Gabriela **ANTOŠOVÁ** • Helmuth Yesid **ARIAS GÓMEZ**

SPATIAL BEHAVIOR OF MANUFACTURING IN COLOMBIA: TERRITORIAL CASE OF CONNECTION IDEAS OF MARSHALLIAN AND NEW ECONOMIC GEOGRAPHY INSIGHTS ABOUT INDUSTRIAL LOCALISATION

Gabriela **Antošová** (ORCID: 0000-0001-5330-679X) – *Akademia Humanitas/Humanitas University, Poland*
Helmuth Yesid **Arias Gómez** (ORCID: 0000-0003-0107-8611) – *Czech Technical University in Prague (Masaryk Institute of Advanced Studies), Czech Republic*

Correspondence address:
Kilińskiego Street 43, 41-200 Sosnowiec, Poland
e-mail: gabriela.antosova@humanitas.edu.pl

ABSTRACT: This study explores the spatial distribution of manufacturing in Colombia through the lenses of Marshallian and New Economic Geography theories on industrial localisation. Employing spatial econometric methods with geo-referenced data reveals that industrial presence at the municipal level is positively impacted by the service sector connections and population density. Additionally, a panel data analysis suggests that traditional manufacturing hubs benefit from unobservable fixed effects, enhancing the persistence of initial locational advantages. These findings underscore the ongoing spatial concentration of manufacturing, influenced by both historical industrial activities and contemporary urban factors.

KEYWORDS: industrial localisation, industrial agglomeration, new economic geography, manufacturing sector

 DOI: 10.34659/eis.2024.91.4.838

Introduction

This article pursues to answer some questions arising from the uneven regional industrial development, stating the hypothesis that in advanced regions, large populations and urban activities reinforce the original advantages bestowed by early industrialization. The early manufacturing localisation eased the irruption of capital funding productive activities and propitiated the emergence of entrepreneurial skills (Hallen, 2008).

Colombian topography purported a high transport cost scheme that has been a strong feature of the economy. Long distances from inland areas to coastal ports hindered export development (Melo, 1978). This regional system evolved around urban places and rural markets exerted an important role as a source of regional demand, mainly for consumption goods that were characteristic of these stages of industrialization (Moncayo, 2002; Moncayo, 2007).

The historical path came about over two centuries and really influenced the current spatial distribution of manufacturing. Therefore, our interest is to confirm if the early manufacturing tradition of some departments (Colombian administrative regions) made up some unobserved traits explaining the actual location of manufacturing and its current trends. Several time-invariant factors, can be assumed as unobserved features deserving deep analysis and interpretation. With this aim, we built a panel data of Colombian departments using the most relevant up-to-date variables associated with the industrial data, extracted from the Colombian manufacturing survey.

Another point of interest is the identification of Marshallian forces explaining the agglomeration of manufacturing plants. This approach has an explicit spatial concern and the strategy includes the estimation of econometric models applying the spatial econometrics with data arranged at municipal level and the panel data technique with regional databases. The detailed information at a municipal level could be consulted provided the available information coming from the national census conducted during 2005 (The source for the data panel was the Annual Manufacturing Survey conducted by the Colombian Statistical Office DANE), and for spatial econometrics, geo-referenced data were extracted by Census 2005 which is available in the SHP format at the Colombian Geographical Office (IGAC). The industrial data at a regional level from the Annual Manufacturing Survey were used for the panel data model.

The article has a synthetic structure. The first task is to expose theoretical bases, the second section develops the empirical strategy based on the techniques of spatial econometrics and panel data. In the last section, we deepen the interpretation and propose some conclusions.

Concepts and Theoretical Background

The manufacturing in Colombia shows a spatial clustered distribution with strong presence of plants in the larger urban cores, a fact explained by traditional theoretical reasons such as the size of the market and the exploitation of proximity to other firms. From the approach of New Economic Geography, the theoretical explanation relies on transport costs, economies of scale and the share of manufacturing labour in the number of employees (Krugman, 1979; Krugman, 1992; Krugman, 2008). Adequately interpreted, such parameters can explain accurately the Colombian manufacturing development during the 19th and 20th centuries.

Some random events give rise to an accumulative causation that triggers manufacturing development. Meanwhile, the parameters must attract footloose factors and the plants into the original cores, relegating the farmer's localisation and the agricultural production to peripheral zones.

As figure 1 shows, the Colombian cordillera system crosses the whole country and divides regions and territories. Main cities are important hubs of regional activities and concentrate manufacturing development. Bogotá came up in the highness of the inner mountains as a political centre and more populated city, in a location distant from both coasts but with a strong influence on the central area of the territory. Medellín emerged as a regional centre in a region highly endowed with gold and silver mining and prolific in manufacturing activities, essentially in the productive chain of textiles. Barranquilla and Cartagena flourished as coastal cities and consequently, as important international ports for trading finished manufactured products and inputs. Cali in the west, experienced a late upsurge during the 20th century based on an agricultural advantage and, the proximity to the port of Buenaventura, in the Pacific Sea.

DOI: 10.34659/eis.2024.91.4.838

NOTE: Information refers to unities of small scale.

Figure 1. Colombia: Number of Industrial Unities by Municipalities (2005)

Source: Own elaboration based on DANE and IGAC.

From a historical perspective, since the early colonial age, economic localisation was associated with populated indigenous settlements and actually was constrained into isolated regional markets bounded geographically (Haveman & Nonnemaker, 2000). In this context, the space burst onto the scene, conveying transport costs as the key localisation factor. Afterwards, when industrialisation evolved during the age of Substitutive Industrialization during the 20th century, the localisation of firms relied on the transformation of natural resources and the proximity to cities. Finally, as a consequence of economic liberalisation at the end of the 20th century, some barriers were dropped, and the transport costs diminished. Nevertheless, the advantages of the urban production cores were enhanced as a consequence of historical inertia or by the forces of "cumulative causation".

In this context, Meyer (1983) identifies the conformation of some subsystems operating around urban centres with the economic function of providing industrial goods to vast hinterlands. Such processes are analysed by Krugman (1992; 2008) and Davis and Weinstein (1996). Therefore, the "Home Market Effect" emerged as a key concept that lies beneath the strengthening of economies of scale, being applied to firms under monopolistic competition. In the case of Colombian development, this effect spurred the internal scale economies and forged productive systems gravitating around large urban centres. Afterwards, when such systems evolved, and communication was possible, trade with other cities and regions ended up in multiregional, or in some cases multinational systems. In such terms, the export of goods originated in a local process of economies of scale, but the final outcome is the export of goods for which there was an initial strong local demand (Krugman, 1980).

