Angel Alañon-Pardo [1,2], Patrick J. Walsh[3], Rafael Myro[4]

Do neighbouring municipalities matter in industrial location decisions? Empirical evidence from Spain

[1] Departamento de Economía Aplicada I, Universidad Complutense de Madrid, 28223 Madrid, Spain

[2] Instituto Complutense de Estudios Internacionales (ICEI)

[3] Landcare Research-Manaaki Whenua, 231 Morrin Road St Johns, Auckland 1072, New Zealand

[4] Department of Applied Economics II, Universidad Complutense de Madrid, 28223 Madrid, Spain

Corresponding author: Email: aalanonp@ucm.es Telephone number: 34-1-3942470

Angel Alañon-Pardo 16-digit ORCID: 0000-0001-8985-0324

Rafael Myro 16-digit ORCID: 0000-0003-1926-5001

Abstract

This paper focuses on industrial location, assuming that entrepreneurs not only consider the advantages associated with a certain municipality, but also those coming from nearby areas. Exploratory analysis reflects the existence of spatial patterns in the creation of manufacturing establishments and sheds light on the geographical scope on which agglomeration economies operate in industrial location. Spatial Probit models and Standard Probit models with spatially lagged explanatory variables are estimated to test whether neighbouring municipalities' location decisions and characteristics, including agglomeration economies, matter in industrial location choices. Results show that neighbouring municipalities location decisions and characteristics help to explain location decisions of new establishments for 11 manufacturing industries in Spanish municipalities (NUTS V) over the period 1991-1995.

Key words: spatial location models, geographical scope agglomeration economies

JEL Classification Code: L6, R3

Acknowledgments

The authors would like to thank to two anonymous referees and to the editors of this journal for their useful comments and suggestions. The usual disclaimer applies.

1. Introduction

Since Alfred Marshall's pioneering *Principles of Economics*, a common theme in Urban and Regional Economics has been that the agglomeration of similar firms can boost firm productivity. Thus agglomeration economies are a key variable in the location decision process. Usually, only firms located in reduced areas, such as the city of Prato (Italy), or Silicon Valley (U.S.A.) (often referred to as Marshallian industrial districts), are supposed to get the advantages of agglomeration economies. However, one can expect that spillovers and other advantages derived from agglomeration economies might also provide benefits to plants locating in nearby areas, in addition to those in the same immediate town or municipality (Ellison and Glaeser, 1997). This issue is related to the so called geographical scope of agglomeration economies commonly assumed to attenuate over distance. In that perspective, the aim of this paper is to analyze whether the location decisions of manufacturing plants in Spanish municipalities are related to the location decisions taken in surrounding or neighbouring municipalities, and to give insight into the reasons for this agglomerative behaviour. In order to do so we will apply Spatial Econometric techniques to study the location decisions of 11 industries in Spanish municipalities.

Firms may cluster due to many reasons, such as history, random events, natural advantages or agglomeration economies (Marshall 1890; Krugman 1991a; Krugman 1991b; Ellison and Glaeser 1997; Ellison et al. 2010)[1]. The most usual classification of agglomeration economies comprises urbanization economies, when the industrial mix is diverse and firms also benefit from the services and facilities of urban areas, and localization economies or Marshallian external economies, when the advantages of clustering derive from the same industry (Hoover 1948)[2]. According to Marshall (1890), the sources of the so called agglomeration economies are: shared input markets[3], labour market pooling[4]; and human capital and knowledge spillovers[5]. A

---

[1] See Rosenthal and Strange (2004) for a review on the nature and sources of agglomeration economies.

[2] Hoover's classification also included internal economies of scale.

[3] "And presently subsidiary trades grow up in the neighborhood, supplying it with implements and materials, organizing its traffic, and in many ways conducing to the economy of its material. (Marshall 1890)".

[4] "A localized industry gains a great advantage from the fact that it offers a constant market for

similar concept to localization economies are the so called MAR externalities -named after Marshall (1890), Arrow (1962) and Romer (1986)- when the agglomeration of firms arises in an oligopolistic environment (Glaeser et al. 1992).

