commuting flows.

The distinction between endogenous and exogenous interaction effects is important for model interpretation.

from and to specific locations, whereas different measures of distance and costs, which we denote as $f(d_{i:j})$, represent the deterring effects of geographical separation of spatial flows (Rowe and Dennett, 2022). A SIM specification can be written as,

$$E(Y_{i:j}) = K_{i:j} U_i V_j f(d_{i:j}) \tag{1}$$

where $E(Y_{i:j})$ is the expectation of a flow from origin i to destination i and $K_{i:j}$ is an origin–destination-specific constant of proportionality or scaling factor.

[2]Fotheringham (1983) shows that the inclusion of a variable that measures each destination's accessibility to all other destinations can greatly reduce the spatial variation in local distance-decay parameter estimates and remove unintuitive patterns.

In spatial autoregressive interaction models, a change in a characteristic at a given location gives rise to possibly numerous impacts on the OD flow matrix (in contrast with the classical non-spatial linear model). Indeed, due to the presence of the filter matrix, the parameters of the explanatory variables no longer coincide with the changes of the dependent variable resulting from increments of the independent variables.[3] This has two consequences: 1) changes are no longer the same for all locations, and 2) an increment of a given explanatory variable at a location may result in non-null variations at all other locations. LeSage and Thomas-Agnan (2015) have analyzed the impact measures in the spatial autoregressive interaction model with endogenous spatial interaction; in turn, LeSage and Fischer (2016) have studied the impact decomposition allowing for exogenous spatial interactions. These two papers decompose the total effects into intraregional, origin, destination, and network effects.

However, both papers impose two restrictions: that the list of origins coincides with the list of destinations and that the characteristics of origins are the same as the characteristics of destinations.

Dargel and Thomas-Agnan (2023) introduce a more general version of the spatial interaction model allowing for different lists of origins and destinations and for possibly different origin and destination characteristics.[4] The goal of the present paper is to examine the impact decomposition in this new framework. Being able to lift these two restrictions when computing impact measures is particularly appealing in numerous applications. For instance, in an application with bilateral remittance flows between country pairs, a country that receives money from abroad may not necessarily send money abroad (for example, if it is a low-income country). As another example, consider an application with OD transfer passengers between home and work. In this case, the number of inhabitants at the origins may be a relevant characteristic for the residential locations, whereas one would like to include the number of employees to characterize the destinations. A third possible illustration is a geomarketing application where the OD flows occur between consumers and retail stores. Also in this case, it is reasonable to expect that the factors characterizing the retail stores (such as the surface of the stores) will not coincide with the destination attributes (e.g. population of the areas where consumers live). The above illustrations motivate the present research.

The contribution of our paper is three-fold. First, we extend the impact decomposition of LeSage and Thomas-Agnan (2015) and LeSage and Fischer (2016) in the framework of a spatial autoregressive Durbin model (with endogenous and exogenous interactions), where the list of origins does not necessarily coincide with the list of destinations and where the sets of origin and destination characteristics are different. This impact generalization results in three distinct cases which we study extensively in the paper: the standard square case, that is, when the lists of origins and of destinations are the same; the rectangular case, i.e. when the list of origins does not coincide (or coincides partially) with the list of destinations, but all possible flows are observed; and the non-cartesian case, that is, when some of the possible flows between the origins and the destinations are not observed or have no meaning. To the best of our knowledge, we are the first to decompose the impacts in this general setting. Our second contribution is that, by exploiting some properties of the filter matrix, we propose a procedure to compute the impacts that considerably reduces the computation time in the cartesian case, under few hypotheses. This is achieved by avoiding to compute the inverse filter matrix of size $N \times N$, where N is the total number of flows. Instead, we need the inversion of a $\sqrt{N} \times \sqrt{N}$ matrix. The computation time reduction is particularly interesting in applications where the total number of flows is large. For instance, with a data set of $N = 3600$

[3] The model specification with exogenous effects only can be interpreted without considering changes in the long-run equilibrium state of the system of flows (LeSage and Fischer, 2016).

[4] In a typical spatial interaction model, the sample involves n locations being origins and destinations at the same time and k variables characterizing the origins as well as the destinations.

flows, computational time (using a Intel Xeon(R) CPU E5, 2.20GHz) is around $1s$ using our algorithm, $5s$ using a classical algorithm taking into account the sparsity and $12s$ using a classical algorithm with no adjustment. Finally, we propose alternative representations of the local effects which can be aggregated at different levels of analysis depending on the research interest. As an application of our work, we present a case study consisting of (a rectangular dataset of) bilateral remittance flows among macro-economic zones all over the world, where the list of origins is different from the list of destinations.

To compute the impacts, in this paper, we first determine the local impacts associated to a location s when changing a characteristic at that particular location s. We define the cumulative local impact of changing $\mathbf{Z}$ at a single location s as the sum of the local intraregional effect, the local origin effect, the local destination effect, and the local network effect, when these terms have a meaning (they are zero otherwise). Precisely, the local intraregional effect measures the local impact of changing a given characteristic $\mathbf{Z}$ at location s on the flow originating and arriving at that location s. This of course has a meaning provided location s is both an origin and a destination site. In turn, the local origin (resp: destination) effect corresponds to the sum of local effects due to a change in $\mathbf{Z}$ at a single location s on the flows starting from (resp: arriving at) s. This effect only has a meaning for origin (resp: destination) characteristics. Finally, the local network effect is the sum of effects of changing a given characteristic $\mathbf{Z}$ at a location s on the flows that do not start from neither arrive at s. The local impact decomposition that we propose in this paper should be of interest to researchers seeking to isolate the local or regional impact of shocks, interventions or policies on a specific subset of positions (including areas, regions, countries). For example, one may be interested to know what has been the impact of a negative economic shock on the migration, knowledge or remittance flows involving certain groups of countries, such as emerging economies or low-to-middle-income countries. These would then be summary measures of impacts with an intermediate level of aggregation.

Going from the local to the global effects, we then define the total impact of changing $\mathbf{Z}$ at all locations on all flows as the sum over all locations of the local total impacts (or as the sum over all locations of the local intraregional, origin, destination and network effects, when these terms have a meaning). Note that our procedure of obtaining the total impacts as sums of local impacts contrasts with the existing literature (LeSage and Pace, 2006, as an example), which defines the total impact directly. As in LeSage and Thomas-Agnan (2015), we also present scalar summary measures for the total intraregional, origin, destination, and network effects by normalizing these total effects by the total number of flows. These scalar summary measures represent an average effect over all possible flows, or the effect on a typical flow. The present definitions reduce to the standard ones in the square framework with characteristics common to origin and destination.

The structure of this paper is as follows. Section 2 presents the spatial autoregressive interaction model with possibly different origin and destination sets. Section 3, in turn, examines the model interpretation in the square, the rectangular, and the non-cartesian cases. The analysis starts by presenting the decomposition of local impacts; next, we propose two graphical representations of the local impact measures; lastly, we present the total effects by aggregating the local impacts over all possible locations. Section 4 shows how to simplify the impact computations in the cartesian case, whereas Section 5 proposes simplified prediction formulas, also in the cartesian case. Section 6 illustrates our results with a case study of bilateral remittance flows among regions all over the world. Section six concludes with a discussion of the possible applications of our results.

2 A spatial autoregressive interaction model with possibly different origin and destination sets

Interaction or gravity models attempt to explain the interaction between origin and destination locations using: (i) origin-specific attributes characterizing the ability of the origins to generate flows; (ii) destination-specific characteristics representing the attractiveness of destinations; (iii) variables that characterize the way spatial separation of origins from destinations constrains the interaction. Acknowledging that using spatial separation variables, such as distance between origin and destination locations, is generally not enough to eradicate the spatial effects in the sample of origin-destination (OD) flows, spatial autoregressive interaction models augment the gravity equation with spatially lagged dependent and independent variables.

In a typical spatial autoregressive interaction model, the sample involves n locations being origins and destinations at the same time and k variables characterizing the origins as well as the destinations. In this paper, we will consider a more general case. We first relax the assumption that the same variables characterize both the origins and the destinations (let k_o and k_d be the numbers of origin and destination characteristics, respectively). We then allow for n_o origins and n_d destinations resulting in $N = n_o \times n_d$ pairs (o, d) of origin-destination (OD) flows respectively. For example, this would be the case in a typical geomarketing application where the origins would be the locations of the customers, the destinations the locations of the stores; the characteristics of origins would be some socio-economic variables describing the places of residence of customers and the characteristics of destination would describe the stores (area, number of employees, ets.). An even more general and realistic case, which we study in this paper, would be to allow some pairs (o, d) to be missing.

Let $\mathbf{Y}$ be the flow matrix, where the n_d columns represent the destination locations indexed by 1 to n_d and the n_o rows correspond to the origin locations indexed by 1 to n_o, organized as follows

$$\begin{pmatrix} o_1 \to d_1 & o_1 \to d_2 & ... & o_1 \to d_{n_d} \\ o_2 \to d_1 & o_2 \to d_2 & ... & ... \\ & & & o_{n_o-1} \to d_{n_d} \\ & & & o_{n_o} \to d_{n_d} \end{pmatrix} \tag{2}$$

We denote by O the set of origins locations, D the set of destinations locations, and $\mathbf{Y_{o:d}}$ an OD flow from origin o to destination d. Note that the OD flow matrix we consider in this paper, equation (2) above, is an oriented flow matrix, in the sense that $\mathbf{Y_{o:d}} \neq \mathbf{Y_{d:o}}$. Let us denote by S the union of O and D and by n_S its cardinality. Two possible vectorizations of the flow matrix $\mathbf{Y}$ are possible, depending on whether we stack its columns (destination centric) or its rows (origin centric). Without loss of generality, we adopt in the following a destination centric ordering and denote by $\mathbf{y}$ the flow vector, of length $N \times 1$. Hence, the first n_d elements of $\mathbf{y}$ represent flows from origin 1 to all n_d destinations. All formulas below can be adapted to the origin-centric scheme.

Let $\mathbf{OW}$ be a $n_o \times n_o$ proximity matrix characterizing the proximity in the set of origin locations, with the proximity being defined, for instance, based on m-nearest neighbours or contiguity, among others. $\mathbf{OW}$ represents a non-negative, row-normalized, sparse matrix, with element $ow_{i,j} > 0$ if site j is one of the neighbours to site i and $\sum_j ow_{i,j} = 1$. Similarly, let $\mathbf{DW}$ of dimension $n_d \times n_d$, be a proximity matrix characterizing the proximity in the set of destination locations. We also assume that $\mathbf{DW}$ is row-normalized i.e. $\sum_j dw_{i,j} = 1$.[5] We consider

[5]Note that we allow for different proximity matrices for the sets of origin and destination locations.

the following three types of neighborhood structures

- $\mathbf{W_o} = \mathbf{OW} \bigotimes \mathbf{I_{n_d}}$ is the origin based spatial neighborhood matrix,
- $\mathbf{W_d} = \mathbf{I_{n_o}} \bigotimes \mathbf{DW}$ is the destination based spatial neighborhood matrix,
- $\mathbf{W_w} = \mathbf{OW} \bigotimes \mathbf{DW}$ is the origin-to-destination based spatial neighborhood matrix,

with $\bigotimes$ standing for the Kronecker product of two matrices. Note that the three weight matrices $\mathbf{W}_o$, $\mathbf{W}_d$ and $\mathbf{W}_w$ are of dimension $N \times N$.

Regarding the origin and destination characteristics, let $\mathbf{OX}$ be the matrix of the k_o characteristics of the origin locations, which is of dimension $n_o \times k_o$, and $\mathbf{DX}$ that of the k_d destination characteristics, with dimension $n_d \times k_d$. On top of that, the characteristics of the origins (respectively, destinations) may appear in their lagged form in the model (depending on the application, we could eventually spatially lag a subset of the origin and destination characteristics). We hence construct the following four matrices

- $\mathbf{X_o} = \mathbf{OX} \bigotimes \iota_{\mathbf{n_d}}$, of dimension $(N \times k_o)$, the characteristics of the origin locations,
- $\mathbf{X_d} = \iota_{\mathbf{n_o}} \bigotimes \mathbf{DX}$, of dimension $(N \times k_d)$, the characteristics of the destination locations.
- $\mathbf{LX_o} = \mathbf{LOX} \bigotimes \iota_{\mathbf{n_d}}$, with $\mathbf{LOX} = \mathbf{OW} \times \mathbf{OX}$ being the lagged characteristics of the spatial units acting as origins.
- $\mathbf{LX_d} = \iota_{\mathbf{n_o}} \bigotimes \mathbf{LDX}$, with $\mathbf{LDX} = \mathbf{DW} \times \mathbf{DX}$ being the lagged characteristics of the spatial units acting as destinations.