The most privileged place for propitiating economies of scale is precisely the city and its function as the pivot of the regional economic system. According to Camagni (2005), the criteria for urban recognition is the sudden change in the density of land amid broad open extensions of hinterlands, and the explanation for such spatial behaviour is the social convenience of closer human relationships. For the sake of labour mobility, the footloose workers are attracted to the city reinforcing the accumulative causation that spurs the concentration of production around the core of the system

(Krugman, 1992; Krugman, 2008). The Marshallian spillovers or technological osmosis constitute a powerful force derived from the clustering of firms. In his own words, "mysteries" that are implicit in a firm's technical innovations can become social knowledge shared with several enterprises located closely (Whittington et al., 2009). In effect, the agglomeration of firms propitiates an "industrial atmosphere" in specialised areas that brings benefits to firms belonging to the same manufacturing sector (Capello, 1999; Georgallis et al., 2019).

The primary explanation for localisation relies on geographical endowments. Once such basic natural factors were identified, Marshall proposed three sources for spatial concentration of firms, sources that were refined and revisited by several authors of New Economic Geography (Krugman, 1991; Krugman, 1992; Fujita & Krugman, 2004; Ottaviano & Puga, 1997). In fact, Krugman (1992) renames this kind of clustering as an agglomeration propitiated by access to resources and asserts that clustering phenomena are really extended and cover a wide range of economic sectors, including other activities different to technological tasks. Krugman ascribes agglomeration to Pecuniary Spillovers, defined as inter-firm interaction mediated by the market (Broschak, 2004). In his opinion, simple technological spillovers are difficult to model due to their lack of visibility, and as a consequence, they are excluded from the analysis (Krugman, 1992).

Finally, a pooled labour market creates an additional source of agglomeration because firms can find specialised workers with specific skills, and the workers have more opportunities to be hired when firms agglomerate nearby. Krugman explains that such circumstances reduce the risk for enterprises and workers because they offer an environment of high certainty (Krugman, 1992). Marshall (2005) describes that isolated entrepreneurs could find high quantities of people willing to work but do not easily find specialised employees for their activities, recognising that firms will prefer moving to places where the labour market is denser and diverse. In the Bond-Smith (2024) analysis, we recognize the combination of Marshall and Krugman's categories embedded in growth theory. It leads to the application of non-rival knowledge and increasing returns models. However, in some particular cases, the models are specified describing agglomeration economies of innovation (Katila & Chen, 2008); however, those models work even when no scale effect is considered inside.

Krugman (2008) exposes three basic parameters that influence the agglomeration. Producers have the objective of meeting the requirements of two markets, and then they could settle in both places or concentrate production in one region. In those terms, the sign S^* represents the shipments to both regions. For transportation of commodities, they incur transport costs per unit (τ), then one way to eliminate transport costs is forging a symmetric equilibrium and distributing production into both regions, but a second plant requires a fixed cost (F). In order to minimise transport costs, a new plant will be located in the larger market, and the condition to generate a concentration process will be: $F>\tau S^*$. If expressed in terms of a fixed cost per unit $F/S>\tau$, then the larger market will maintain all industries if scale economies (F/S^*) are more powerful than transport costs.

The inclusion of labour mobility enables the model to deal with the concentration of productive factors. In doing so, a parameter μ is defined, representing the share of footloose workers in production. It opens the possibility to concentrate industrial production in a manufacturing core, while in peripheral areas, the immobile factor will generate a demand represented by $S(1-\mu)/2$. Summing up, when the model includes transport costs and economies of scale and introduces a mechanism to insert the footloose factors, the location of firms becomes endogenous. In this context, the economy can propitiate equilibria characterised by the concentration or dispersion of firms according to the behaviour of parameters. When production concentrates on space, a core-periphery scheme is propitiated under the condition: $F/S > \tau\ (1-\mu)/2$.

Revisiting the recent Marshall's interpretation, two quantitative approaches to explain space have emerged. The first stream aims to quantitatively describe the underlying economic mechanisms behind localisation, starting from spatial data and panel data. From the beginning, this kind of approach defines a representative spatial unit and estimates covariate values in order to explain the economic phenomena around such defined spatial unit (Arbia et al., 2009). Capoani (2023) applies a Smithian interpretation of the spatial concentration of European activities. This gravitational approach points on the map the "blue banana" as the most developed continental area, defined where the vector sum of the individual is the maximum, the product of the masses is the maximum, and the distances are minimised.

DOI: 10.34659/eis.2024.91.4.838

Research methods – empirical strategy

We will apply some econometric techniques in order to figure out the determinants of the location of Colombian manufacturing firms using the log of the number of industrial units by municipalities as endogenous. Secondly, an important aspect in the deep unequal spatial distribution of production is the seeming existence of unobserved time-invariant effects, acting in favour of more traditional manufacturing centres and assuming a sort of early aptitude for entrepreneurship, a solidly rooted manufacturing tradition and the geographical distance regarding the economic cores.

For the first purpose, we estimate the lagged spatial model and the spatial error model. In addition, the estimation of the spatial Durbin model contributes to identifying the indirect effect coming from spatial lags.

The data source is a geo-referenced database elaborated with the basic information of the 2005 General Census conducted by the Statistical Office (DANE), which is available in geospatial terms on the web of Geographical Office (IGAC). The information is very detailed, and the geographical scope is very specific because it comprises the entire local, territorial structure of the municipalities.

In the second phase, we conduct an exercise of panel data applied to Colombian departments (the administrative regional breakdown). The source of information is the Annual Manufacturing Survey (EAM) conducted by DANE. Regarding panel data, municipal coverage of the analysis is ruled out due to the confidentiality of the information and the geographical units used, which are the regions (departamentos).

The econometric models

Spatial Econometrics

In the context of employment growth, Helsen (2008) applies a Kaldorian analysis corroborating the impulse that industrial growth transmits to overall activity, including data from 50 states in the USA. Another application of spatial econometrics in the manufacturing context appears in Bernat (1996), who connects industrial growth with the overall growth of the economy. There, the author attributes the poor performance of former models to the presence of spatial autocorrelation and the lack of procedures to correct it. Several authors analysed the influence of natural conditions and endowments on specialisation. Arias-Gomez and Antošová (2023) blended a classical interpretation based on Ricardian principles with a disruptive approach reliant on Krugman's new trade theory. In doing so, the manufacturing profiles of Czech regions are understood as moulded by natural determinants and economies of scale.