Most analyses of Marshallian externalities have usually focused on the aforementioned sources of agglomeration economies[6], and on the so called industrial scope, which deals with the distinction between localization economies and urbanization economies[7]. However, as it is pointed out in Rosenthal and Strange (2004), less attention has been paid to the other dimensions over which agglomeration economies extend: the temporal scope and the geographic scope. The temporal scope is related to whether the effects of these economies are felt immediately or whether there may be any time lag, since there may be static agglomeration economies and dynamic agglomeration economies (see Glaeser et al. (1992) or Henderson (1997)). The geographic scope deals with the attenuation of the benefits of agglomeration with physical distance, since, ceteris paribus, when economic agents are closer there is more potential for interaction. This paper is focused on this geographical dimension of agglomeration economies, using data from Spanish municipalities.

There is not much work done on the geographic scope of agglomeration economies, with existing studies exhibiting only limited evidence of benefits extending beyond town limits. Using US zip codes, Rosenthal and Strange (2003) show that the geographic scope of localization economies seems larger than urbanization economies. They found that employment outside the industry of focus had an inconsistent and frequently insignificant effect. For the Spanish municipalities, Viladecans-Marsal (2004), who limits her

---

skill. ... Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require; while men seeking employment naturally go to places where there are many employers who need such skills as theirs.(Marshall 1890)".

[5] "great are the advantages which people following the same skilled trade get from near neighbourhood to one another. The mysteries of the trade become no mysteries; but are as it were in the air … if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas. (Marshall 1890)".

[6] See Holmes (1999) or Bartleman et al (1994) for evidence about shared input markets. Jaffee (1989), Acs et al (1992), Jaffe et al. (1993), or Audretsch and Feldman (1996) provide evidence on the relevance of human capital and knowledge spillovers. Evidence on labour market pooling can be found in Baumgartner (1988), Diamond and Simon (1990), Moretti (2000) or Costa and Kanh (2001).

[7] See Henderson (2003), Glaeser et al (1992) or Duranton and Puga (2005).

analysis to the most crowded Spanish cities (over 15,000 inhabitants), found that urbanization economies influence location in most industries, while localization economies played a minor role, and the agglomeration effects only spilled over the city borders in three of the six manufacturing industries analyzed. Using similar techniques, but studying Catalan municipalities, Jofre-Montseny (2009) found evidence on the geographical scope of localization economies for the textile and wood and furniture industries, and for urbanization economies in medical, precision and optical instruments, chemical products and metal products except for machinery industries.

On the other hand, Van Soest et al (2006), working with zip code data from a Dutch province, conclude that agglomeration economies may well operate on a geographic scale that is smaller than a city, since they only found evidence for interurban externalities for manufacturing, which is analysed as a single industry. Simmie (1998), Suarez-Villa and Alrod (1998), and Arita and McCann (2000) also cast doubts on the spatial extent of agglomeration.

According to Tobler's first law of Geography "everything is related to everything else, but near things are more related than distant things" (Tobler 1970)[8]. That sentence is often used to explain the concept of spatial dependence or spatial autocorrelation, and to justify the need to check for spatial autocorrelation when dealing with spatial data and processes. There is spatial dependence or autocorrelation when the values of a variable in a certain location are related to the values of the same variable in neighbouring locations. Surprisingly spatial autocorrelation is seldom taken into consideration in industrial location decision analysis. Therefore, most of the studies referenced above are mainly based on non-spatial regression analysis[9], which limits their findings. To properly capture the geographical scope of agglomeration economies, controls for spatial dependence should be used[10]. Spatial tools allow location decisions to be influenced by the decisions

---

[8] See Miller (2004) for more information on Tobler's law.

[9] See Arauzo-Carod et al. (2010) for a review on methods and results of empirical studies in industrial location.

[10] There are some works following this way such as Viladecans-Marsal (2004), Autant-Bernard (2006) or LeSage et al (2010). While LeSage et al (2010) addresses spatial autocorrelation by estimating a spatial autorregresive probit model to study the decisions of reopen after Hurricane Katrina, the other papers model spatial effects including spatially-lagged explanatory variables. However, these other papers do not fully control for spatial dependence through the error term or

of firms in neighbouring or nearby municipalities. Ignoring these influences can cause a variety of issues in an empirical analysis.

The aim of this paper is to analyze the extent of dependence in location decisions between neighboring municipalities. Instead of building or testing a comprehensive or sophisticated location decision model, we focus on the similarities or dissimilarities of those location decisions among neighbouring municipalities.

We apply Spatial Econometrics (Spatial Probit models and Non Spatial Probit models with spatially lagged explanatory variables) to estimate a simple location decision model and Spatial Statistics techniques (BB Join Count Statistics and Moran' I Statistic) to analyse the spatial allocation of new manufacturing establishments in Spanish municipalities. Both methods examine spatial dependence in location decisions. Our dataset comprises the continental Spanish municipalities and 11 industries.