Let $\mathbf{G}$ be a matrix of variables characterizing the pairs of origins and destinations (for example, distance). The spatial autoregressive Durbin model, hereafter SDM, in its reduced form can be written as follows

$$(\mathbf{I_{N\times N}} - \rho_o \mathbf{W_o} - \rho_d \mathbf{W_d} - \rho_w \mathbf{W_w})y = \iota_{\mathbf{N}}\alpha + \mathbf{X_o}\beta_{\mathbf{o}} + \mathbf{X_d}\beta_{\mathbf{d}} + \mathbf{LX_o}\delta_{\mathbf{o}} + \mathbf{LX_d}\delta_{\mathbf{d}} + \mathbf{G}\omega + \epsilon, \tag{3}$$

where the parameters ρ_o, ρ_d and ρ_w capture the strength of the origin, destination, and origin-to-destination spatial dependence, respectively; $\beta_{\mathbf{o}}$, $\beta_{\mathbf{d}}$, $\delta_{\mathbf{o}}$, $\delta_{\mathbf{d}}$ and ω are the vectors of parameters whose dimensions correspond to the number of variables in $\mathbf{X_o}$, $\mathbf{X_d}$, $\mathbf{LX_o}$, $\mathbf{LX_d}$ and $\mathbf{G}$, respectively.

One difficulty that often arises when modelling OD flows is the presence of large flows within regions (that is, those located on the diagonal of the OD flow matrix $\mathbf{Y}$), relative to smaller flows between regions (those that are on the off-diagonal elements of $\mathbf{Y}$). One approach used in empirical studies is to model the diagonal elements of the flow matrix separately (Tiefelsdorf, 2003; Fischer et al., 2006) reflecting the view that intraregional flow elements represent a nuisance. While this approach is fine in the case of the ordinary lineal model, it can have an adverse impact in the case of our spatial interaction model.

LeSage and Pace (2008) suggest an alternative approach to dealing with the large intraregional flow magnitudes which involves adding a specific set of explanatory variables for these observations (a possibly distinct set of covariates). As detailed in LeSage and Thomas-Agnan (2015), this separate model is embedded into the specification by adjusting the matrices $\mathbf{X_o}$, $\mathbf{X_d}$ to have zero values for the observations associated with the intraregional flows in the flow matrix $\mathbf{Y}$. They then propose to introduce an additional matrix of explanatory variables containing non-zero observations for those associated with the intraregional flows (which were set to zero in the matrices $\mathbf{X_o}$, $\mathbf{X_d}$.) We label this matrix of intra-regional factors $\mathbf{X_i}$.

The previously described procedure aims at allowing the parameters associated with the matrices $\mathbf{X_o}$, $\mathbf{X_d}$ to more accurately model the variation in interregional flows, and those associated with the matrix $\mathbf{X_i}$ to capture variation in intraregional flows. To account for intraregional flows, model (3) in its reduced form needs to be modified as follows,

$$(\mathbf{I_{N\times N}} - \rho_o \mathbf{W_o} - \rho_d \mathbf{W_d} + \rho_w \mathbf{W_w})y = \iota_\mathbf{N}\alpha + \mathbf{1_{o=d}}\alpha_i + \mathbf{X_o}\beta_\mathbf{o} + \mathbf{X_d}\beta_\mathbf{d} + \mathbf{X_i}\beta_\mathbf{i} + \mathbf{LX_o}\delta_\mathbf{o} + \mathbf{LX_d}\delta_\mathbf{d} + \mathbf{G}\omega + \epsilon, \quad (4)$$

where $\mathbf{X_o}$ and $\mathbf{X_d}$ have been adjusted to account for the intra-regional flows and $\beta_\mathbf{i}$ is the vector of parameters, of length k_i, with k_i being the number of intraregional explanatory variables and α_i being the intra constant.

Let us denote by $\mathbf{A}(\mathbf{W}) = (\mathbf{I_{N\times N}} - \rho_o \mathbf{W_o} - \rho_d \mathbf{W_d} - \rho_w \mathbf{W_w})^{-1}$ the $N \times N$ inverse filter matrix. It could be any of the other filters for the nine flow submodels described in LeSage and Pace (2009). Note that when there is no lagged explanatory variable, the Durbin model is called a LAG model.

3 Model interpretation

Interpretation of explanatory variable's effects in simultaneous spatial autoregressive models requires the computation of the so-called direct and indirect impacts introduced in LeSage and Pace (2006). Indeed, due to the presence of the filter matrix, the parameters of explanatory variables no longer coincide with the increment of the dependent variable resulting from an additive increment of the explanatory as is the case in the classical (non-spatial) linear models. The two consequences are that this increment is no longer the same for all locations and that an increment of a given explanatory variable at a location may result in non null increments at all other locations. Note that, due to the linearity of the expected value $\mathbb{E}(\mathbf{y} \mid \mathbf{X})$, these increments can be indifferently thought of as infinitesimal or finite increments in the case of continuous covariates. These impact measures have been extended to the spatial interaction models by LeSage and Thomas-Agnan (2015) for models with endogenous spatial interaction and later by LeSage and Fischer (2016) to the case of exogenous spatial interactions. The last two papers further decompose the total effects into origin, destination and network effects, but with the following restrictions: the list of origins coincides with the list of destinations and the characteristics of origins are the same as the characteristics of destination.

In this paper, we examine the impact decomposition and computation in a more general spatial autoregressive interaction model, where the list of origins does not necessarily coincide with the list of destinations and where the sets of origin and destination characteristics are different. To build intuition and for ease of exposition, we split the analysis into three parts. To begin with, Section 3.1 focuses on the square case, that is, when the list of origins coincides with the list of destinations (therefore, $n_o = n_d$ and our flow matrix is square). In turn, Section 3.2 concentrates on the rectangular case, which occurs when the list of origins does not coincide with the list of destinations, but all possible flows between the origin and destination locations are indeed observed (the set of origin-destination couples is a cartesian product). Lastly, Section 3.3 examines the non-cartesian case, that is, when some of the possible flows between the origins and the destinations are not observed or have no meaning.

The fact that a location may be an origin without being a destination (or reversely) makes it difficult to define properly in a unified manner the marginal impact $\frac{\partial \mathbb{E}Y_{o:d}|\mathbf{X}}{\partial X_s}$ of changing characteristic $\mathbf{X}$ at location s on the flow $Y_{o:d}$ from origin o to destination d. Indeed, we need to distinguish three types of characteristics. A characteristic of type oc is a characteristic of origins without being a characteristic of destinations (i.e. corresponds to a column of $\mathbf{OX}$ but not to a column of $\mathbf{DX}$); then changing $\mathbf{X}$ at location s has a meaning only when s is an origin. Respectively, a characteristic of type dc is a characteristic of destinations without being a characteristic of origins (i.e. corresponds to a column of $\mathbf{DX}$ but not to a column of $\mathbf{OX}$); then changing $\mathbf{X}$ at location s has a meaning only when s is a destination. Finally, a characteristic of type odc qualifies both origins and destinations, then it corresponds to a column of $\mathbf{OX}$ and to a column of $\mathbf{DX}$. In this situation, changing $\mathbf{X}$ at location s has a different meaning depending on whether s belongs to $O \setminus D$, to $D \setminus O$ or to $O \cap D$. In the last case, when changing a characteristic that describes both origins and destinations at a location which is both an origin and a destination, the change corresponds to changing a column of $\mathbf{OX}$ and a column of $\mathbf{DX}$ simultaneously. For this reason, and in order to maintain some unity in the formulas, we denote by $\mathbf{Z}$ a generic characteristic which may be of type oc, dc or odc. We denote by $\mathbf{Z}^{\mathbf{OX}}$ a generic column of $\mathbf{OX}$, and by $\mathbf{Z}^{\mathbf{DX}}$ a generic column of $\mathbf{DX}$. A characteristic $\mathbf{X}$ of type oc corresponds to a given $\mathbf{Z}^{\mathbf{OX}}$, a characteristic of type dc, to a given $\mathbf{Z}^{\mathbf{DX}}$ and a characteristic of type odc will be associated to both a given $\mathbf{Z}^{\mathbf{OX}}$ and a given $\mathbf{Z}^{\mathbf{DX}}$. Therefore, for a generic variable $\mathbf{Z}$, the impact on the expected flow $\mathbb{E}Y_{o:d}$ of changing $\mathbf{Z}$ at the location s will be given by $\frac{\partial \mathbb{E}Y_{o:d}|\mathbf{X}}{\partial Z_s}$.

Before entering into the analysis of the impact decomposition in the three cases defined above (the square, the rectangular, and the non-cartesian cases), we exhibit some graphical representations to illustrate these three cases. Figure 1 is a flow map representation, as described in Laurent et al. (2022a), of two toy datasets. Figure 1 a) illustrates the square case: each of the eight sites $s_1, \ldots, s_8$ is both an origin and a destination; each site has seven outflows, seven inflows and one intra-flow. There are hence $N = n^2 = 64$ observed flows. The spatial weight matrices are $\mathbf{OW} = \mathbf{DW} = \mathbf{W}$, with $\mathbf{W}$ being the square contiguity matrix of size $n = n_o = n_d$ (meaning s_1 is a neighbour of s_2, s_2 is a neighbour of s_1 and s_3, etc.). The explanatory variable $\mathbf{Z}$ of type odc observed on $\{s_1, \ldots, s_8\}$ is $(40, 30, 20, 10, 7, 10, 15, 25)'$. In turn, Figure 1 b) illustrates the rectangular case, with: $O = \{o_1, o_2, o_3, o_4, od_1\}$, $D = \{d_1, d_2, od_1\}$. Hence, o_1, o_2, o_3, o_4 are origin sites only, d_1, d_2 are destination sites only and od_1 is both an origin and a destination site. There are $N = n_o \times n_d = 15$ observed flows. $\mathbf{OW}$ is the contiguity matrix for the origin sites, which is of size $n_o \times n_o$, and $\mathbf{DW}$ is the contiguity matrix for the destination sites of size $n_d \times n_d$. The explanatory variable $\mathbf{Z}$ of type odc observed on $\{o_1, o_2, o_3, o_4, od_1, d_1, d_2\}$ is $(40, 20, 10, 7, 10, 15, 25)'$.

Finally, Figure 1 c) illustrates the non cartesian-case. In this example, the list of origins $O = \{o_1, o_2, o_3, o_4, od_1\}$ differs from the list of destinations $D = \{d_1, d_2, d_3, d_4, od_1\}$, with only od_1 being both an origin and a destination site. However, it corresponds to the non-cartesian case because, as Figure 1 c) indicates, origin sites o_1, o_2, o_3, o_4 can reach only 3 destination sites, whereas od_1 can reach all of the destination sites, including itself. There are hence $N = 3+3+3+3+5 = 17$ observed flows. $\mathbf{OW}$ is based on the 3-nearest neighbors, as $\mathbf{DW}$. Site od_1 differs from the other sites as it is a neighbour of all the other sites. The explanatory variable $\mathbf{Z}$ of type odc observed on $\{o_1, o_2, o_3, o_4, od_1, d_1, d_2, d_3, d_4\}$ is $(40, 20, 10, 7, 10, 15, 25, 25, 15)'$. To simulate the flow data in Figure 1 a), b), and c), we consider model (4) assuming the following parameters: $\rho_o = 0.4$, $\rho_d = 0.4$, $\rho_w = -\rho_o \times \rho_d$, $\alpha = 0$, $\beta_o = 0.5$, $\beta_d = 1$, $\beta_i = 2$, $\delta_o = 0.1$, $\delta_d = 0.3$, $\omega = -0.5$.

Please insert Figure 1 here

We now decompose the impacts. To do this, we start by defining the local impacts indexed by $s \in S$ as the result of changing a characteristic at that particular location s. Section 3.1 focuses on the square case, Section 3.2 on the rectangular case, and finally, Section 3.3 presents the local impacts in the non-cartesian case. In the three sections, we then obtain the total impact as sums of local impacts. Our approach contrasts with the previous literature that defines a total impact directly as in LeSage and Pace (2006). In addition, note that a separate analysis of the distribution of the local impacts may be of interest.