Bernat runs regressions using OLS procedures, and later, in three separate models, he submits data to standard correction of spatial autocorrelation, either by a lag model or by a spatial error model. Applying three different contrasts, he finds strong evidence about the presence of spatial autocorrelation, and once the correction is applied, his results contribute to a better fit of the model. Other econometric exercises are run in the context of Marshall's agglomeration forces as they appear in Lu and Tao (2008), Rosenthal and Strange (2001) and Fu and Hong (2010).

In other fields, spatial econometrics has been applied to figure out the behaviour of house prices in the frame of real state analysis (Osland, 2010). There also appears a short description of previous works dealing with spatial behavior and empirical strategies.

In order to give a local approach to the analysis, we run a spatial model using the same manufacturing tradition. The set of data reports information on 1118 municipalities, including all municipalities and the capital city Bogotá; therefore, there are profusely crammed manufacturing centres and small towns as well. The sources of data are DANE (Departamento Administrativo Nacional de Estadística) and IGAC (Instituto Geográfico Agustín Codazzi), and the objectives of the analysis are enriched by an approach at a local level. Appropriate procedures of spatial econometrics are conducted to deal with spatial interaction and spatial structure of data, applying standard procedures following the pioneer text of Anselin (2001) and the recent recommendation of Elhorst (2010).

The census data of 2005 provides information at a municipal level that contributes hugely to the analysis in spatial terms, and we gain an enormous advantage in coverage. At such a level of a geo-

DOI: 10.34659/eis.2024.91.4.838

graphical scale, it´s possible to know the spatial dynamics of a firm´s distribution, combined with a theoretical explanation of economic localisation. Some proxies are used following specific literature in order to represent such an agglomerative dynamic, particularly for industrial locations. For our purposes, census data offer two proxies representing linkages of industry with different economic sectors, namely the number of sector services unities set out each municipality , and the local agricultural production. (Sources of data are Ministerio de Agricultura y Desarrollo Rural, DANE and IGAC. DANE collected several variables during the 2005 general Census)

To incorporate the size of local markets, we used the population as a proxy of demand linkages that contribute to attracting firms to specific municipalities. Within this context, there is empirical evidence demonstrating that larger cities have an important advantage in productivity and, more specifically, diversified metropoles have a great deal of economic activity. Elhorst (2010) proposes to include additional spatial interactions in order to figure out the real process of spatial interdependence between endogenous, exogenous and exogenous variables omitted in the model. Other exogenous variables are the number of units in the service sector, the total local population, the existence of technical or university labour, and an additional human capital variable related to educational enrollment. Estimations try to model the effects of market size, the complementarity with other economic sectors and the human capital on the stock of manufacturing firms in each municipality.

Next, we run the spatial lag and spatial error model to see whether autocorrelation really existed. Prior to estimation, standard contrasts of Lagrange Multiplicator and Robust LM are tested in order to discern which specification of spatial dependence turns out adequate.

The phenomena of autocorrelation come up when the residual terms of different geographical areas are correlated. In such cases, data must be submitted for correction. This phenomenon can respond to the presence of systemic variables correlating in space, or a phenomenon of spatial dependence in residuals (Moreno & Vayá, 2002) and eventually, the assumption about uncorrelated error terms is violated (Osland, 2010). Spatial autocorrelation is derived from the omission of some variables in a specified model, and consequently, the error term collects this kind of spatial influence. This phenomenon can come about by the existence of dependence between endogenous variables across spatial units, independent of the influence of exogenous variables, or in the case when the political-administrative partition of data has no economic significance and fails to collect the economic interaction (Helsen, 2008). This last drawback is acknowledged by Krugman (1992) and described by Duranton and Overman (2005) because the specific breakdown of spatial units according to administrative parameters has no economic relevance. Arbia et al. (2009) define this drawback in terms of a Modifiable Areal Unit Problem.

Two models can be proposed to correct spatial autocorrelation. The first model can include a spatial lagged dependent variable, known as the spatial lag of endogenous variable (Wy). A second model assumes a structure of dependence in error term ($E[\varepsilon i \varepsilon j] \neq 0$). In the first case, it is assumed that spatial dependence is linked to the spatial model and is applied when the interest is to bear out the intensity of spatial correlation. In the second case, spatial dependence is understood as nuisance dependence, and it´s useful for correcting the biasing influence of spatial autocorrelation (Anselin, 2001). Moreover, correction for spatial autocorrelation conveys an additional advantage represented by an improvement in the model fit (Bernat, 1996).

In the first model of substantial spatial autocorrelation, spatial lag must be assumed as an endogenous variable, and a suitable method must be applied, considering that the OLS model offers biased and inconsistent estimators (Anselin, 2001). In the second model, when spatial interaction is identified as a spatial error type, estimators are inefficient and, in consequence, statistical inference is invalid, although it could be unbiased (Helsen, 2008). On the other hand, in the case of substantive spatial autocorrelation, estimations will be biased and inconsistent even if the error term is not correlated (Moreno & Vayá, 2002).

Dubin (1998) insists on a careful treatment when spatial autocorrelation is detected, given that the presence of spatial autocorrelation in the context of the regression model has important implications. This situation is almost ubiquitous when data are distributed in space or when location is a fundamental criterion. In this case, the classical OLS regressors turn out to be unbiased but inefficient estimators, and the variance of estimators is biased as well (Dubin, 1998).

DOI: 10.34659/eis.2024.91.4.838

OLS estimation does not incorporate spatial effects, and consequently, it has no big relevance, although it is used as a benchmark for comparison with the remaining models. Such a non-spatial linear regression model is defined as:

$$y = \alpha_{IN} + X\beta + \varepsilon \tag{1}$$

being y a n x 1 vector of observations of dependent variable associated with each spatial unit (k=1,..., n), corresponds to a n x 1 vector of ones associated with constant parameters α, x a n x K matrix representing exogenous variables with the set of associated parameters β and ε_i as an independent and identically distributed error term for all i, with zero mean and variance σ^2.

On previous consideration, any correction to autocorrelation contributes to obtaining more accurate estimates and to improving the reliability of the hypothesis test. When the structure of autocorrelation is estimated, its information is included in the prediction in order to improve accuracy. In order to do this, Maximum Likelihood techniques are commonly applied to model the autocorrelation parameters and estimations of regression (Dubin, 1998).

The presence of spatial autocorrelation in the regression model can be explained in two ways: it's possible that exogenous or endogenous variables be correlated spatially, or that error term has an autocorrelation scheme (Moreno & Vayá, 2002).