This paper is organized as follows. Section 2 provides data, the methodology both for the exploratory analysis and for the confirmatory analysis, and a simple location decision model is presented. Results are shown in section 3. Finally, the main conclusions of this research are set out in section 4.

2. A simple location model, the statistical methodology, the spatial unit of analysis and the data

In this section we introduce the model, the spatial econometrics and spatial statistics techniques that will be implemented in the next section, some considerations about the spatial unit of analysis, and the data.

2.1. Econometric specification

the likelihood function (Anselin, 1988). Viladecans-Marsal (2004) use an OLS IV estimator to analyse the role of agglomeration economies in most crowded Spanish municipalities. Autant-Bernard (2006) analyses the location of R&D establishments in French NUTS 2 using a conditional logit model. However, neither of the latter two papers use a full spatial econometric model, as we do here.

Usually, location models are constructed considering the location decision problem as one of "random" profit maximization[11] (Figueiredo et al, 2002). Following McFadden (1974) and Carlton (1983) it is considered that if an entrepreneur, who previously decided to open a new establishment in manufacturing industry *j*, locates in municipality *i* it will produce a potential profit of π_{ij}. Formally,

(1) $$\pi_{ij} = X_i + \varepsilon_{ij}$$

where X_i reflects internal characteristics of municipality *i* and ε_{ij} stands for a random variable, which is expected to be distributed independently. So, this entrepreneur will locate in municipality *i* if the potential profit is greater than in other municipalities, *m*, for instance, that is

(2) $$\pi_{ij} > \pi_{mj}$$

where $i \neq m$. This profit depends on a set of local characteristics, and it is usually expressed as a linear combination of these characteristics (Figueiredo et al. 2002). Thus, in our case this profit would also depend on the characteristics of the neighbouring area

(3) $$\pi_{ij}\ f(X_i, WX)$$

where the explanatory variables X_i and *WX* account for the local characteristics which impact on profits and for the relevant characteristics of the neighbouring municipalities respectively. *W* is a spatial weights matrix *(SWM)*, where w_{ij} is set to 1 if municipality i and municipality are considered neighbours, and to zero otherwise. So, *WX* could be substituted by $W\pi_{ij}$

(4) $$\pi_{ij}\ f(X_j, W\pi_{ij})$$

As it is not possible to observe π_{ij} (Ellison and Glaeser 1997), the dependent variable of location models is usually the number of new establishments or new firms created over a period of time, *LOC*. So, we may express *LOC* as a linear combination of independent variables from equation (3)

(5) $$LOC_{ij} = \Sigma_n \beta_n X_{ni} + \Sigma_n \rho_n WX_{ni} + \varepsilon_{ij}.$$

---

[11] Called random, since it follows from the random utility framework. See Guimarães et al. (2004) for an extension of the random utility framework.

Location decision models are usually estimated using limited dependent variable models, i.e., Logit, Probit or Poisson specifications[12]. However, there are potentially a variety of unobserved (or difficult to quantify) influences that could cause location decisions to be spatially dependent. For instance, some areas may have better infrastructure or road networks that are conducive to manufacturing. If LOC_{ij} depends on what happens in neighbouring municipalities, the assumption of an independently distributed ε_{ij} is too strong. Two popular tests of spatial dependence are described in Section 2.3.The existence of spatial autocorrelation invalidates the use of most usual statistical and econometric techniques, such us ordinary least squares, or the basic logit or probit models[13]. If those models are used on spatially dependent data, biased or inefficient results will be obtained.

Spatial autocorrelation in data and processes may be treated in different ways. A simple approach may be to try to remove it from the dataset[14], but this is often not sufficient. Alternatively, spatial controls can be included in the specification of the model. The two most common approaches to the later method are the spatial autoregressive model (SAR) and the spatial error model (SEM)[15].

Three models will be estimated for each manufacturing industry: a standard Probit with spatially lagged explanatory variables, (PLEV), a Bayesian spatial autoregressive probit, (SARP), and a Bayesian spatial error probit, (SEMP).

As changes in explanatory variables for municipality $_i$ will have a direct impact on the location decisions of municipality $_i$, as well as an indirect or spatial

---

[12] See Arauzo-Carod (2002), Holl (2004a) and (2004b) or Guimaräes et al. (2004).

[13] See Anselin (1988) for more information about spatial autocorrelation and Spatial Econometrics techniques.