3.1 The square case

In the square cartesian case, we have $S = O = D$. A consequence of the fact that the set of origin-destination couples is a cartesian product is that the number of outflows (respectively, inflows) per site $s \in S$ is constant and equal to the number of sites n. The total number of flows is hence $N = n^2$. Let us define the local total impact of changing $\mathbf{Z}$ at a location $s \in S$ as

$$TE(s) = \sum_{o \in S, d \in S} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s}. \quad (5)$$

$TE(s)$ measures the cumulative effect due to changing $\mathbf{Z}$ at a single location s, which makes it a local effect. Following LeSage and Thomas-Agnan (2015) but locally, we can decompose $TE(s)$ into four terms:

$$TE(s) = \frac{\partial \mathbb{E}(Y_{s:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{d \in S, d \neq s} \frac{\partial \mathbb{E}(Y_{s:d} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in S, o \neq s} \frac{\partial \mathbb{E}(Y_{o:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in S, d \in S, o \neq s, d \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \quad (6)$$

The first term is the local intra-regional effect denoted by $IE(s)$; the second term corresponds to the local origin effect or $OE(s)$, that is, the sum of effects on the outflows of s (the flows starting from s), so-called ; the third one is the local destination effect or $DE(s)$, namely the sum of effects on the inflows (the flows arriving to s); finally, the last term corresponds to the local network effect or $NE(s)$ (the flows that do not start from neither arrive to s).

Going from the local to the global effect, let us now define the total impact of changing $\mathbf{Z}$ at all locations on all flows as $TTE = \sum_{s \in S} TE(s)$. It corresponds to the sum of the local total impacts and can be decomposed as follows:

$$TTE = \sum_s TE(s) = \sum_s IE(s) + \sum_s OE(s) + \sum_s DE(s) + \sum_s NE(s) := TIE + TOE + TDE + TNE \quad (7)$$

The total origin (respectively, destination) effect TOE (resp. TDE) is the sum of all changes due to a change in a given characteristic on all the flows that originate (arrive) at the locations of change. The total intra-regional effect can be interpreted as the cumulated impact of changing a given characteristic at all locations on the flows originating and arriving at that location. In turn, the total network effect measures the sum of all changes due to a variation in a given characteristic on all the flows that do not originate nor arrive at the locations of change.

To summarize the local effects by averaging, one needs to count the number of terms in each sum TIE, TOE, TDE and TNE: we have n terms for the intra-regional effect, $n(n-1)$ for the origin and destination effects and $n(n-1)^2$ terms for the network effect. LeSage and Thomas-Agnan (2015) introduce scalar summary measures for the intra-regional, origin, destination, and network effects by normalizing these total effects by the total number of flows N. They hence represent an average effect over all possible flows, or the effect on a typical flow (see Table 1).

Please insert Table 1 here

These scalar measures summarize: i) the average impact of changing a characteristic at a representative origin on all the flows departing from and arriving to that given location (intra-regional effect); ii) the average impact of changing a characteristic at a representative origin on the flows departing from that given location (origin effect); iii) the average impact of changing a characteristic at a representative destination on the flows arriving at that given location (destination effect); iv) the average impact of changing a characteristic at a representative location on the flows neither departing from nor arriving to that given location (network effect).

3.1.1 Alternative local summaries

In order to summarize the effects differently, we propose two graphical representations which provide more detailed information. The first representation considers the decomposition based on the local impacts presented in equation (6). Indeed, after computing, for each location s, the local effects $IE(s), OE(s), DE(s), NE(s)$ and their sum $TE(s)$, results can be presented in a table where the rows correspond to the sites and the columns to the local effects. This table can in turn be used to produce a barplot or a shares' graphic (see for instance Figure 2). These representations have at least two advantages: on the one hand, one can observe the contributions to the total impact of each site $s \in S$; on the other hand, they allow the comparison of the local impacts among the sites. The differences observed between the sites are due to the network structure and will be explained in detail in the next section.

Please insert Figure 2 here

The second additional representation consists in drawing the distributions of each of these local effects across the sites. Using, for example, a non-parametric kernel estimator, one can plot the probability density functions of the local effects in each of the intra-regional, origin, destination and network terms, thus resulting in five density curves for each site $s \in S$. In case of a large number of sites, we can summarize one step further with the distributions of effects over all sites as represented by the last line in Figure 3. It is worth adding that when the number of geographical sites in the dataset is large, it may not be very informative to represent the local effects for each location. However, in some applications, it may be revealing to only focus on the sites of interest. For example, suppose policymakers in a municipality, a region, or a country wish to evaluate the impact of a public policy leading to the increase of a given explanatory variable. In that case, it may be appealing to only exhibit the local effects associated with the location under consideration. One could also imagine grouping several sites (for example, all the municipalities belonging to the same region) that have decided to implement the policy at the same time and then aggregating all the individual local effects to measure the total impacts of this joint decision. In other words, with this type of representation, the user has the possibility to focus on sites linked to a particular economic issue.

Please insert Figure 3 here

One can also compute other scalar summaries which are

- the local average intra-regional $AIE(s)$ effect is equal to $IE(s)$ (because there is only one intra-regional flow)
- the local average origin effect $AOE(s) = \frac{OE(s)}{n-1}$ (because there are $n-1$ flows starting from s)
- the local average destination effect $ADE(s) = \frac{DE(s)}{n-1}$ (because there are $n-1$ flows arriving to s)
- the local average network effect $ANE(s) = \frac{NE(s)}{(n-1)^2}$ (because there are $(n-1)^2$ flows that do not start from, neither arrive to s)

- the local average total effect $ATE(s) = \frac{TE(s)}{n^2}$ (because there are n^2 flows that are impacted by a change of $\mathbf{X}$ at s)

The means of these local averages are identified by a red dot below the density curves on Figure 3.

3.2 The rectangular case

When the list of origins does not coincide with the list of destinations, the definition of the local total impact $TE(s)$ itself is unchanged but its decomposition has to be adapted, depending on the nature of site s (whether it belongs to $O \setminus D$, to $D \setminus O$ or to $O \cap D$). In the example of Figure 1 b), $O \setminus D = \{o_1, o_2, o_3, o_4\}$ are origin sites only, $D \setminus O = \{d_1, d_2\}$ are destination sites only and $O \cap D = \{od_1\}$ is an origin and a destination site.
Depending on the nature of s, the decomposition of $TE(s)$ can be expressed as follows:

- If $s \in O \cap D$ (s is both an origin and a destination site), the decomposition of $TE(s)$ into the four terms $IE(s)$, $OE(s)$, $DE(s)$ and $NE(s)$ can be written as before:

$$\begin{aligned} TE(s) &= IE(s) + OE(s) + DE(s) + NE(s) \\ &= \frac{\partial \mathbb{E}(Y_{s:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{d \in D, d \neq s} \frac{\partial \mathbb{E}(Y_{s:d} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in O, o \neq s} \frac{\partial \mathbb{E}(Y_{o:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in O, d \in D, o \neq s, d \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \end{aligned} \quad (8)$$

  In the example of Figure 1 (b), this is only the case for site od_1. Precisely, the change of the flow $od_1 : od_1$ is included in $IE(od_1)$, the changes of the flows $od_1 : d_2$ and $od_1 : d_3$ are included in $OE(od_1)$, the changes of the flows $o_1 : od_1$, $o_2 : od_1$, $o_3 : od_1$, and $o_4 : od_1$ are included in $DE(od_1)$ and the changes of all other flows are included in $NE(od_1)$.

- If $s \in O \setminus D$ (s is an origin site only), there is no local intra effect (as the intra flow does not exist), neither local destination effect (as inflows do not exist). $TE(s)$ can then be written as follows

$$TE(s) = OE(s) + NE(s) = \sum_{d \in D} \frac{\partial \mathbb{E}(Y_{s:d} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in O, d \in D, o \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \quad (9)$$

  In our example of Figure 1 (b), this is the case for sites o_1, o_2, o_3, o_4. For example, for site o_1, the changes of the flows $o_1 : d_1$, $o_1 : d_2$, $o_1 : od_1$ are included in $OE(o_1)$ and the changes of all other flows are included in $NE(o_1)$.

- If $s \in D \setminus O$ (s is a destination site only), there is no local intra effect (as the intra flow does not exist) neither local origin effect (as the outflows do not exist). $TE(s)$ can be written as follows

$$TE(s) = DE(s) + NE(s) = \sum_{o \in O} \frac{\partial \mathbb{E}(Y_{o:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in O, d \in D, d \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \quad (10)$$

  In the example of Figure 1 (b), this is the case for sites d_1, d_2. For example, for site d_1, the changes of the flows $o_1 : d_1$, $o_2 : d_1$, $o_3 : d_1$, $o_4 : d_1$, $od_1 : d_1$ are included in $DE(d_1)$ and the changes of all other flows are included in $NE(d_1)$.

We denote the total impact of changing $\mathbf{Z}$ at all locations on all flows by $TTE = \sum_{s \in S} TE(s)$, which is the sum of the local total impacts. We can decompose TTE as follows:

$$
\begin{aligned}
TTE =& \sum_{s \in S} TE(s) \\
=& \sum_{s \in S \setminus D} TE(s) + \sum_{s \in S \setminus O} TE(s) + \sum_{s \in O \cap D} TE(s) \\
=& \sum_{s \in S \setminus D} (OE(s) + NE(s)) + \sum_{s \in S \setminus O} (DE(s) + NE(s)) + \sum_{s \in O \cap D} (IE(s) + OE(s) + DE(s) + NE(s)) \\
=& \sum_{s \in O \cap D} IE(s) + \sum_{s \in O} OE(s) + \sum_{s \in D} DE(s) + \sum_{s \in S} NE(s) \qquad (11) \\
=& TIE + TOE + TDE + TNE
\end{aligned}
$$

The last two equalities in (11) coincide with the equation (7) for the TTE in the square case $O = D = S$.

To summarize these effects by averaging, we need to provide some additional definitions. Let $nd(s)$ be the number of outflows i.e. the number of destinations that one can reach from site $s \in O$. Similarly, let $no(s)$ be the number of inflows i.e. the number of origins that can lead to destination $s \in D$. Let

$$
n_S = \begin{cases} n_o, & \text{if } X \text{ is of type oc (an origin characteristic only)} \\ n_d, & \text{if } X \text{ is of type dc (a destination characteristic only)} \\ n_{O \cup D}, & \text{if } X \text{ is of type odc (an origin and destination characteristic)} \end{cases}
$$

where $n_{O \cap D}$ is the number of elements of $O \cap D$. Finally, we continue to denote N by the total number of flows, with this being $N = \sum_{s \in O} nd(s) = \sum_{s \in D} no(s)$.

The number of terms in $TE(s)$ is equal to N. The number of terms in $IE(s)$ is equal to:

$$
\begin{cases} 1, & \text{if } s \in O \cap D \text{ (and intra flow } s - s \text{ exists)} \\ 0, & \text{otherwise.} \end{cases}
$$

The number of terms in $OE(s)$ is equal to

$$
\begin{cases} nd(s), & \text{if } s \in O \setminus D \\ nd(s) - 1, & \text{if } s \in O \cap D \\ 0, & s \notin O. \end{cases}
$$

The number of terms in $DE(s)$ is equal to

$$
\begin{cases} no(s), & \text{if } s \in D \setminus O \\ no(s) - 1, & \text{if } s \in O \cap D \\ 0, & s \notin D. \end{cases}
$$

The number of terms in $NE(s)$

$$\begin{cases} N - nd(s), & \text{if } s \in O \setminus D \\ N - no(s), & \text{if } s \in D \setminus O \\ N - no(s) - nd(s) + 1, & \text{if } s \in O \cap D \ . \end{cases}$$

Let us now turn our attention to the number of terms in each component of the total effects. The number of terms in TE is $n_S N$ when none of them is void. The number of terms in TIE is $n_{O\cap D}$ when none of the intra flows is void. The number of terms in $TOE = \sum_{s\in O} OE(s)$ is:

$$\begin{cases} \sum_{s\in O} nd(s) - n_{O\cap D} = N - n_{O\cap D}, & \text{if } \mathbf{Z} \text{ is of type oc or odc} \\ \sum_{s\in O\cap D} nd(s) - n_{O\cap D}, & \text{if } \mathbf{Z} \text{ is of type dc} \end{cases}$$

The number of terms in $TDE = \sum_{s\in D} DE(s)$ is,

$$\begin{cases} \sum_{s\in D} no(s) - n_{O\cap D} = N - n_{O\cap D}, & \text{if } \mathbf{Z} \text{ is of type dc or odc} \\ \sum_{s\in O\cap D} no(s) - n_{O\cap D}, & \text{if } \mathbf{Z} \text{ is of type oc.} \end{cases}$$

Finally, the number of terms in TNE is $\sum_{s\in O\cap D}(N-no(s)-nd(s)+1)+\sum_{s\in O\setminus D}(N-nd(s))+\sum_{s\in D\setminus O}(N-no(s))$. Note that the 2nd term in the previous expression for the TNE disappears if $\mathbf{Z}$ is of type dc (as $S = D$) and the 3rd term disappears if $\mathbf{Z}$ is of type oc (as $S = O$). Hence, the number of terms in TNE is:

$$\begin{cases} (n_o - 1)N + n_{O\cap D} - \sum_{s\in O\cap D} no(s), & \text{if } \mathbf{Z} \text{ is of type oc} \\ (n_d - 1)N + n_{O\cap D} - \sum_{s\in O\cap D} nd(s), & \text{if } \mathbf{Z} \text{ is of type dc} \\ (n_{O\cup D} - 2)N + n_{O\cap D}, & \text{if } \mathbf{Z} \text{ is of type odc.} \end{cases}$$

Table 2 summarizes the number of local terms and total terms for our example, using the above formulas. In the cartesian case, we can group the sites with respect to their nature. Indeed, in that case, the number of terms is identical for each site belonging to one of the three situations: $O \setminus D, D \setminus O, O \cap D$.