In the first case it´s necessary to specify the following model:

$$y = \rho Wy + X\beta + u$$
$$u \approx N(0, \sigma^2 I) \tag{2}$$

where:
y – is a vector (nx1),
Wy – is spatial lag of exogenous variable,
X – is the exogenous variables matrix,
u – is a white noise error term,
N – is the number of observations and ρ is an autoregressive parameter.

In equation (1), it can be observed that if spatial lag is omitted in exogenous or endogenous variables, the feature of spatial dependence will be transmitted to the error term, which is quoted as substantive spatial autocorrelation. In facing this situation, it is necessary to include the spatial lag of variables with spatial autocorrelation.

This model are also known as *Models of Communication or Contagion* and integrates the whole autocorrelation structure into spatial lag as an explicative argument of the endogenous variable. In the case of omission of weight matrix in this type of model, estimation commits a specification error that biases estimators and drives to invalid inference (Moreno & Vayá, 2002).

In the second case, if spatial autocorrelation is present only in the error term, the model to estimate corresponds to:

$$y = X\beta + \varepsilon$$
$$\varepsilon = \lambda\, W\, \varepsilon + U \tag{3}$$

where:
u – is a white noise term,
λ – becomes the autoregressive parameter.

It means that we incorporate an autoregressive process in the error terms and consider that ξ is related not only to a stochastic term of error, but it is also a function of non-included exogenous variables of neighbouring places and therefore, there are several omitted variables spatially correlated to one another.

In equation 3, if there is no omission of lag in variables of the model, it corresponds to the case of residual spatial autocorrelation, and then a scheme of spatial dependence in error term must be included (Moreno & Vayá, 2002). In performing the spatial error model, we assume that there is an autoregressive process in the error terms, and we suppose that any kind of spillover effect is present on the residuals.

DOI: 10.34659/eis.2024.91.4.838

The standard treatment of econometrics in the presence of spatial autocorrelation runs a regression via Maximum Likelihood. As it is known in this technique, regressors are calculated, maximising the logarithm of the function of likelihood associated with the spatial model (Moreno & Vayá, 2002).

We apply an additional perspective. Elhorst (2010) evokes LeSage's suggestion to apply innovative models with an endogenous spatially lagged variable. In addition, the structure of lags in all exogenous variables under the assumption that the behaviour of a neighbor affects the model under estimation globally. In this last case, we estimate direct and indirect effects rendered in the respective coefficient of the contemporaneous spatial effect, and the lagged influence of exogenous variables belonging to neighbour spaces. Such a procedure is widely labelled as the Spatial Durbin Model and its specification corresponds to following model:

$$y = \rho X\beta + \alpha t_N + X\beta + WX\Theta + \varepsilon, \quad (4)$$

where:
$\varepsilon = (\varepsilon_i, \ldots, \varepsilon N)$ – is a vector of disturbances, in which εi are error terms distributed independently and identically with 0 mean and variance σ^2.

Panel Data

During the analysis, we were concerned about some fixed effects of time-invariant variables that are embodied in some favourable conditions for advanced manufacturing regions. We look for the existence of a fixed and random effect using the technique of panel data, and our purpose is to arrange data on a regional scale (Colombian Departments) covering recent years (2011-2015). The regional breakdown of data is extracted from the EAM (The Colombian Manufacturing Survey, compiled yearly) conducted by DANE (Colombian Statistical Office) and collecting standard variables reported by manufacturing units.

Our purpose was to estimate a fixed effect model maintaining a similar structure to our proposed model of spatial econometrics while using the panel data technique, using manufacturing number of employees as an endogenous variable as a proxy of the industrial presence in each department.

We alternated diverse manufacturing endogenous variables, calibrating which one casts the best results. The final model of the panel data ended up using the number of employees as an endogenous variable and as exogenous variables, we included the regional production of the financial sector, the net enrollment in the educational system and the regional mining production. All exogenous variables were demonstrated to be highly significant in all models.

We proved several variables to be endogenous, and the most optimal results were derived from models of manufacturing employment. The data panel specification is as follows:

$$y_{it} = X'_{it}\beta + (\alpha_i + \varepsilon_{it}), \quad (5)$$

where:
α_{it} – Collects the unobservable factors that do not change on time,
ε_{it} – is the idiosyncratic error collecting all unobservable factors but that can change in time.

Having the data arranged as a panel data, we need to figure out the optimal procedure for estimating the fixed effects. In doing so, we will determine whether the most suitable estimation must be done by Fixed or Random Effects.

In the first case, we assume that ε_{it} can be correlated with X_{it}, so this regressor can be endogenous. In such situation, OLS estimations of β are inconsistent but the estimators regressed by fixed effects are consistent.

In the case of Random Effects, we assume that α_i is a randomly generated process unrelated to X_{it}; consequently, such regressor is exogenous and all estimations render consistent parameters.

DOI: 10.34659/eis.2024.91.4.838

Data and results

Spatial Econometrics

In this context, Table 1 presents the descriptive statistics of basic data. Comparing all variables, technical and professional levels of education exhibit the highest dispersion (looking at standard deviation), indicating that there are broad differences in educational terms across municipalities of Colombia. We arrive at an equivalent conclusion by checking the variable representing human capital representation, indicating heterogeneity in educational attainment. Hall (2013) points out in his research that geographical areas having a high level of human capital exhibit a higher probability of receiving firms and, in consequence, show superior levels of productivity.

Table 1. Descriptive Statistics Spatial Model

	Observ	Mean	Std. Dev.	Min	Max
Lnind	1118	3.153302	1.71034	0	10.51070
Lnagr	1118	8.900346	2.28957	0	14.90450
Lnser	1118	4.274094	1.80796	0	11.78070
Lnpobl	1118	9.413448	1.49491	0	15.73832
Tecprof	1118	3.983148	3.251993	0	2.82e+07
Khumano	1118	2.41e+07	3.183349	0	3.41e+07

Other statistically dispersed variable is the number of service units, showing that in terms of economic diversity, Colombian municipalities are highly heterogeneous. As endogenous variable, we use the log of the number of industrial units.

For spatial econometrics, Anselin proposes arranging a sequence of models for incorporating several spatial interactions in order to apply contrast that could help to identify the true spatial data generation process. The suggested order starts with OLS estimation as a benchmark basis for generating a test on residuals, and gradually, a set of Lagrange multiplicators are calculated to determine the type of spatial interaction, namely spatial lag, error lag or spatial Durbin model. The spatial Durbin model goes back to 1988, looking for the contribution of Anselin, who includes spatially lagged dependent variable (WY) and spatially lagged exogenous variables (WX).