[14] By implementing robust estimation techniques, applying spatial filters or enlarging or improving the dataset, etc.

[15] SAR models include a spatially lagged dependent variable, Wy, as one of the explanatory variables, that is $y = \rho Wy + X\beta + \varepsilon$, where y is a n_x1 vector of observations on the dependent variable and Wy is an n_x1 vector of spatial lags for the dependent variable (where again, W is an *SWM)*. The parameter ρ is the spatial autoregressive coefficient that indicates the strength of spatial dependence, X is an n_xk matrix of observations on the (exogenous) explanatory variables with an associated β k_x1 vector of regression coefficients, and ε is an n_x1 vector of normally distributed (N(0, σ^2)) random error terms.
SEM models deal with spatial dependence through a spatially lagged error term, which uses a non-spherical error: $y = X\beta + u$, where $u = \lambda Wu + \varepsilon$, and $\varepsilon \sim N(0,\sigma^2 I_n)$. λ is a coefficient on the spatially correlated errors. See Anselin (1988) for additional details.

spillover impact on neighbours, following Lesage and Page (2009) we will estimate total, direct and indirect effects of SARP models.

However, since the indirect and indirect effects of SAR models are global (Lesage and Page, 2009) and that location processes may seem more localized, we will also estimate SEMP models with spatially lagged explanatory variables.

As a dependent variable, we use LOC_{ij}, a binary variable which is set to 1 if the location decision industry *j* is implemented in municipality *i* over the period 1991- 1995[16] and to 0 otherwise. We estimate an equation for each one of the eleven manufacturing industries considered. The normal approach to this type of data would be to use a probit or logit model[17]. In the presence of spatial autocorrelation however, standard Logit and Probit models are not very useful since ε does not follow a normal distribution. The majority of spatial econometric models with a continuous dependent variable use maximum likelihood techniques. However, with a binary dependent variable, there is no closed form solution to probit or logit probabilities (Anselin 2002, LeSage and Pace 2009).

We therefore use an alternative approach, which employs Bayesian methods to control for spatial dependence (Lesage, 1997 and 2000; Smith and Lesage 2002). Although there are other less popular alternatives[18], such as the generalized methods of moments (GMM) estimation (Pinkse and Slade 1988); or the EM (expectation maximization) approach for error models (McMillen, 1995), Bayesian methods represent the most comprehensive approach with a range of support and previous literature. This approach, proposed in Lesage (1997, 2000) and Smith and Lesage (2002) "is the most flexible of the spatially dependent models because it can incorporate spatial

[16] We choose that period because of the availability of data for the dependent and independent variables.

[17] If we were not interested on the location decisions but in the creation of new manufacturing establishments there are several ways to estimate spatial count data models. Kaiser and Cressie (1997) developed a Poisson auto-model which allows positive spatial dependencies in multivariate count data by specifying conditional distributions as truncated or Winsorized Poisson probability mass functions, and Poisson spatial interaction models are estimated in Lesage et al (2007) and in Fischer and Griffith (2008) to analyse origin-destination patent citation data.

[18] See Fleming (2004) for a more complete discussion on the advantages and disadvantages of different spatial Probit estimation techniques.

lag dependence and spatial error dependence in addition to general heteroskedasticity, of unknown form (Fleming, 2004, p.166-167)".

The Bayesian approach used here has its foundations in a non-spatial paper by Albert and Chib (2003), who model the binary dependent variable *y* as an indicator of unobserved latent utility *y** (LeSage and Pace, 2009). The relationship between *y* and *y** is as follows: $y_i = 1$ if $y_i^* \geq 0$, and $y_i = 0$ if $y_i^* < 0$. In the present application, when the net utility ($y_i^* \geq 0$) of locating in municipality *i* is positive, $y_i = 1$ and the firm selects *i* for its location. Albert and Chib (1993) recognized that $p(\beta,\sigma^2 \mid y^*) = p(\beta,\sigma^2 \mid y^*,y)$, since if you have *y** you have all the information needed to create *y*. This significantly simplifies the problem, because if *y** is added as an additional parameter to be estimated, then the joint conditional posterior distribution of β and σ^2 can be modelled as the same form as a continuous dependent variable Bayesian regression (LeSage and Pace, 2009; LeSage et al. 2011).