Please insert Table 2 here

3.2.1 Alternative local summaries

We propose the same representations to summarize the local origin, destination and network effects as in the square case: the contributions of each location to the corresponding total effects and the visualization of the distributions of these local effects. Regarding the first representation, Figure 4 exhibits, for each location $s \in S$, the decomposition of the local total effect into the different effects $IE(s)$, $OE(s)$, $DE(s)$ and $NE(s)$. It is interesting to group the sites according to their nature: $s \in O \setminus D$, $s \in D \setminus O$ and $s \in O \cap D$.

Please insert Figure 4 here

Figure 5 exhibits the non-parametric kernel estimators for the densities of the local effects. The Figure distinguishes among the intra, origin, destination, network and total effects resulting in five density curves per site and for all sites. Finally, Table 3 presents the local average effects, with respect to the nature of s.

Please insert Figure 5 here

Please insert Table 3 here

3.3 The non cartesian case

In the non-cartesian case, some of the possible flows between the origins and the destinations are not observed or have no meaning. In this situation, we define for each site s two subsets of O and D: if s is a destination, $O(s)$ contains the list of origins of flows arriving to destination s and if s is an origin, $D(s)$ is the list of destinations of flows starting from origin s. As in the previous section, we denote by $no(s)$ (respectively, $nd(s)$) the size of $O(s)$ (respectively, $D(s)$). In the example of Figure 1 c), o_1 is an origin site only that can reach only the sites d_1, d_3, od_1. Hence, $O(o_1) = \varnothing$, $D(o_1) = \{d_1, d_3, od_1\}$, $no(o_1) = 0$ and $nd(o_1) = 3$. d_1 is a destination site only that is reached only by the sites o_1, o_2, od_1. Hence, $O(o_1) = \{o_1, o_2, od_1\}$, $D(o_1) = \varnothing$, $no(o_1) = 3$ and $nd(o_1) = 0$. od_1 is both an origin and destination site that can reach and can be reached by all the sites. Hence, $O(od_1) = \{o_1, o_2, o_3, o_4, od_1\}$, $D(od_1) = \{d_1, d_2, d_3, d_4, od_1\}$, $no(od_1) = 5$ and $nd(od_1) = 5$.

The definition of the different impacts presented in the rectangular case remains unchanged, but its decomposition needs to be adapted. The decomposition of $TE(s)$ can be expressed as follows depending on the nature of s:

- If $s \in O \cap D$ (s is both an origin and a destination site) and under the hypothesis that the intra-flow $s - s$ exists, the decomposition of $TE(s)$ into the four terms $IE(s)$, $OE(s)$, $DE(s)$ and $NE(s)$ can be written as before:

$$\begin{aligned} TE(s) &= IE(s) + OE(s) + DE(s) + NE(s) \\ &= \frac{\partial \mathbb{E}(Y_{s:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{d \in D(s), d \neq s} \frac{\partial \mathbb{E}(Y_{s:d} \mid \mathbf{X})}{\partial Z_s} + \sum_{o \in O(s), o \neq s} \frac{\partial \mathbb{E}(Y_{o:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{t \in S, t \neq s} \sum_{o \in O(t), d \in D(t), o \neq s, d \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \end{aligned} \tag{12}$$

- If $s \in O \setminus D$ (s is an origin site only), there is no local intra effect (as there is no intra flow) neither local destination effect (as there are no inflows). $TE(s)$ can be written as follows

$$TE(s) = OE(s) + NE(s) = \sum_{d \in D(s)} \frac{\partial \mathbb{E}(Y_{s:d} \mid \mathbf{X})}{\partial Z_s} + \sum_{t \in S, t \neq s} \sum_{o \in O(t), d \in D(t), o \neq s, d \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \tag{13}$$

- If $s \in D \setminus O$ (s is a destination site only), there is no local intra effect (as there is no intra flow) neither local origin effect (as there are no outflows). $TE(s)$ can be written as follows

$$TE(s) = DE(s) + NE(s) = \sum_{o \in O(s)} \frac{\partial \mathbb{E}(Y_{o:s} \mid \mathbf{X})}{\partial Z_s} + \sum_{t \in S, t \neq s} \sum_{o \in O(t), d \in D(t), o \neq s, d \neq s} \frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} \tag{14}$$

We can compute the total impact TTE in the non cartesian case using formula (11) but replacing the terms $TE(s)$, using formulas (12), (13) and (14). Finally, one can use the same formula proposed in the rectangular case to compute the number of terms in $OE(s)$, $DE(s)$, $NE(s)$ and $TE(s)$. In the non-cartesian case, the summaries $IE(s), OE(s), DE(s)$ and $NE(s)$ may have different number of terms according to the nature of the corresponding site, which requires counting the number of terms site by site. To conclude, the formulas of the local average effects presented in Table 3 remain unchanged in the non cartesian-case.

4 Impact computations

Cartesian case

We now provide details about the impact computations in the SDM model (3). Using the reduced equation of this model, it is easy to see that the impacts on the flow $Y_{o:d}$ due to a change of Z at location s equals:

$$\begin{aligned}\frac{\partial \mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} =& \beta_o^Z \sum_{l \in D} A(W)_{o:d;s:l} + \beta_d^Z \sum_{l \in O} A(W)_{o:d;l:s} \\ &+ \delta_o^Z \sum_{t \in O} \sum_{l \in D} OW_{ts} A(W)_{o:d;t:l} + \delta_d^Z \sum_{t \in D} \sum_{l \in O} DW_{ts} A(W)_{o:d;l:t} \end{aligned} \quad (15)$$

where $\beta_o^Z, \beta_d^Z, \delta_o^Z, \delta_d^Z$ are the parameters associated to covariate Z. Calculating (15) a priori requires the full computation of $\mathbf{A(W)}$, that is, the inversion of the $N \times N$ matrix $(\mathbf{I_{N\times N}} - \rho_o \mathbf{W_o} - \rho_d \mathbf{W_d} - \rho_w \mathbf{W_w})$. The computation may be demanding in time and in memory especially when N is large (which is often the case with spatial flows). To avoid computing $\mathbf{A(W)}$, we rely on some properties of this matrix that hold in the cartesian case (square or rectangular). In what follows, we first present two theorems regarding the computation of the terms $\sum_{l \in D} A(W)_{o:d;s:l}$ and $\sum_{l \in O} A(W)_{o:d;l:s}$ in equation (15) and then explain how we apply them to simplify the computations of $\frac{\partial \mathbb{E}(Y_{o:d}|\mathbf{X})}{\partial Z_s}$. Lastly, we discuss how to adapt computations in the case of a SDM including an intraregional term.

For the next three theorems, we use the following set of assumptions (A)

- the model is given by (3)
- **OW** and **DW** are row-normalized
- the data is cartesian

Theorem 1. *Under assumptions (A), for any d in D, and for $o, s \in O$, we have $\sum_{l \in D} A(W)_{o:d;s:l} = \lambda_{os}$, where λ_{os} is the element os of the $n_o \times n_o$ matrix $\mathbf{\Lambda} = (\mathbf{I_{n_o}} - \mathbf{\Sigma_M})^{-1}$ with $\mathbf{M} = \mathbf{I_{N\times N}} - \mathbf{A(W)} = \rho_o \mathbf{W_o} + \rho_d \mathbf{W_d} + \rho_w \mathbf{W_w}$. For $i, j \in O$, the element σ_{ij}^M of the matrix $\mathbf{\Sigma_M}$ is the row-sum (for any row) of the $n_d \times n_d$ matrix $\mathbf{M_{ij}}$*

corresponding to a block of as follows

$$\mathbf{M} = \begin{pmatrix} \mathbf{M_{o_1o_1}} & \dots & \mathbf{M_{o_1o_{no}}} \\ \mathbf{M_{o_2o_1}} & \dots & \mathbf{M_{o_2o_{no}}} \\ \vdots & & \vdots \\ \mathbf{M_{o_{no}o_1}} & \dots & \mathbf{M_{o_{no}o_{no}}} \\ \end{pmatrix} = \rho_o \begin{pmatrix} ow_{11}\mathbf{I_{n_d}} & \dots & ow_{1n_o}\mathbf{I_{n_d}} \\ ow_{21}\mathbf{I_{n_d}} & \dots & ow_{2n_o}\mathbf{I_{n_d}} \\ \vdots & & \vdots \\ ow_{n_o1}\mathbf{I_{n_d}} & \dots & ow_{n_on_o}\mathbf{I_{n_d}} \\ \end{pmatrix} + \rho_d \begin{pmatrix} \mathbf{DW} & 0 & \dots & 0 \\ 0 & \mathbf{DW} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \mathbf{DW} \\ \end{pmatrix}$$

$$+ \rho_w \begin{pmatrix} ow_{11}\mathbf{DW} & \dots & ow_{1n_o}\mathbf{DW} \\ ow_{21}\mathbf{DW} & \dots & ow_{2n_o}\mathbf{DW} \\ \vdots & & \vdots \\ ow_{n_o1}\mathbf{DW} & \dots & ow_{n_on_o}\mathbf{DW} \\ \end{pmatrix} \tag{16}$$

Proof. We have $(\mathbf{I_N} - \mathbf{M})^{-1} = \mathbf{I_N} + \sum_k \mathbf{M^k}$. Let us define $\sigma_{ij}^{(k)}$ to be the sum of any row of the matrix $\mathbf{M^k_{ij}}$. When $k = 1$, it is obvious from equation (16) that the sum of any row of the matrix $\mathbf{M_{ij}}$ is constant and equal to $\sigma_{ij}^{(1)} = \rho_o ow_{ij} + \rho_d \mathbf{1}_{i=j} + \rho_w ow_{ij}$. Now, consider two block matrices $\mathbf{A}$ and $\mathbf{B}$ with the same dimensions $N \times N$ and with n_o row/column partitions, each block being of size $n_d \times n_d$. Let us assume moreover that the sum of any row of the block $\mathbf{A_{ij}}$ (resp: $\mathbf{B_{ij}}$) is constant and equal to σ_{ij}^A (resp: σ_{ij}^B). Let $\mathbf{C} = \mathbf{AB}$ be the block matrix that has the same dimension and partition as $\mathbf{A}$ and $\mathbf{B}$. One can show that the sum of any row of the block $\mathbf{C_{ij}}$ is constant and equal to $\sum_k \sigma_{ik}^A \sigma_{kj}^B$. Indeed, since $\mathbf{C_{ij}} = \sum_k \mathbf{A_{ik}B_{kj}}$, the sum of the rows of $\mathbf{C_{ij}}$ are obtained by $\mathbf{C_{ij}1_{n_d}} = \sum_k \mathbf{A_{ik}}(\mathbf{B_{kj}1_{n_d}}) = \sum_k \mathbf{A_{ik}}(\sigma_{kj}^B \mathbf{1_{n_d}}) = \sum_k (\mathbf{A_{ik}1_{n_d}})\sigma_{kj}^B = (\sum_k \sigma_{ik}^A \sigma_{kj}^B)\mathbf{1_{n_d}}$. Using this property recursively for $k = 2, ..., \infty$, we can see M^k as the product of M^{k-1} times M that are two block matrices with the same dimensions $N \times N$ and with n_o row/column partitions, each block being of size $n_d \times n_d$. Furthermore, the sum of any row of the block $\mathbf{M^{k-1}_{ij}}$ is constant and equal to $\sigma_{ij}^{(k-1)}$ and the sum of any row of the block $\mathbf{M_{ij}}$ is constant and equal to $\sigma_{ij}^{(1)}$. Hence, the sum of any row of the block $\mathbf{M^k_{ij}}$ is constant and equal to $\sum_k \sigma_{ik}^{(k-1)} \sigma_{kj}^{(1)}$. The consequence is that the sum of any row of the block $(\mathbf{I_N} - \mathbf{M})^{-1}_{ij}$ is constant and equal to $\lambda_{ij} = \mathbf{1}_{i=j} + \sum_k \sigma_{ij}^{(k)}$. $\square$