To cover the spatial relationship across geographical units, we needed to define a weight matrix that describes the spatial arrangement of data within the sample. We generated a spatial weight matrix (W) as a contiguity binary matrix, hence the direction of spatial association flow is similar to the queen movement in chess. The matrix W should be a non-negative matrix of known constants, and the diagonal elements are null, denoting that units cannot be their own neighbour. Matrix W describes the spatial structure of data, and so far, it is generally understood as non-stochastic in consequence, it`s assumed as previously known, with all results conditional upon its definition (Dubin, 1998). We assume matrix W, according to the method of the nearest neighbour defining the off-diagonal elements as wij =1 in case when no observation is closer to either i or j, assuming a null value otherwise.

Following Elhorst (2010), we assess two paths to deal with spatial autocorrelation namely, a specific-to-general treatment and a general-to-specific approach for detecting the true-data generation process, each of them dealing with a different sequence of models performed. The Florax approach favours the first approach, starting from a standard linear regression and afterwards applying a statistical contrast to support the incorporation of spatial interactions and spatially lagged variables. The decision criteria for choosing the adequate data-generating process is guided by the Lagrange Multiplicator Test (and the Robust one), proposed by Anselin and based on the residuals stemming from OLS estimation that follows a Chi square distribution.

The particular-to-general procedure will guide our approach. In doing so, this paper first performs OLS regression assuming null values of autocorrelation parameters ($\rho=\lambda=0$). In these terms, the contrast of spatial autocorrelation will be applied, and in order to confirm the presence of auto-

DOI: 10.34659/eis.2024.91.4.838

correlation, the Lagrange multiplier test must be performed to choose the suitable structure for spatial dependence (Helsen, 2008).

Table 2 shows the results of the OLS model, indicating the basic behaviour of the model and standard contrast applied in the presence of spatial autocorrelation in order to evaluate the pertinence of spatial lag or spatial error pattern.

The adjusted R2 0.80 indicates that the fit of the model jet does not include relevant information about some explicative variables on model but is an acceptable value for the cases of cross-section or panel data arrangement.

Table 2. OLS Regression

	Coef.	Std. Err	t value	P>|t|
Lnagr	-0.02753	0.01182	-2.329	0.02004 *
Lnser	0.082251	0.01705	48.245	< 2e-16 ***
Lndensidad	0.07900	0.02226	3.549	0.000403***
Tecprof	0.02102	0.01069	1.9666	0.049515 *
Khumano	-0.03146	0.01101	-2.856	0.004368
_cons	-0.18601	0.25773	-0.722	0.470617
Adjusted R-squared:	0.8074			
Multiple R-squared:	0.8083			

The next step, according to Elhorst (2010), is to evaluate if OLS must be ruled out in favour of the spatial lag process and spatial error process or if the spatial Durbin Model must also be performed incorporating spatially lagged exogenous variables.

According to Table 3, strong evidence of autocorrelation exists in the spatial distribution of manufacturing units all around the country, and this argument is confirmed by the Moran Index, despite that this test is described as a very general contrast about spatial autocorrelation (Bernat, 1996). Consequently, two additional tests provide an explanation of the character of spatial dependence: LMlag and LMerr, taking into account that the definition of spatial dependence guides the interpretation of spatial interaction. The Robust Lagrange Multiplier contrast confirms that the spatial error model must be selected because it corroborates the significance of statistics and because the values of such statistics are higher. The right choice of spatial dependence is crucial taking into account that each selection drives a diverse interpretation of the spatial dependence.

Table 3. Global Index I of Moran: variable lnind

Moran I index	Expected Index	Variance	Z score	p-value
0,028021	-0,000893	0,000016	7,222305	0,00000

Table 3 contains the results of the test for spatial autocorrelation that confirms the existence of such a phenomenon in our data. The confirmation of autocorrelation suggests that nearby municipalities are commonly influenced by similar values in close spaces. Such a phenomenon leads to correlation in error terms (Dubin, 1998). According to our results, we discover spatial autocorrelation in our set of data.

The next step requires the application of criteria to distinguish which one really describes the nature of the generation data process, and the procedure corresponds to the LM-test and the Robust LM-test that stem from the residuals of our former OLS estimation (Elhorst, 2010).

The Lagrange multiplicator test provides the decision criteria for choosing the true generation of a data process. The contrast operates based on a test of significance based on OLS residuals following chi square distribution with K degrees of freedom.

DOI: 10.34659/eis.2024.91.4.838

The contrast operates under a likelihood ratio, which can be used to prove several assumptions. The first hypothesis is H0: $\theta = 0$ which assesses if the spatial Durbin model can be reduced to a spatial lag model. The second hypothesis evaluates the following assertions: H0: $\theta + \rho\beta = 0$ aiming to verify whether the spatial Durbin Model can be collapsed into the spatial error model. In the case of a joint rejection in both hypotheses H0: $\theta = 0$ and H0: $\theta + \rho\beta = 0$, so the spatial Durbin model makes up as the best description of the data generation process (Elhorst, 2010).

Table 4. Robust Lagrange Multiplicator Test

		df	p-value
RLMerr	15.559	1	7.995e-05
RLMlag	10.073	1	0.001505

Spatial Lag and spatial error models

Tables 4 and 5 present two results of models, the first one corresponding to the case of the spatial lag model and the second reports the spatial error model.

Table 5. Spatial Lag Model

	Coef.	Std. Err	Z	P>|z|	[95% Conf. Interval]	
Lnagr	-.0245034	.0187193	-1.31	0.191	-.0611926	.0121858
Lnser	.8824198	.0301237	29.29	0.000	.8233784	.9414612
lndensidad	.1240626	.0375861	3.30	0.001	.0503951	.1977301
tecprof	.0237517	.0159181	1.49	0.136	-.0074471	.0549506
khumano	-.0354482	.0197402	-1.80	0.073	-.0741383	.0032418
_cons	.419993	.6146108	0.68	0.494	-.7846221	1.624608
Rho	-.1408266	.1746909	-0.81	0.420	-.4832146	.2015613
Number of obs	759					