Instead of having to numerically integrate over the conditional distributions, Albert and Chib's (1993) contribution allows us to use Bayesian Markov chain Monte Carlo methods to sample each parameter from its conditional distribution. After numerous iterations of this sampling algorithm, a set of draws is produced that converges to the unconditional joint posterior distribution (full details are contained in LeSage and Pace (2009)). For instance, the conditional distributions

of ρ in the SAR model, and λ in the SEM model, as follows[19].

$$p(\rho \mid \beta, y^*) \propto |I_n - \rho W| \exp\left(-\frac{1}{2}\big((I_n - \rho W)y^* - X\beta\big)'\big((I_n - \rho W)y^* - X\beta\big) \right)$$

(6)

$$p(\lambda \mid \beta, y^*) \propto |I_n - \lambda W| \exp\left(-\frac{1}{2}(y^* - X\beta)'(I_n - \lambda W)'(I_n - \lambda W)(y^* - X\beta) \right)$$

For the number of iterations, we use 10,000 draws along with a 2,500 draw "burn in", which is discarded, but used to better calibrate the initial parameter values. To determine if this number of draws is sufficient, Raftery-Lewis convergence diagnostics are employed. Although we implement several tests of

[19] Following LeSage and Pace (2009), we employ a normal prior distribution for the β parameters, which are conditional on an inverse gamma distribution for σ^2. The spatial parameters, λ and ρ, have uniform prior distributions.

spatial dependence below, there is not a robust method of choosing between the SAR and SEM models in the context of a binary dependent variable[20]. Consequently, both models are presented below.

2.2. Data Sources and Location Determinants

Location models try to explain how certain variables may influence location decisions. Most empirical work usually groups these variables into categories such as supply factors, demand factors, external economies and diseconomies, etc. (Guimaräes et al. 2004). Since our central focus is the spatial influence of neighbouring municipalities, we do not carry out an extensive analysis of location determinants[21]. As explained later, this is also due to the lack of data for NUTS V in Spain with regard to location factors such as labour cost, land prices or taxes[22] etc. The location determinants we are taking into consideration are: human capital as a supply factor; municipality product as a demand factor; local external economies (localization and urbanization); and the role of neighbouring municipalities' location decisions characteristics.

The human capital index, HC_i, is defined as the percentage of population with at least a secondary school degree in municipality *i* in 1991. The expected sign is positive since it reflects the skilled labour market. Municipality product in 1991, MP_i, reflects the volume of economic activity in the municipality, the potential market for new firms, so its expected sign is positive.

External economies are represented by the classic location quotient and by a diversity index.

The location quotient, LQ_{ij} represents the advantages of geographical specialization of municipality *i* in industry *j*, that is, traditional localization economies, Marshallian externalities or MAR's agglomeration economies in 1990. Its expected sign is positive. Since higher LQ_{ij} may be caused both by a large number of small firms and by a small number of large firms, besides

[20] Unlike the case of a continuous dependent variable, where Lagrange multiplier tests can be used to choose between the two models.

[21] See Hayter (1997), Guimaräes et al. (2000), Figueiredo (2002) or Guimaräes et al. (2004) for more information about locational determinants.

[22] Local tax data are not available for small municipalities due to statistical secrecy, and, as argued before, we should not use NUTS III in order to avoid MAUP or ecological fallacy problems.

localization externalities it may also reflect the effects of concentration or internal returns of scale. It is defined as follows:

(7) $$LQ_{i,j} = (E_{ij} / E_i) / (E_J / E_T)$$

where E_{ij} accounts for total employment in manufacturing activity *j* in municipality *i*, E_i for total employment in municipality *i*, E_J for national employment in manufacturing activity *j*, and E_T total national employment in all manufacturing activities.

DI_i is a manufacturing diversification index for municipality *i* in 1990. The expected sign of this variable is positive since manufacturing diversity may reflect the existence of inter-industrial external economies, such as the Jacobs type (Jacobs 1969; Glaeser et al. 1992), and also because the creation of new plants is biased towards more diversified cities (Duranton and Puga 2000). This index is based on the correction for differences in sectoral employment shares at the national level of the inverse of a Hirschman-Herfindahl index proposed in Duranton and Puga (2000):

(8) $$DI_i = 1 \Big/ \sum_j \left| s_{ij} - s_i \right|$$

where, s_{ij} is the share of manufacturing industry *j* in manufacturing employment in municipality *i*, and s_j is the share of manufacturing industry *j* in total national manufacturing employment.