Theorem 2. *Under assumptions (A), for any $o \in O$, and for $d, s \in D$, we have $\sum_{l \in O} A(W)_{o:d;l:s} = \gamma_{ds}$, where γ_{ds} is the element ds of the $n_d \times n_d$ matrix $\mathbf{\Gamma} = (\mathbf{I_{n_d}} - \mathbf{\Sigma_N})^{-1}$ and for $i, j \in D$, the element σ_{ij}^N of the matrix $\mathbf{\Sigma_N}$*

is the row-sum (for any row) of the matrix $\mathbf{N_{ij}}$ *corresponding to a block of* $\mathbf{N}$ *as follows*

$$\mathbf{N} = \begin{pmatrix} \mathbf{N_{d_1d_1}} & \dots & \mathbf{N_{d_1d_{nd}}} \\ \mathbf{N_{d_2d_1}} & \dots & \mathbf{N_{d_2d_{do}}} \\ \vdots & & \vdots \\ \mathbf{N_{d_{nd}d_1}} & \dots & \mathbf{N_{d_{nd}d_{nd}}} \end{pmatrix} = \rho_o \begin{pmatrix} \mathbf{OW} & 0 & \dots & 0 \\ 0 & \mathbf{OW} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \mathbf{OW} \end{pmatrix} + \rho_d \begin{pmatrix} dw_{11}\mathbf{I_{n_o}} & \dots & dw_{1n_d}\mathbf{I_{n_o}} \\ dw_{21}\mathbf{I_{n_o}} & \dots & dw_{2n_d}\mathbf{I_{n_o}} \\ \vdots & & \vdots \\ dw_{n_d1}\mathbf{I_{n_o}} & \dots & dw_{n_dn_d}\mathbf{I_{n_o}} \end{pmatrix}$$
$$+ \rho_w \begin{pmatrix} dw_{11}\mathbf{OW} & \dots & dw_{1n_o}\mathbf{OW} \\ dw_{21}\mathbf{OW} & \dots & dw_{2n_o}\mathbf{OW} \\ \vdots & & \vdots \\ dw_{n_o1}\mathbf{OW} & \dots & dw_{n_on_o}\mathbf{OW} \end{pmatrix} \tag{17}$$

As for matrix $\mathbf{M}$, note that the matrix $\mathbf{N}$ corresponds to the matrix $(\rho_o\mathbf{W_o} + \rho_d\mathbf{W_d} + \rho_w\mathbf{W_w})$ that has been ordered with respect to the destination sites. The proof is similar to the previous proof.

The following theorem combines the results of Theorem 1 and 2 to simplify the computation of $\frac{\partial\mathbb{E}(Y_{o:d}|\mathbf{X})}{\partial Z_s}$.

Theorem 3. *Under assumptions (A), for any* $o \in O$ *and* $d \in D$ *and for any* $s \in O \cup D$*, the impact on flow* $o : d$ *due to a change of* Z *at location* s *can be written as follows:*

$$\frac{\partial\mathbb{E}(Y_{o:d} \mid \boldsymbol{X})}{\partial Z_s} = \beta_o^Z\lambda_{os} + \beta_d^Z\gamma_{ds} + \delta_o^Z\sum_{t\in O} OW_{ts}\lambda_{ot} + \delta_d^Z\sum_{t\in D} DW_{ts}\gamma_{dt} \tag{18}$$

Importantly, Theorem 3 shows that computing $\frac{\partial\mathbb{E}(Y_{o:d})}{\partial Z_s}$ only requires the evaluation of the two matrices $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$, of sizes $n_o \times n_o$ and $n_d \times n_d$, respectively, instead of computing the $N \times N$ matrix $\mathbf{A}(\mathbf{W})$. Another interesting consequence of Theorem 3 is that, since $\beta_o^Z\lambda_{os} + \delta_o^Z\sum_{t\in O} OW_{ts}\lambda_{ot}$ does not depend on d, all flows starting from a given origin o are affected in the same way by the origin covariates $\mathbf{X_o}$ and $\mathbf{LX_o}$. Similarly, since $\beta_d^Z\gamma_{ds} + \delta_d^Z\sum_{t\in D} DW_{ts}\gamma_{dt}$ does not depend on o, all flows arriving to d are impacted in the same way by the destination covariates $\mathbf{X_d}$ and $\mathbf{LX_d}$.

Adding the intraregional characteristics $\mathbf{X_i}$ in the SDM specification (equation (4)) requires the computation of $\beta_i^Z\mathbf{A}(\mathbf{w})_{\mathbf{o:d,s:s}}$. In other words, to evaluate the impacts on all flows $Y_{o:d}$, due to a change of $\mathbf{Z}$ at location s, we need to compute the ss column of matrix $\mathbf{A}(\mathbf{w})$. To avoid the full computation of $\mathbf{A}(\mathbf{w})$ for extracting column ss, one can alternatively solve the equation:

$$(\mathbf{I_{N\times N}} - \rho_o\mathbf{W_o} - \rho_d\mathbf{W_d} + \rho_w\mathbf{W_w})x = \mathbf{1_{ss}} \tag{19}$$

where the N-vector $\mathbf{1_{ss}}$ equals to 1 for flow ss and to 0 otherwise.

Let the N vector $\psi_\mathbf{s}$ be the solution of the previous equation. In the model specification with intraregional characteristics, under the condition that the intraregional flow ss exists, equation (18) then becomes:

$$\frac{\partial\mathbb{E}(Y_{o:d} \mid \mathbf{X})}{\partial Z_s} = \beta_o^Z\lambda_{os} + \beta_d^Z\gamma_{ds} + \beta_i^Z\psi_{\mathbf{s}\,o:d} + \delta_o^Z\sum_{t\in O} OW_{ts}\lambda_{ot} + \delta_d^Z\sum_{t\in D} DW_{ts}\gamma_{dt}. \tag{20}$$

Finally, to get the different local impacts in formulas (8), (9), or (10), one has to solve equation (19) for each $s \in O \cap D$ (under the condition that the intraregional flow ss exists) and use equation (20).

Adjustment needed when the dependent variable is in log

One issue is that the effects we are computing are derivatives of expected values. Assume $\log(Y)$ is modelled instead of Y (as is often the case) but Y still is the variable of interest. We know that the expected value and the log do not commute, hence, information about the derivative of the expected log of Y is not informative about the derivative of the expected value of Y. One way out is to use another interpretation: expected values coincide with predicted values in the ordinary linear model in the sense of the best unbiased prediction. If we define the impact as an impact on the predicted value, then we can argue that the exponential of the predicted value of $\log(Y)$ can be used as an admissible prediction for Y, i.e. $\widehat{\log(Y_{o:d})} = \log(\widehat{Y_{o:d}})$ and hence $\widehat{Y_{o:d}} = \exp(\widehat{\log(Y_{o:d})})$. Then if we take $TE(s)$, for example in equation (5), and adapt it to predictions, we can write

$$TE(s) = \sum_{o \in O, d \in D} \frac{\partial \widehat{Y_{o:d}}}{\partial \log \widehat{Y_{o:d}}} \frac{\partial (\widehat{\log Y_{o:d}})}{\partial Z_s} = \sum_{o \in O, d \in D} \widehat{Y_{o:d}} \frac{\partial (\widehat{\log Y_{o:d}})}{\partial Z_s} \tag{21}$$

Note that formula (21) is a version of (5) weighted by the flows.

Non-cartesian case

In the non-cartesian case, the spatial weight matrices $\mathbf{W_o}$, $\mathbf{W_d}$, and $\mathbf{W_w}$ are not anymore the cartesian products of the initial weight matrices $\mathbf{DW}$ and $\mathbf{OW}$. Still, they can be obtained by removing the rows and columns of the flows which are missing or which do not exist, without forgetting to renormalize by row the new matrices. For that reason, Theorems 1 and 2 do not extend easily to the non-cartesian framework. The non-cartesian framework requires *a priori* the full computation of $\mathbf{A(W)}$ in equation (15).

5 Prediction computations

Following Goulard et al. (2017), the trend-corrected (TC) formula adapted to model (3) yields the following prediction :

$$\hat{\mathbf{Y}}^{TC} = \mathbf{A(W)}(\iota_\mathbf{N}\hat{\alpha} + \mathbf{X_o}\hat{\beta}_\mathbf{o} + \mathbf{X_d}\hat{\beta}_\mathbf{d} + \mathbf{LX_o}\hat{\delta}_\mathbf{o} + \mathbf{LX_d}\hat{\delta}_\mathbf{d} + \mathbf{G}\hat{\omega}), \tag{22}$$

The following Theorem is an extension of Theorems 1 and 2, applied to the prediction problem for Model (3).

Theorem 4.

$$\begin{aligned} \hat{Y}^{TC}_{o:d} =& \hat{\alpha} \sum_s \lambda_{os} + (\mathbf{\Lambda} \times \mathbf{OX} \times \hat{\beta}_\mathbf{o})_o + (\mathbf{\Gamma} \times \mathbf{DX} \times \hat{\beta}_\mathbf{d})_d + (\mathbf{OW} \times \mathbf{\Lambda} \times \mathbf{OX} \times \hat{\delta}_\mathbf{o})_o + (\mathbf{DW} \times \mathbf{\Gamma} \times \mathbf{DX} \times \hat{\delta}_\mathbf{d})_d \\ &+ \mathbf{A(W)} \times \mathbf{G}_{od} \times \hat{\omega}. \end{aligned} \tag{23}$$

The advantage of relying on equation (23) is the reduction in the computation time compared to formula (22) provided the term $\hat{\omega}$ is null; in other words, if there are no variables characterizing the pairs of origin-destination. Indeed, in that case it is not necessary to compute the last term containing $\mathbf{A(W)}$. Instead, if $\hat{\omega} \neq \mathbf{0}$, equation (23) remains appealing from a practitioner point of view as it avoids storing the matrices $\mathbf{X_o}, \mathbf{X_d}, \mathbf{LX_o}, \mathbf{LX_d}$. In the model specification with intraregional characteristics, equation (23) becomes:

$$\hat{Y}^{TC}_{o:d} = \hat{\alpha}\sum_s \lambda_{os} + (\mathbf{\Lambda} \times \mathbf{OX} \times \hat{\beta}_{\mathbf{o}})_o + (\mathbf{\Gamma} \times \mathbf{DX} \times \hat{\beta}_d)_d + (\mathbf{OW} \times \mathbf{\Lambda} \times \mathbf{OX} \times \hat{\delta}_o)_o + (\mathbf{DW} \times \mathbf{\Gamma} \times \mathbf{DX} \times \hat{\delta}_{\mathbf{d}})_d$$
$$+\mathbf{A}(\mathbf{W}) \times (\mathbf{G}_{od} \times \hat{\omega} + \mathbf{X_i} \times \hat{\beta}_{\mathbf{i}}). \tag{24}$$

In this case, there is no gain in computation time; still, we do not need to store the full matrices $\mathbf{X_o}, \mathbf{X_d}, \mathbf{LX_o}, \mathbf{LX_d}$.

6 Case study

Our supplementary material (see XX) contains the **R** code that allows to reproduce the results we present in this paper. In addition, we have implemented **R** functions available on Github, which allow to compute the impacts for the different model specifications presented above: the model without intra-regional terms (equation (18)) and the model specification including intra-regional terms (equation (20)). Also, we have developed functions that implement prediction formulas for the model specification without and with intra-regional terms (equations (23) and (24), respectively). These functions are useful to compute the impact decomposition even when the dependent variable is transformed by the logarithm. To illustrate the impact computations in this general spatial autoregressive interaction model, we consider two case studies: the first one relies on a sample of bilateral remittance flows used in Laurent et al. (2022b). The second one consists of home-to-work commuting flows within the 71 municipalities around Paris. This second dataset is part of the **spflow R** package for modeling spatial interactions. In what follows, we detail the first case study. The analysis of the second case study can be found in Section 5 of the supplementary material.