Table 6. Spatial Error Model

	Coef.	Std. Err	Z	P>|z|	[95% Conf. Interval]	
Lnagr	-.0043338	.0196958	-0.22	0.826	-.0429369	.0342693
Lnser	.8901162	.0303201	29.36	0.000	.8306899	.9495425
lndensidad	.1186596	.0382055	3.11	0.002	.0437783	.1935409
Tecprof	.0231848	.0158944	1.46	0.145	-.0079677	.0543373
khumano	-.0344205	.0203035	-1.70	0.090	-.0742146	.0053737
_cons	-.4892415	.6053039	-0.81	0.419	-1.675615	.6971323
Lambda	.9308156	.067874	13.71	0.000	.797785	1.063846
Wald test of lambda=0:	chi2(1) = 188.070 (0.000)					
Likelihood ratio test of lambda=0:	chi2(1) = 26.906 (0.000)					
Lagrange multiplier test of lambda=0:	chi2(1) = 62.556 (0.000)					

DOI: 10.34659/eis.2024.91.4.838

The magnitude of the estimators shifts slightly between OLS estimations and two models of correction of spatial autocorrelation, but the more outstanding feature is found in variable "techprof" that becomes non-significant in the last estimations. It can demonstrate that the OLS estimator of such a variable was really biased (Chasco, 2008). In such terms, Bernat (1996) warns about a similar situation in the presence of spatial autocorrelation because that effect invalidates standard tests when OLS regression is estimated in a similar way as serial autocorrelation or heteroskedasticity do. As mentioned earlier, the statistical features of both models are similar although each one leads to a different final interpretation. In the lag model, a rho parameter appears, indicating the magnitude of the effect originating in neighbour entities on endogenous variables.

The error model is the more appropriate specification for describing the spatial dependence, so the influence of neighbouring areas on local manufacturing presence requires a special interpretation. In this context, the existence of an important number of manufacturing units would depend on surrounding manufacturers' presence (Audia et al., 2006), only to the extent to which the neighbouring municipalities have an important number of manufacturing units above or below of what is considered a normal level. Then, the influence originating in neighbour entities affects local values if a variable in nearer places deviates strongly from the expected value (Bernat, 1996).

As is apparent in Table 6 coefficient λ is significant and has a high magnitude, reinforcing the strong influence of a neighbouring industrial activity on the number of industrial units at a local level. The magnitude of λ indicates that a deviation of 1% in the number of industrial units in the neighbourhood has a similar effect on local industrial presence than the number of local sector services units. Moreover, it can be observed that a deviation of 1% in neighbour industrial presence has a seven times stronger impact than population density in each locality.

The coefficient of the number of service units is very high and strongly significant, revealing a close inter-sectorial interaction. The manufacturing production can easily build links to the urban sector of services. The concept of skilled labour measured on the basis of education enrollment has a negative effect on manufacturing presence, indicating that other kinds of human skills must be incorporated. Population density is revealed to be strongly significant and has a high coefficient, so manufacturing units tend to prefer to locate where important human agglomeration can generate attractive markets.

The literature recommends an additional informational criterion in order to choose adequately the pattern of spatial dependence, and to reinforce the selection of the model that has been based on the Lagrange multiplicator. In the presence of spatial autocorrelation, the criteria derived from r squared can not be applied trustfully, so it is necessary to use other criteria for evaluating the good fit of the model.

Table 7. The Durbin Model with Lagged Exogenous Variables

	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.204313	0.373284	0.547	0.584256
Lnagr	-0.013920	0.012343	-1.128	0.259682
Lnser	0.832514	0.017786	46.806	<2,00E16***
Lnpobl	0.085148	0.022677	3.755	0.000183***
tecprof	0.024266	0.012034	2.016	0.043997*
Khuman	-0.017508	0.013919	-1.258	0.208701
lag.lnagr	-0.041816	0.022397	-1.867	0.062165
lag.poligonos.lnser	-0.022331	0.031514	-0.709	0.478710
lag.poligonos.lnpop	-0.022746	0.038160	-0.596	0.551260
lag.poligonos.tecprof	-0.029640	0.019572	-1.514	0.130202
lag.poligonos.Khuman	-0.006404	0.020208	-0.317	0.751366
Adjusted R-squared:	0.8105			

DOI: 10.34659/eis.2024.91.4.838

Table 8. Spatial Durbin Model. Direct and Indirect Impacts

	Direct	Indirect	Total
Lnagr	-0.01391973	-0.041816310	-0.05573604
	(-1.127721)	(-1.8670083)	(-2.4574300)
Lnser	0.83251425	-0.022331422	0.81018283
	(46.806085)	(-0.7086197)	(25.9085830)
Lnpobl	0.08514751	-0.022745644	0.06240186
	(3.754771)	(-0.5960557)	(1.5564979)
Tecprof	0.02426615	-0.029640149	-0.00537400
	(2.016422)	(-1.5144299)	(-0.3021726)
Khuman	-0.01750832	-0.006404193	-0.02391251
	(-1.257877)	(-0.3169177)	(-1.4576786)

t – value in parentheses.

In the Durbin model, we find out the effect of lagged exogenous variables on the endogenous variable, and considering that such information is worthy of our purposes, we proceed to interpret it. As shown in Table 8, the summary of direct and indirect effects informs about the contemporaneous effect and the lagged influence coming from the X values in neighbouring spaces.

We found an interesting interpretation of the significance in the set of diverse effects according to the p values. As mentioned earlier, three variables have a significant direct effect, namely the presence of sector services units, the local population, and one variable related to people's education. The only lagged exogenous effect exerting a significant influence is the indirect effect of agricultural production.

We assume that the significance of spatially lagged agricultural production is related to the eminently urban structure of Colombian manufacturing. If the core of industrial production is located in big cities, it is unlikely that the raw material used as input could be absorbed from the scarce urban territory. It is more feasible to use the close agricultural municipalities as providers of primary and agricultural inputs for industry.

Panel Data

According to the Breusch and Pagan contrast, we can reject the null hypothesis of the inexistence of fixed effects; therefore, we reinforce the suitability of the panel data technique. Such a test corroborates the existence of time-invariable effects influencing the behaviour of industrial number of employees in the Colombian regions. The model used this variable as a proxy of the industrial presence, and the financial and mining regional production appears as exogenous as proxies of inter-sectorial links in each department. We also used the net education enrollment of the population to represent the level of regional human capital.

Table 9. Panel Data Model. Descriptive Statistics

	Observ	Mean	Std. Dev.	Min	Max
lnempl	140	8.917577	1.823369	4.094345	12.27558
lnfinance	140	7.229552	1.400469	2.639057	10.81874
lnmin	140	5.748543	1.81754	2.302585	9.509185
net_enrrol	140	40.99019	6.808037	18.43	55.01

DOI: 10.34659/eis.2024.91.4.838

Once we confirmed the existence of the time-invariable fixed effect, we performed both estimations using the within estimator and by random effects. Afterward, we applied the Hausman test, comparing the estimators derived from the within model with those of random effects.