The aim of the first case study is to provide an illustration of the estimation of a SIM specification and of the impact computations in a rectangular dataset. Precisely, our sample comprises 14 macro economic regions, which act as origins, and 15 macro economic regions, which are destinations,[6] for which workers' remittances are reported in 2017.[7] In this rectangular dataset, 14 macro economic zones act both as origin and destination regions (South Asia is a recipient zone only). The bilateral remittance dataset hence involves $14 \times 15 = 210$ flows, and 14 intra-regional remittance flows. The map flow in Figure 6 exhibits our data. The way we read it is the following. On the one hand, the left (respectively, right) barplot located at a given zone gives information on the total outflows (total inflows) for the given zone. On the other hand, the region-pair remittance flows are represented

[6] In their initial study, Laurent et al. (2022b) work with OD remittance flows by country pairs. In contrast, in this work, we focus on aggregated remittance flows by macro-economic regions. The higher order of aggregation we propose in this paper is to facilitate the interpretation and visualisation of the the impact decomposition.

[7] The macro economic zones are: **North America** (Canada, United States), **Central America** (Costa Rica, Guatemala, Honduras, Mexico, Nicaragua, Panama, El Salvador, Belize), **South America** (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Peru, Paraguay, Uruguay, Venezuela, Guyana, Suriname), **Carribean** (Antigua and Barbuda, Aruba, Bahamas, The, Dominican Republic, St. Kitts and Nevis, St. Lucia, Barbados, Dominica, Grenada, Haiti, Jamaica, Trinidad and Tobago), **European Union pre 2004 expansion** (Austria, Belgium, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Italy, Netherlands, Portugal, Sweden, Luxembourg), **European Union post 2004 expansion** (Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Malta, Poland, Romania, Slovenia, Slovakia, Croatia, Lithuania, Latvia), **Europe other** (Bosnia and Herzegovina, Switzerland, Iceland, Norway, Albania, Macedonia, Serbia), **Ex-soviet union** (Russia, Ukraine, Armenia, Georgia, Moldova), **Middle East** (Israel, Jordan, Turkey, Iran, Lebanon, Oman, Saudi Arabia, Syria, Yemen), **North Africa** (Egypt, Libya, Mauritania, Algeria, Morocco, Tunisia), **Sub-Saharan** (Kenya, South Africa, Angola, Burkina Faso, Burundi, Benin, Ivory Coast, Cameroon, Cape Verde, Ethiopia, Ghana, Gambia, The, Guinea, Guinea-Bissau, Liberia, Mali, Mauritius, Malawi, Mozambique, Niger, Nigeria, Namibia, Rwanda, Sudan, Sierra Leone, Senegal, Togo, Uganda, Zambia), **South Asia** (Bangladesh, India, Sri Lanka, Nepal, Pakistan), **East Asia** (Japan, China, South Korea, Mongolia), **South East Asia** (Malaysia, Philippines, Thailand, Indonesia, Cambodia, Laos), **Pacific** (Australia, New Zealand, Fiji)

by arrows, whose color depends on the origin region and whose width is proportional to the flow value.

There are several elements to highlight from the map flow. First, the closer the zones are to each other, the larger the remittance flows are. Second, the three biggest flows correspond to two inter-regional flows (North America-Central America and North America-East Asia, with respectively 48 and 25 billion dollars) and one intra-regional flow (European union before 2004 expansion, with 47 billion dollars). Third, North America and the European Union (before 2004 expansion) have sent a large fraction of total remittances worldwide (precisely, 75% of of world remittances, which amounts to 391 billions dollars). Fourth, Central America, East Asia, South East Asia and South Asia have received a lot of remittances, but they have not sent so much money. Finally, North America, is the largest sender of remittances; however, it has not received an equally considerable amount of money.

Please insert Figure 6 here

To model the bilateral remittance flows, we consider two different model specifications: a SDM which does not include intra-regional terms (equation 3) and a SDM specification that includes intra-regional characteristics and an intra-regional constant (equation 4). In both cases, the dependent variable is the logarithm of the remittance flows between macro-economic regions for normalizing reasons.

Following Laurent et al. (2022b), we consider the following explanatory variables. As type odc determinants (characterizing both the origin and the destination countries in the remittance flows), we consider electricity use per capita (in logarithm), which is a proxy for economic activity (source Central Intelligence Agency, USA), and the total logged population, which proxies for the size of the macro-economic zones. As type oc variables (characterizing only the origin countries), we include the average transfer fee associated with sending remittances from each origin macro-economic zone (source: Western Union website). These transaction costs are computed assuming a remittance of 200 Euros, they are quoted in nominal terms. As type dc control (characterizing only the destination countries), we include the political stability index (source World Bank). Lastly, we include one variable typically used in the gravity models to characterize the pair of OD zones, namely, the average distance among the pairs of origin and destination countries involved in each pair of macro-economic zones.

To define neighborhood between macro-economic zones, we consider proximity in the set of origin or destination regions, with the proximity being defined based on the $3-$nearest neighbors. As spatially lagged control variables, we include the spatial lag of electricity use per capita, both at the origin and the destination (spatial neighborhood matrices $\mathbf{W_o}$ and $\mathbf{W_d}$, respectively). Please refer to Laurent et al. (2022b) for details on the variables, data sources, and the descriptive statistics of the control variables.

In order to illustrate the impact computation in this general spatial interaction model, Table 4 first presents the model estimates of the spatial Durbin model considering the model specifications in equations (3) (left panel) and (4) (right panel). For comparison, the table also exhibits the model estimates of the gravity model and of the SLX model (which is the gravity model with spatially lagged explanatory variables). For estimation of the spatial models, we follow a Bayesian approach coupled with Markov chain Monte Carlo (MCMC) with two neighborhood matrices, at the origin and at the destination (matrices $\mathbf{W_o}$ and $\mathbf{W_d}$, respectively). The Bayesian MCMC estimates are based on 5,500 draws with 2,500 omitted draws for start-up. Inference is made from the distribution of the parameters obtained with the 3,000 retained replications. More precisely, Table 4 reports the average parameter estimates; in parentheses, it reports the t-statistics, which are the ratio between the mean and the standard deviation computed over the 3,000 replications.

Results in Table 4 show, first, that the estimated spatial dependence parameters at the origin in the two model estimates are statistically significant, which suggests the presence of origin-based spatial dependence in the remittance flows between macro-economic zones. Second, the coefficients related to the intra-regional constant and the intra-regional variables are all statistically significant, thus indicating that the intra-regional flows behave differently from the inter-regional flows. In addition, the mean squared error (MSE) is smaller in the model with intra-regional terms, thus suggesting that this model should be preferred when considering remittance flows between region-pairs. Third, the coefficient estimates for distance are negative and statistically significant. This is consistent with the evidence presented in Figure 6: the smaller the distance between the zones, the larger the remittance flows. Finally, when comparing the SDM estimates with those of the gravity specification and the SLX specification, we find that the signs of the coefficients are identical and their values are similar to those of the SDM estimates.[8]

Please insert Table 4 here

Interpreting the parameter estimates requires implementing our impact decomposition. As the dependent variable in our model specifications is expressed in logarithm, we need to consider equation (21) which requires the computation of the predicted values given in expressions (23) and (24). In addition, we rely on the Bayesian estimates presented in Table 4 to implement formulas (18) and (20). Figure 7 exhibits the local impacts $IE(s)$, $OE(s)$, $DE(s)$ and $NE(s)$ for each macro-economic zone and for the variables we consider in this case study: electric use per capita, population, transfer cost and political stability. Importantly, the computation time using our simplifications is 7 times smaller than the standard way which requires inverting the complete inverse filter matrix $\mathbf{A}(\mathbf{W})$.

Please insert Figure 7 here

The first element to highlight from Figure 7 is that the model without intra-regional terms over-estimates the intra-regional effects of the biggest remittance flows (East-Asia, North America, and European Union before 2004 expansion). It thus confirms that we should prefer the model specification including intra-regional terms. For the reasons above, in what follows, we discuss the impact decomposition of the model specification including the intra-regional terms (right column in Figure 7).

Focusing on the impacts of electricity use per capita (top panel in Figure 7), which is a variable of type odc (characterizing both the origin and the destination countries in the remittance flows) that is expressed in logarithm, we find that, consistent with economic intuition, the macro-economic zones that send more remittances than what they receive (North America, European Union before expansion, other countries in Europe, Pacific) exhibit larger local origin effects $OE(s)$, relative to the intra, destination and network effets ($IE(s)$, $DE(s)$, or $NE(s)$, respectively). On the contrary, the macro-economic zones that receive more funds than what they send (East Asia, South Asia, North Africa or Africa) have higher destination effects $DE(s)$ compared to their corresponding origin, intra or network effects, (respectively, $OE(s)$, $IE(s)$ or $NE(s)$). In particular, North America is the macro-economic zone that is impacted the most by a change of the electricity use per capita. To build intuition

[8] In the same line, we carry out Moran's tests on the residuals of these two models and we could not reject the hypothesis of the absence of spatial autocorrelation. This in turn provides another motivation for the use of the SDM model. Moreover, the calculation of the mean squared errors also suggests that spatial models fit the data better.

on the decomposition of local impacts, we now focus on this macro economic region. Our results indicate that a 1% expansion in the electricity use per capita in North America increases the remittance flows starting from North America by 881 million dollars ($OE(N.America) = 881$). In turn, the intra-regional flows increase by 19.2 millions dollars ($IE(N.America) = 19.2$), while the remittance flows arriving to North America expand by 55.7 million dollars ($DE(N.America) = 55.7$). Finally, our results show that the impacts on all the other remittance flows in the network (not arriving to nor starting from) expand by 19.9 million dollars ($NE(N.America) = 19.9$) when electricity use per capita augments by 1%. Averaging the local impacts as proposed in Table 3 leads to an increase in the outflows starting from North America of 62.9 million dollars on average ($AOE(N.America) = \frac{OE(N.America)}{nd(N.America)-1} = \frac{881}{15-1} = 62.9$) if North America augments by 1% the electricity use per capita. In turn, the average destination effect amounts to 4.3 million dollars ($ADE(N.America) = \frac{DE(N.America)}{no(N.America)-1} = \frac{55.7}{14-1} = 4.3$) and the average network effect, to 0.1 million dollars ($ANE(N.America) = \frac{NE(N.America)}{(no(N.America)-1)(nd(N.America)-1)} = \frac{19.9}{14\times 13} = 0.1$).

It is possible to aggregate the local impacts across regions to obtain the total effects TTE, TDE, TOE, TNE. Table 5 presents the estimated scalar summaries TTE, TDE, TOE, TNE, which we obtain by them by the total number of flows N. The table shows that an increase of 1% in the electricity use per capita in each of the macro-economic regions in the world has an impact of $\frac{TTE}{N} = 18.1$ million dollars on average. In turn, the intra-effect amounts to $\frac{TIE}{N} = 0.6$ million dollars increase, the origin-effect to $\frac{TOE}{N} = 11.1$ million dollars expansion, the destination-effect to $\frac{TDE}{N} = 3.5$ million dollars hike, and the network-effect, to $\frac{TDE}{N} = 2.9$ million dollars. Hence, even if this variable is of type odc, we find that the origin impact is stronger on average than its destination effect.

Please insert Table 5 here

Regarding the estimated impacts for transfer cost, which is a variable of type oc characterizing the origin countries, our results indicate that it has a negative impact for all the macro-economic zones and regardless of the type of impact. Furthermore, the local origin effect $OE(s)$ for a given region s is higher in absolute value than the corresponding $DE(s)$, $IE(s)$ or $NE(s)$ for the same given region. The reason for this is that in our model specification, there is only a β_o term referring to this variable. Importantly, as the variable transfer cost is in USD (not in logarithm), the interpretation is slightly different. For instance, in North America, an additive increase in the transfer cost of 0.12 (which corresponds to 1/100 of the difference between the maximum and the minimum) has an estimated negative impact of 877 million dollars on the remittance flows starting from North America, of 93.7 million dollars on the intra-regional flow, of 8.78 on the flows arriving to North America, and of 39 million dollars on the network not including North America. Furthermore, the total estimated impact exhibited in Table 5 is negative and mainly due to the origin effect.

Finally, concerning the political stability variable, which is of type dc characterizing the destination countries, the local destination effect $DE(s)$ for a given region s is higher than the corresponding $OE(s)$, $IE(s)$ or $NE(s)$. This is because our model specification only includes a β_d term for this variable. Results show that East Asia and European Union are the most impacted regions. In East Asia, for instance, an additive increase of the political stability index of 0.023 (which corresponds to 1/100 of the difference between the maximum and the minimum), has a positive impact of 705 million dollars on the remittance flows arriving to East Asia, of 150 million dollars on the intra-regional flow, of 2 million dollars on the flows leaving from East Asia, and of 17 million dollars in the remaining network. The scalar summary measures reported in Table 5 reveal that the estimated total effect is mainly explained by the destination effect.

7 Conclusions

In this paper, we present local measures for the impact decomposition in a general spatial autoregressive interaction model including endogenous and exogenous interaction effects, and allowing for a different list of origins and destinations, as well as for different origin and destination characteristics. Importantly, we show how to compute the impact measures for all the possible types of flow data (square, rectangular and non-cartesian) and for any characteristic type (type oc if origin only, type dc if destination only or type odc if both). Our approach allows us to considerably reduce the computation time of the vector of impacts of size N (given a change of a characteristic in a particular site) in the square and rectangular cases. This is thanks to some simplifications in the computation formulas. However, our simplifications only work when there are no intra-regional coefficients. From an applied point of view, most geomarketing applications do not include intra-flows. In the case study we propose, we show the advantage of considering local rather than global effects. This is because aggregations (when computing total effects) hinder the identification of heterogeneous behaviors of certain geographical areas (or entities), with respect to the remaining sites (or entities). Although we do not report the significance associated to our local impact measures, inference can be done using the MCMC estimates (see LeSage and Thomas-Agnan, 2015; Laurent et al., 2022b, for instance). We aim at including all these new tools in an upcoming version of the R package spflow (Dargel and Laurent, 2021). Finally, one venue of future research is to extend the impact computations to panel and longitudinal datasets, where the entities are allowed to evolve through time. Moreover, Oshan (2021) questions the veracity of network effects in real applications. From a single cross-section, one cannot observe the impact of a change in one characteristic and, therefore, it may be difficult to prove or disprove their existence outside of a spatio-temporal framework. Nevertheless, this investigation should be an additional interesting direction of future research.

Acknowledgement

The authors are grateful to Lukas Darkel for his helpful comments. Thibault Laurent and Christine Thomas-Agnan acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

References

Chun, Y. and Griffith, D. A. (2011). Modeling network autocorrelation in space–time migration flow data: an eigenvector spatial filtering approach. *Annals of the Association of American Geographers*, 101(3):523–536.

Curry, L. (1972). A spatial analysis of gravity flows. *Regional Studies*, 6(2):131–147.

Curry, L., Griffith, D. A., and Sheppard, E. S. (1975). Those gravity parameters again. *Regional Studies*, 9(3):289–296.

Dargel, L. and Laurent, T. (2021). *spflow: Spatial Econometric Interaction Models*. R package version 0.1.0.

Dargel, L. and Thomas-Agnan, C. (2023). Efficient estimation of spatial econometric interaction models for sparse OD matrices. *TSE Working Paper 23-1409*.

Fischer, M. M. and Griffith, D. A. (2008). Modeling spatial autocorrelation in spatial interaction data: an application to patent citation data in the european union. *Journal of Regional Science*, 48(5):969–989.

Fischer, M. M., Scherngell, T., and Jansenberger, E. (2006). The geography of knowledge spillovers between high-technology firms in europe: Evidence from a spatial interaction modeling perspective. *Geographical Analysis*, 38(3):288–309.

Fotheringham, A. S. (1981). Spatial structure and distance-decay parameters. *Annals of the association of American geographers*, 71(3):425–436.

Fotheringham, A. S. (1983). A new set of spatial-interaction models: the theory of competing destinations. *Environment and Planning A: Economy and Space*, 15(1):15–36.

Goulard, M., Laurent, T., and Thomas-Agnan, C. (2017). About predictions in spatial autoregressive models: Optimal and almost optimal strategies. *Spatial Economic Analysis*, 12(2-3):304–325.

Griffith, D. A. and Jones, K. G. (1980). Explorations into the relationship between spatial structure and spatial interaction. *Environment and Planning A*, 12(2):187–201.

Hunt, L. M., Boots, B., and Kanaroglou, P. S. (2004). Spatial choice modelling: new opportunities to incorporate space into substitution patterns. *Progress in Human Geography*, 28(6):746–766.

Laurent, T., Margaretic, P., and Thomas-Agnan, C. (2022a). Investigating graphical representations of origin-destination spatial flow data. *Mimeo*.

Laurent, T., Margaretic, P., and Thomas-Agnan, C. (2022b). Neighbouring countries and bilateral remittances: a global study. *Spatial Economic Analysis*, 0(0):1–28.

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

LeSage, J. P. and Fischer, M. M. (2016). Spatial regression-based model specifications for exogenous and endogenous spatial interaction. In Patuelli, R. and Arbia, G., editors, *Spatial econometric interaction modelling*, pages 15–36. Springer.

LeSage, J. P. and Pace, R. K. (2006). Interpreting spatial econometric models. In Fischer, M. and Nijkamp, P., editors, *Handbook of regional science*, pages 1535–1552. Springer, Berlin, Heidelberg.

LeSage, J. P. and Pace, R. K. (2008). Spatial econometric modeling of origin-destination flows. *Journal of Regional Science*, 48(5):941–967.

LeSage, J. P. and Thomas-Agnan, C. (2015). Interpreting spatial econometric origin-destination flow models. *Journal of Regional Science*, 55(2):188–208.

Oshan, T. M. (2021). The spatial structure debate in spatial interaction modeling: 50 years on. *Progress in Human Geography*, 45(5):925–950.

Rowe, F. L. and Dennett, R. (2022). A., 2022. spatial interaction modelling: A manifesto. *Wolf, L.*

Tiefelsdorf, M. (2003). Misspecifications in interaction model distance decay relations: A spatial structure effect. *Journal of Geographical Systems*, 5(1):25–50.

Tiefelsdorf, M. and Griffith, D. A. (2007). Semiparametric filtering of spatial autocorrelation: the eigenvector approach. *Environment and Planning A*, 39(5):1193–1221.

8 Appendix

8.1 Tables

Table 1: Scalar measures summary as cumulative effect and average effect as presented by LeSage and Thomas-Agnan (2015).

	Intra	**Origin**	**Destination**	**Network**	**Total**
Average	$AIE = \frac{TIE}{N}$	$AOE = \frac{TOE}{N}$	$ADE = \frac{TDE}{N}$	$ANE = \frac{TNE}{N}$	$ATE = \frac{TTE}{N}$

Table 2: Number of terms in $IE(s)$, $OE(s)$, $DE(s)$ and $NE(s)$ with respect to s, for our previous example in Figure 1 (b).

Z is of type oc, i.e. $S = O$

				Number of elements				
Nature	Sites	$nd(s)$	$no(s)$	**Intra**	**Origin**	**Destination**	**Network**	**Total**
$s \in O \setminus D$	o_1, o_2, o_3, o_4	3	0	0	3	0	12	15
$s \in O \cap D$	od_1	3	5	1	2	4	8	15
Total	O	15	5	1	14	4	56	75

Z is of type dc, i.e. $S = D$

				Number of elements				
Nature	Sites	$nd(s)$	$no(s)$	**Intra**	**Origin**	**Destination**	**Network**	**Total**
$s \in D \setminus O$	d_1, d_2	0	5	0	0	5	10	15
$s \in O \cap D$	od_1	3	5	1	2	4	8	15
Total	D	3	15	1	2	14	28	45

Z is of type odc, i.e. $S = O \cup D$

				Number of elements				
Nature	Sites	$nd(s)$	$no(s)$	**Intra**	**Origin**	**Destination**	**Network**	**Total**
$s \in O \setminus D$	o_1, o_2, o_3, o_4	3	0	0	3	0	12	15
$s \in D \setminus O$	d_1, d_2	0	5	0	0	5	10	15
$s \in O \cap D$	od_1	3	5	1	2	4	8	15
Total	$O \cup D$	15	15	1	14	14	76	105

Table 3: Local average effects $AIE(s)$, $AOE(s)$, $ADE(s)$, $ANE(s)$, $ATE(s)$.

Configuration	**Intra**	**Origin**	**Destination**	**Network**	**Total**
$s \in O \setminus D$	.	$\frac{OE(s)}{nd(s)}$	.	$\frac{NE(s)}{nd(s)(no(s)-1)}$	$\frac{TE(s)}{N}$
$s \in D \setminus O$	.	.	$\frac{DE(s)}{no(s)}$	$\frac{NE(s)}{no(s)(nd(s)-1)}$	$\frac{TE(s)}{N}$
$s \in O \cap D$	$IE(s)$	$\frac{OE(s)}{nd(s)-1}$	$\frac{DE(s)}{no(s)-1}$	$\frac{NE(s)}{(no(s)-1)(nd(s)-1)}$	$\frac{TE(s)}{N}$

Table 4: Gravity model (OLM), SLX and Bayesian SDM estimates

	Gravity model (OLM) log(remittances)		SLX log(remittances)				SDM without intra log(remittances)				SDM with intra log(remittances)				
$\hat{\rho}_o$	.		.				0.273 (4.13)				0.294 (4.69)				
$\hat{\rho}_d$	.		.				0.108 (1.40)				0.112 (1.51)				
Constant	-12.078 (3.80)		-5.544 (5.59)				-10.499 (-1.84)				-17.591 (-2.99)				
Constant (intra)	.		.				.				30.617 (3.69)				
	$\hat{\beta}_o$	$\hat{\beta}_d$	$\hat{\beta}_o$	$\hat{\beta}_d$	$\hat{\delta}_o$	$\hat{\delta}_d$	$\hat{\beta}_o$	$\hat{\beta}_d$	$\hat{\delta}_o$	$\hat{\delta}_d$	$\hat{\beta}_o$	$\hat{\beta}_d$	$\hat{\beta}_i$	$\hat{\delta}_o$	$\hat{\delta}_d$
Log(Electric use PC)	1.858	0.266	1.860	0.183	-0.116	-0.333	1.386	0.252	-0.204	-0.133	1.454	0.399	-1.843	-0.199	-0.089
	(0.11)	(0.14)	(0.11)	(0.14)	(0.23)	(0.21)	(8.47)	(1.72)	(−0.85)	(−0.62)	(9.18)	(2.79)	(−4.65)	(−0.85)	(−0.45)
Log(Population)	0.524	0.886	0.519	0.846	.	.	0.401	0.801	.	.	0.432	0.856	-0.672	.	.
	(0.084)	(0.086)	(0.085)	(0.089)	.	.	(4.32)	(7.42)	.	.	(4.71)	(8.03)	(−2.05)	.	.
Transfer cost	-0.181	.	-0.185	.	.	.	-0.135	.	.	.	-0.123	.	.	.	
	(0.03)	.	(0.032)	.	.	.	(-3.91)	.	.	.	(-3.60)	.	.	.	
Political stability	.	0.551	.	0.685	.	.	.	0.514	.	.	.	0.459	.	.	.
	.	(0.20)	.	(0.22)	.	.	.	(2.28)	.	.	.	(2.15)	.	.	.
Distance	-1.908 (0.15)		-1.957 (0.15)				-1.386 (-6.40)				-1.205 (-5.08)				
Mean squared error	1.865		1.839				1.742				1.590				
Observations	210		210				210				210				

Electric use PC stands for electricity use per capita. t-stat are reported into the parentheses.