Table 10. Data Panel Model

	(1)	(2)	(3)
	c1	c2	c3
VARIABLES	lnempl	lnempl	Lnempl
Lnfinance	0.652***	0.286**	0.652***
	(0.106)	(0.143)	(0.106)
net_enrroll_me	0.0262	0.0764**	0.0262
	(0.0220)	(0.0306)	(0.0220)
Lnmin	-0.0581	-0.236**	-0.0581
	(0.0718)	(0.0932)	(0.0718)
Constant	3.464***	5.078***	3.464***
	(0.928)	(1.351)	(0.928)
Observations	140	140	140
R-squared	0.308	0.137	0.308
Number of region		27	
State FE	YES	YES	YES
Year FE	YES	YES	YES

The interpretation of the Hausman test points out that the data panel estimations reject the hypothesis of inexistence of correlation between exogenous variables and fixed effects of observations. Consequently, we must extract the final interpretation of the model from the regressors found within the estimation.

Conclusions

The spatial distribution of industry in Colombia has been utterly uneven and concentrated, because large cities have retained modern activities characterised by innovation and skilled workers. Powerful effects have been transmitted from urban manufacturing centres, spurring industrial development in conterminous areas. This spatial contagion and influence of manufacturing activity towards neighbouring spaces work only in larger cities, being weak or inexistent in medium-sized cities and regional nodes. The peripheral spaces produce raw materials and some industrial commodities intended to meet local markets, exploiting reduced-economies of scale.

According to the Moran Test, the endogenous variable "number of industrial units" exhibits autocorrelation in the spatial distribution, confirming dependence across spatial entities. Analysing diverse parameters found out, we can assert that some spatial effects came up. The spatial methods offer useful and timely hints for our analysis even when they are not clearly discerned at first glance (Osland, 2010).

The decision criterion advises to chose the spatial error model, looking at strong significance of the Lagrange Multiplier test. The presence of industrial units in a specific place will be strongly affected by manufacturing presence in closer municipalities, but only if the number of establishments in neighbourhoods deviates considerably from the expected level (Bernat, 1996).

The industrial presence demonstrates strong linkages with the number of units in the services sector, indicating that both activities are predominantly urban and complementary. Population den-

DOI: 10.34659/eis.2024.91.4.838

sity has a significant role, suggesting the link between the industrial location and the power of the market size. Highly populated places spur economies of scale and attract industrial establishments.

On the other hand, in our Spatial Durbin Model, we incorporated a pair of interactions that stem from lagged endogenous and lagged exogenous variables. This model has a strong potential for econometrics because it estimates unbiased regressors, even if the true data generation stems from a spatial lag or a spatial autoregressive error process. The spatial spillover effects could be defined as local or global and are calibrated using different exogenous variable sets (Elhorst, 2010).

The Spatial Durbin Model contributed additional information because the indirect effect of agricultural production is significant for industrial employment and indicates strong links between the industry and rural activities, but not with local agriculture. Rather this is hurled to neighboring agricultural exploitation in other municipalities. In fact, the bulk of manufacturing is located in urban areas with a scarcity of space for agriculture; therefore, spaces in neighbouring municipalities must be exploited as a source of raw materials.

In the panel data, we confirmed that, in fact, unobservable fixed effects were present in the determination of manufacturing employment and that it was necessary to tackle the analysis of such latent effects. The results highlight the definitive influence of historical factors on the industrial location, coherent with the New Economic Geography analysis which stresses the early local development processes of industrialisation as a strong advantage for attracting plants, in a cumulative process reinforced over time (Krugman, 1992).

Certainly, former peripheral regions of Colombia received gradually few industrial branches, but they did not reach the highest levels of technological complexity, namely food, beverages, leather products, clothes, footwear and so on. It is evident that the attractiveness of intermediate departments relies on the proximity to natural resources, in order to transform them and achieve scales of production at a regional level.

Acknowledgements

The authors are deeply thankful to the Academy of Humanitas in Sosnowiec (Poland), to the Prague innovation institute within the framework of the project *Prague Smart Accelerator+* (Czech Republic), and to Masaryk Institute of Advanced Studies at the Czech Technical University in Prague (Czech Republic), by their stimulating academic and cuadruple-helix innovative support.

The contribution of the authors

Conceptualization, G.A. and H.Y.A.G.; theoretical background, G.A. and H.Y.A.G.; results and analysis, G.A. and H.Y.A.G.; conclusions, G.A. and H.Y.A.G.

The authors have read and agreed to the published version of the manuscript.

References

Anselin, L. (2001). Spatial Econometrics. In B. Baltagi (Ed.), *Companion to Theoretical Econometrics* (pp. 310-330). Oxford: Blackwell Publishing.

Arbia, G. (2001). Modelling the geography of economic activities on a continuous space. Papers in Regional Science, 80(4), 411-424. https://doi.org/10.1111/j.1435-5597.2001.tb01211.x

Arbia, G., Espa, G., & Quah, D. (2009). A class of spatial econometric methods in the empirical analysis of clusters of firms in the space. In G. Arbia & B.H. Baltagi (Eds.), *Spatial Econometrics. Studies in Empirical Economics* (pp. 81-103). Physica-Verlag HD. https://doi.org/10.1007/978-3-7908-2070-6_5

Arias-Gomez, H. Y., & Antošová, G. (2023). Regional manufacturing in the Czech Republic. Urban, Planning and Transport Research, 11(1), 2222154. https://doi.org/10.1080/21650020.2023.2222154

Audia, P. G., Freeman, J. H., & Reynolds, P. D. (2006). Organizational Foundings in Community Context: Instruments Manufacturers and Their Interrelationship with Other Organizations. Administrative Science Quarterly, 51(3), 381-419. https://doi.org/10.2189/asqu.51.3.381

Bernat, A. (1996). Does Manufacturing Matter? A Spatial Econometric View of Kaldor's view. Journal of Regional Science, 36(3), 463-477. https://doi.org/10.1111/j.1467-9787.1996.tb01112.x

Bond-Smith, S. (2024). Combining agglomeration economies and endogenous growth without scale effects. Spatial Economic Analysis, 19(3), 287-308. https://doi.org/10.1080/17421772.2023.2280106

Broschak, J. P. (2004). Managers' Mobility and Market Interface: The Effect of Managers' Career Mobility on the Dissolution of Market Ties. Administrative Science Quarterly, 49(4), 608-640. https://doi.org/10.2307/4131492

Camagni, R. (2005). *Economía urbana.* Barcelona: Antoni Bosch. (in Spanish).