Table 5: Scalar measures summary as cumulative effect and average effect as presented

	SDM with intra				
	$\frac{TIE}{N}$	$\frac{TOE}{N}$	$\frac{TDE}{N}$	$\frac{TNE}{N}$	$\frac{TTE}{N}$
Log(Electric Use PC)	0.6	11.1	3.5	2.9	18.1
Log(Population)	2	3.5	8.2	3.4	17.1
Transfer Cost	-3.2	-11.2	-0.7	-5.7	-20.8
Political Stability	2.8	0.2	10	2.1	15.2

8.2 Figures

Figure 1: Toy examples used to illustrate: a) the square case where $O = D = S$, b) the rectangular case where O and D are different with a possible non-null intersection, and c) the non-cartesian case, where some of the possible flows between the origins and the destinations are not observed or have no meaning

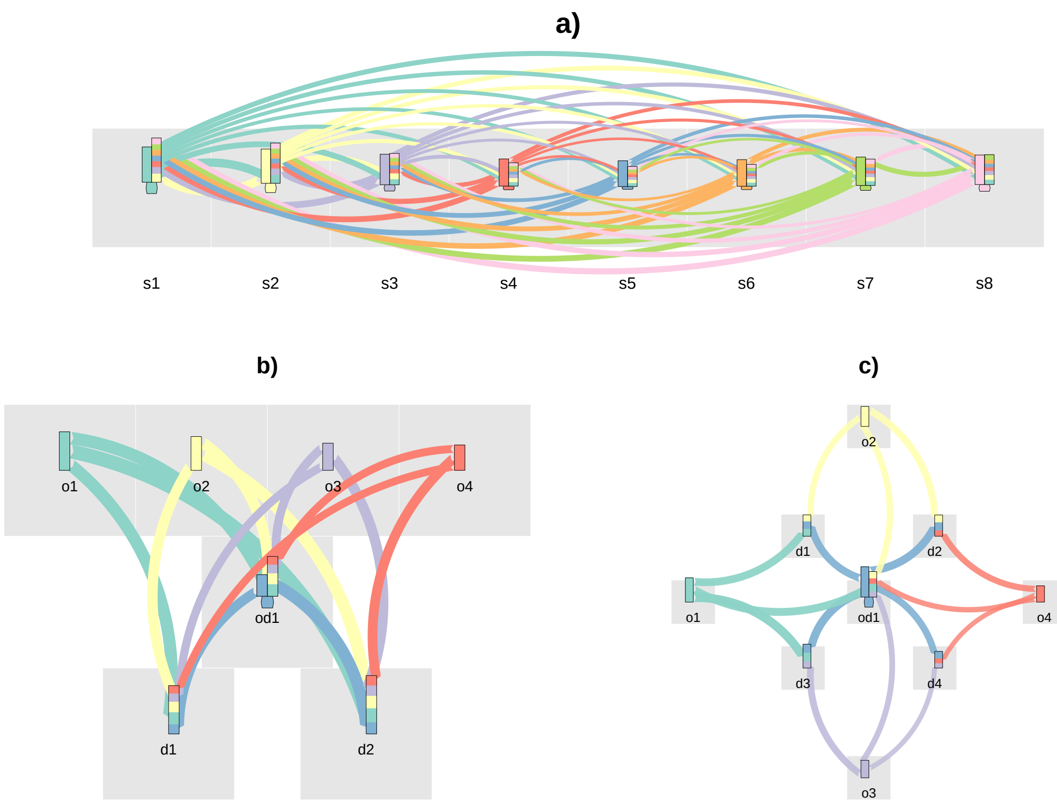

Figure 2: Representing the local effects in the square case, by stacking the $IE(s)$, $OE(s)$, $DE(s)$ and $NE(s)$. The left barplot presents the absolute values of the local impacts; the shares' graphic, on the right, exhibits the relative values.

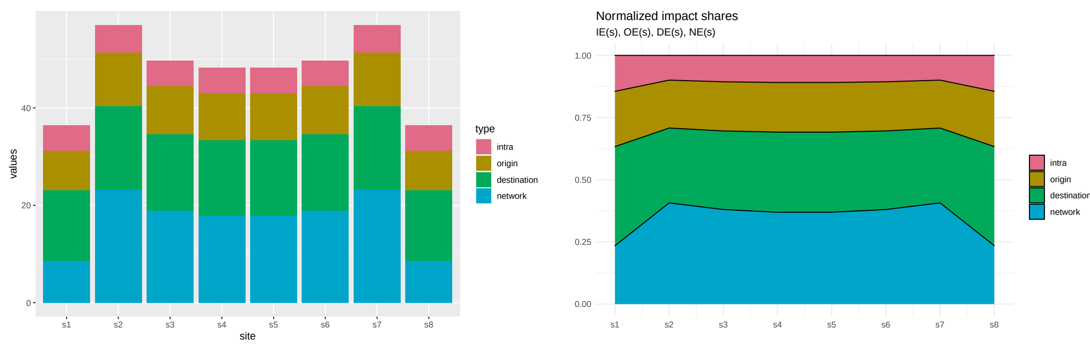

Figure 3: $IE(s)$ and kernel densities for the $OE(s)$, $DE(s)$, $NE(s)$ and $TE(s)$ in the square case

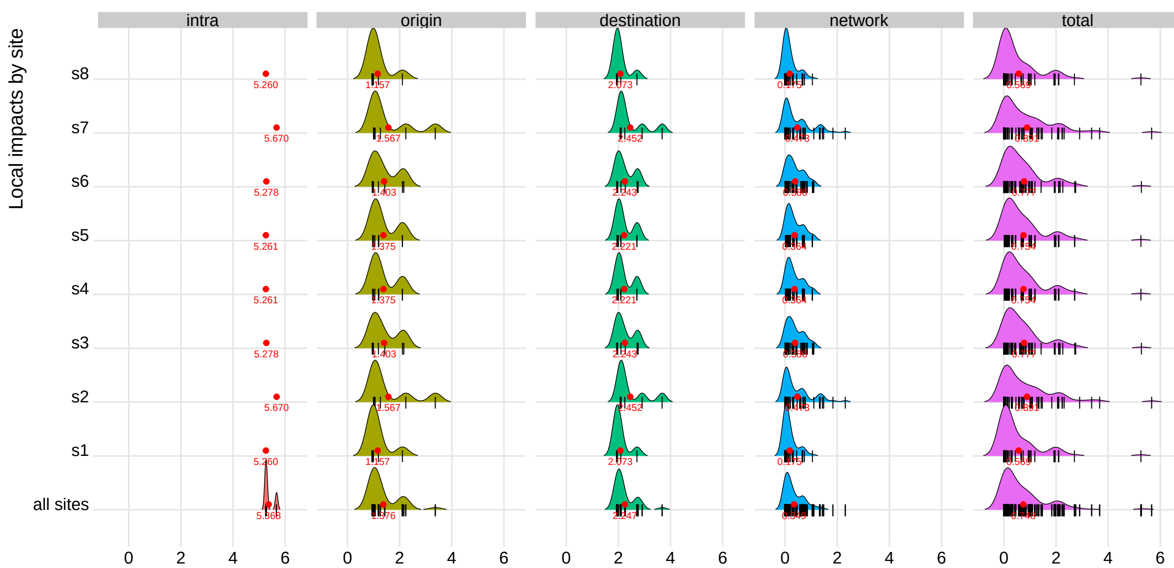

Figure 4: Example of representation of the local effects. In the fictitious example of Figure 1 b), o_1, o_2, o_3, o_4 are origin sites only, d_1, d_2 are destination sites only and od_1 is an origin/destination site. The sites are ordered with respect to their nature.

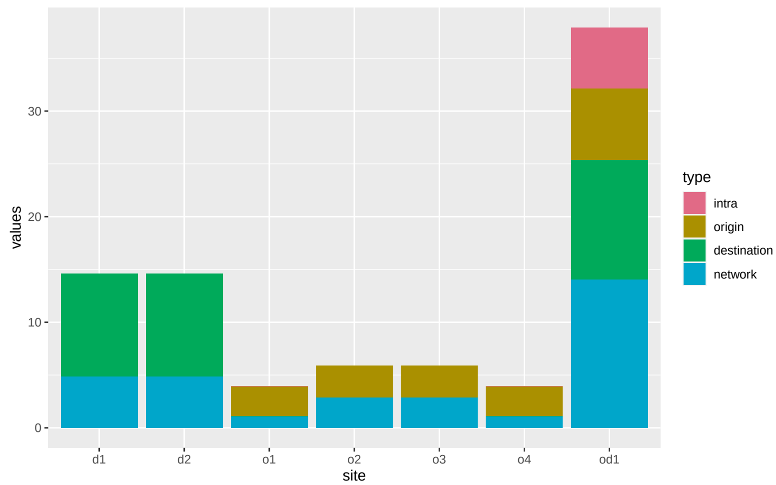

Figure 5: $IE(s)$ and kernel densities for the $OE(s)$, $DE(s)$, $NE(s)$ and $TE(s)$

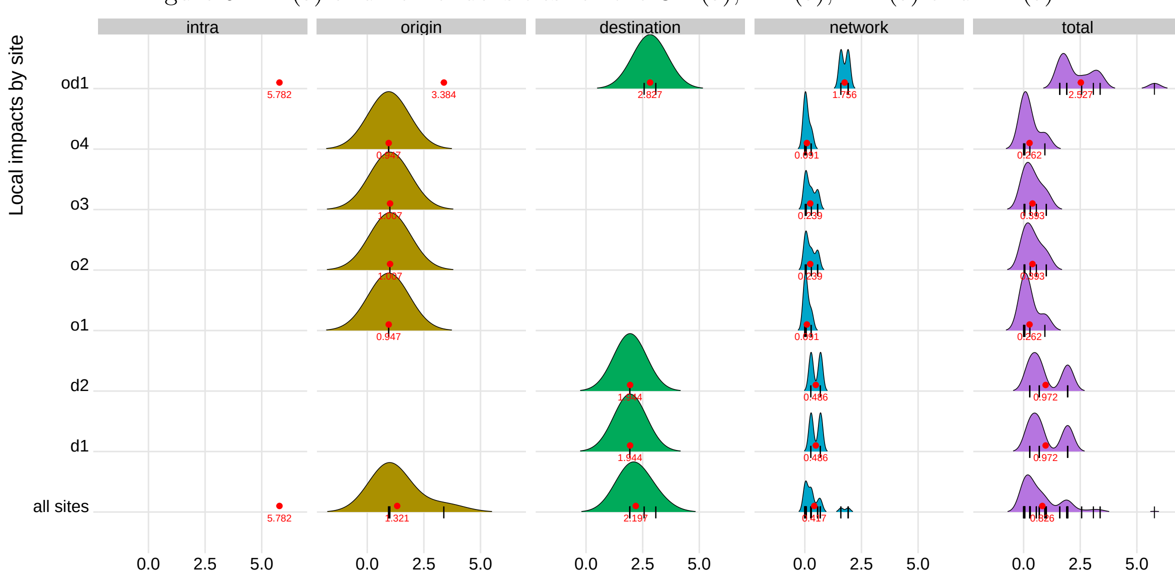

Figure 6: Map flow for the remittances between macro-economic regions

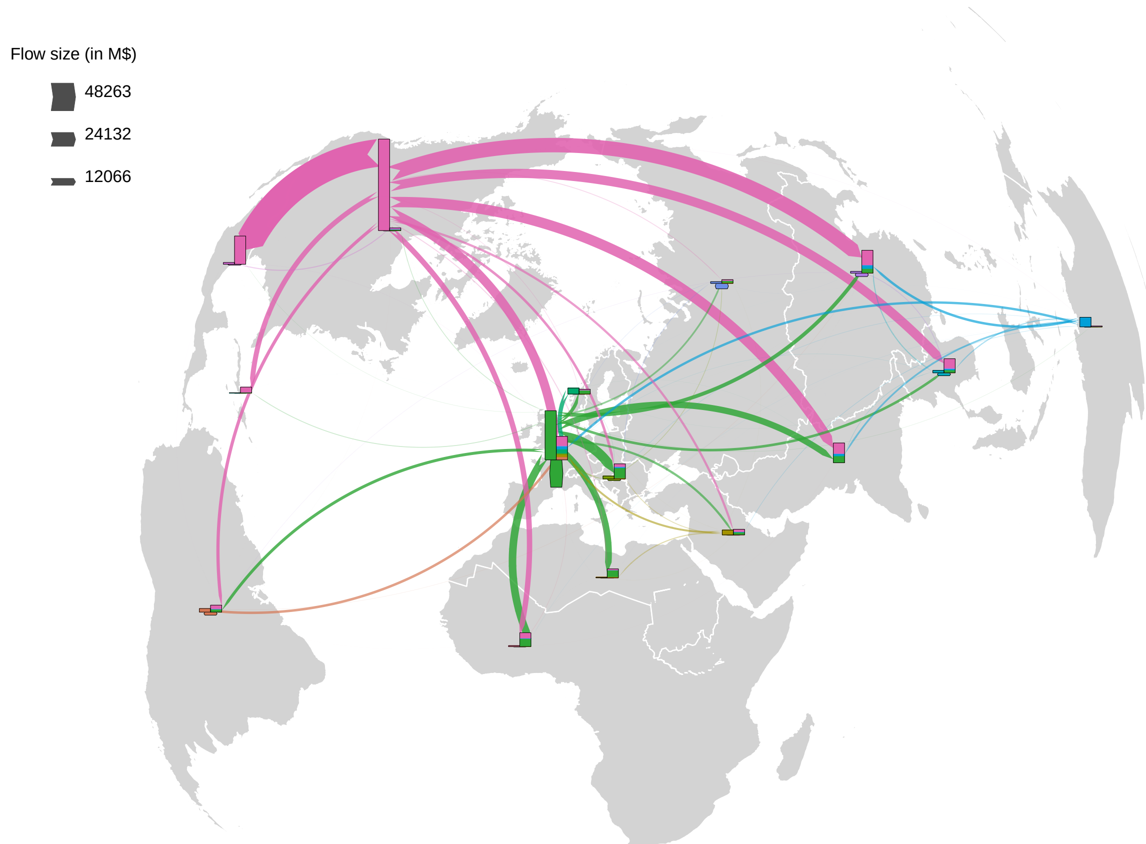

Figure 7: Representation of the local impacts, region by region, and variable by variable (in million dollars)