Capello, R. (1999). Agglomeration Economies and Urban Productivity: Location vs Agglomeration Economies. *Proceedings of the 39th Congress of the European Regional Science Association: "Regional Cohesion and Competitiveness in 21st Century Europe",* Dublin, Ireland. https://hdl.handle.net/10419/114383

Capoani, L. (2023). Theory of Commercial Gravitational Fields in Economics: The Case of Europe. Networks and Spatial Economics, 23, 845-884. https://doi.org/10.1007/s11067-023-09591-2

Chasco, C. (2003). *Métodos gráficos del análisis exploratorio de datos espaciales.* Almería: Asociación Española de Economía Aplicada. (in Spanish).

Davis, D., & Weinstein, D. (1996). Does Economic Geography Matter for International Specialization? NBER Working paper series, 5706. http://www.nber.org/papers/w5706

Dubin, R. (1998). Spatial Autocorrelation: a Primer. Journal of Housing Economics, 7(4), 304-327. https://doi.org/10.1006/jhec.1998.0236

Duranton, G., & Overman, H. (2005). Testing for Localization Using Micro-geographic Data. The Review of Economic Studies, 72(4), 1077-1106. https://doi.org/10.1111/0034-6527.00362

Elhorst, P. (2010). Applied Spatial Econometrics: Raising the Bar. Spatial Economic Analysis, 5(1), 9-28. https://doi.org/10.1080/17421770903541772

Fu, S., & Hong, J. (2010). Testing Urbanization Economies in Manufacturing Industries: Urban Diversity or Urban Size? EERI Research Papers Series, 38. https://hdl.handle.net/10419/142599

Fujita, M., & Krugman, P. (2004). La Nueva Geografía Económica: Pasado, Presente y Futuro. Investigaciones Regionales, 4, 177-206. https://www.redalyc.org/articulo.oa?id=28900409 (in Spanish).

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

Georgallis, P. P., Dowell, G., & Durand, R. (2019). Shine on Me: Industry Coherence and Policy Support for Emerging Industries. Administrative Science Quarterly, 64(3), 503-541. https://doi.org/10.1177/0001839218771550

Hall, Robert E. (2013). „Comment on „Reference Dependence and Labor Market Fluctuations" in: *NBER Macroeconomics Annual 2013*, 28, 201-208, National Bureau of Economic Research, Inc.Hallen, B. L. (2008). The Causes and Consequences of the Initial Network Positions of New Organizations: From Whom Do Entrepreneurs Receive Investments? Administrative Science Quarterly, 53(4), 685-718. https://doi.org/10.2189/asqu.53.4.685

Haveman, H. A., & Nonnemaker, L. (2000). Competition in Multiple Geographic Markets: The Impact on Growth and Market Entry. Administrative Science Quarterly, 45(2), 232-267. https://doi.org/10.2307/2667071

Helsen, J. (2008). *Essays on the Spatial Analysis of Manufacturing* [Doctoral Dissertation]. Ken State University.

Katila, R., & Chen, E. L. (2008). Effects of Search Timing on Innovation: The Value of Not Being in Sync with Rivals. Administrative Science Quarterly, 53(4), 593-625. https://doi.org/10.2189/asqu.53.4.593

Krugman, P. (1979). Increasing Returns, Monopolistic Competition and International Trade. Journal of International Economy, 9(4), 469-479. https://doi.org/10.1016/0022-1996(79)90017-5

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. The American Economic Review, 70(5), 950-959. https://www.aeaweb.org/aer/top20/70.5.950-959.pdf

Krugman, P. (1991). Increasing Return and Economic Geography. Journal of Political Economy, 99(3), 483-499. https://www.jstor.org/stable/2937739

Krugman, P. (1992). *Geografía y Comercio*. Barcelona: Antoni Bosch. (in Spanish).

Krugman, P. (2008). The increasing Returns Revolution in Trade and Geography. Nobel Prize Lecture, 335-348. https://www.nobelprize.org/uploads/2018/06/krugman_lecture.pdf

Lu, J., & Tao, Z. (2008). Trends and Determinants of China's Industrial Agglomeration. Munich Personal RePEc Archive, 6597. https://mpra.ub.uni-muenchen.de/6597/

Marshall, A. (2005). *Principios de Economía*. Madrid: Fundación ICO. (in Spanish).

Melo, J. (1978). La Evolución Económica de Colombia, 1830-1900. In J.J. Uribe (Ed.), *Manual de Historia de Colombia* (pp. 1-64). Bogotá: Instituto Colombiano de Cultura. (in Spanish).

Meyer, D. (1983). Emergence of the american manufacturing belt: an interpretation. *Journal of historical geography* 9, 145 – 174, doi: 10.1016/0305-7488(83)90220-7

Moncayo, E. (2002). Nuevos Enfoques de Política Regional en América Latina: el Caso de Colombia. Enfoques Teóricos y Evidencias Empíricas sobre el Desarrollo Regional en Colombia. Archivos de Economía, 3284. (in Spanish).

Moncayo, E. (2007). *Dinámicas Regionales de la Industrialización*. Bogotá: Universidad Central. (in Spanish).

Moreno, R., & Vayá, E. (2002). Econometría Espacial: Nuevas Técnicas para el Análisis Regional. Una Aplicación a las Regiones Europeas. Investigaciones Regionales, 1, 83-106. https://www.redalyc.org/pdf/289/28900104.pdf (in Spanish).

Osland, L. (2010). An application of Spatial Econometrics in relation to Hedonic House Price Modeling. Journal of Real Estate Research, 32(3), 289-320. https://doi.org/10.1080/10835547.2010.12091282

Ottaviano, G., & Puga, D. (1997). Agglomeration in the Global Economy: a Survey of the New Economic Geography. Center for Economic Performance. Discussion Paper, 356, 1-32.

Rosenthal, S., & Strange, W. (2001). The Determinants of Agglomeration. Journal of Urban Economics, 50(2), 191-229. https://doi.org/10.1006/juec.2001.2230

Silva Colmenares, J., & Padilla Pardo, C. (2017). Colombian Industry 1975-2014: Inequity between Labor and Capital. Apuntes del Cenes, 36(63), 137-167. https://doi.org/10.19053/01203053.v36.n63.2017.5423

Whittington, K. B., Owen-Smith, J., & Powell, W. W. (2009). Networks, Propinquity, and Innovation in Knowledge-intensive Industries. Administrative Science Quarterly, 54(1), 90-122. https://doi.org/10.2189/asqu.2009.54.1.90