Finally, we consider the potential role of neighbouring municipalities NM_i, that is, location decisions of neighbouring municipalities ante the characteristics of neighbouring municipalities. It may be measured by the spatially lagged independent variables in a standard (non spatial) Probit model and in Spatial Error models, (WHC_i, WLQ_{ij}, WDI_i and WMP_i), where *W* is an *SWM*, and by the spatially lagged dependent variable in a Spatial Autoregressive Probit model[23], ($WLOC_i$). While WHC_i and WMP_i account for the human capital and the potential market of neighbouring municipalities, WLQ_{ij} and WDI_i represent the geographical scope of agglomeration economies which are originated in neighbouring municipalities. Location decisions of neighbouring municipalities

[23] The economic interpretation of λWu in SEM models is not so straightforward.

in industry *j* are represented by $WLOC_i$. That is, $WLOC_i$ measures part of the geographical scope of location decisions.

Therefore, location decisions may be explained as a function of local and neighbouring municipalities variables, such as agglomeration economies, human capital and potential market through the following expression:

$$LOC_{ij} = f(HC_i, LQ_{ij}, DI_i, MP_i, NM_i) \quad (9)$$

As Ottaviano and Puga (1998) point out, literature on economic geography identifies economic agglomeration at different levels of aggregation, from the small scale, e.g. a highly specialized industrial district such as the city of Prato in Italy, to the large scale agglomerations that cut across states, such as the US "Manufacturing Belt" or the European "Hot Banana". Since the geographic scope of agglomeration economies do not seem to be very large, as described in the previous section, we focus on Spanish municipalities (NUTS V). It seems a sensible election to study both the location of new manufacturing plants or the geographical scope of agglomeration economies, (as shown in Holl (2004a), in Jofre-Montseny (2009) or in Viladecans-Marsal (2001, 2003 and 2004)), since the average size of Spanish municipalities is 64 km^2, which is 1/3 of the average size of the U.S. zip codes analyzed in Rosenthal and Strange (2003), and around 85 % of the municipalities considered[24] are smaller than 100 km^2.

Nevertheless, working with Spanish municipalities also imposes a hard data constraint since most municipality data are related to socio-demographic characteristics and they are not usually up to date, because they are often produced for decennial census or for other purposes. We could try to overcome this scarcity of data using data related to higher levels of spatial aggregation, such as NUTS III, as done in Holl (2004a) to proxy municipal wages, labour force qualification, sector and industry specialization, and industry share. Unfortunately, as it is widely known in spatial analysis but often ignored in location analysis, our analysis could be wrong due to the so called Modifiable Areal Unit Problem (MAUP)[25], which is a potential source of error that can affect spatial studies which use aggregate data sources, consist of both a scale and an

[24] In order to work with spatially continuous data, we consider 7,906 municipalities, that is, we ignore the municipalities which belong to Balearic Islands or to Canary Islands.

[25] The influence of MAUP on location analysis is addressed in Pablo-Martí and Muñoz-Yebra (2009).

aggregation problem and is related to the concept of ecological fallacy (Unwin 1996; Bailey and Gatrell 1995). Thus, as our target is not to fully explain location decisions or location determinants, but to test whether location decisions in a municipality are related to the ones taken in neighboring municipalities, we will only consider NUTS V data.

The data sources that we will use in our analysis are Registro de Establecimientos Industriales -Industrial Establishments Register- (REI), Censo de Población 1991 (1991 Population Census), Censo de Locales 1990 (1990 Establishments Census 1990), and Alañón (2002). REI data[26] will allow us to study the spatial allocation of new manufacturing establishments in Spanish municipalities for 11 industries at 2 CNAE-93 digit level (Spanish classification of economics activities at 2 digit level). The industries considered are: food and tobacco; clothes and leather; wood and furniture; printing and paper; chemistry; other non metallic minerals; first transf. of metals; machinery; computer, office equipment, etc.; electric and electronic equipment; and transport equipment. We have data from 1980 to 1998[27]. 1991 Population Census and 1990 Establishments Census are the last Spanish Census whose municipality data are available for all municipalities. Census data will allow us to build indicators for the advantages derived from human capital, and agglomeration economies. Alañón-Pardo (2002) provides gross domestic product of Spanish municipalities for 1991.

Due to the restrictions of the data sources referred above, while the spatial exploratory analysis will cover the 1980-1998 period, the regression analysis will be limited to the 1991-1995 period.

2.3. The spatial statistics tools

In this section we introduce the BB Join Count statistic and Moran's I statistic that will be applied to study the spatial allocation of new manufacturing plants in Spanish municipalities.

---

[26] All manufacturing establishments must be registered in REI before starting up its activities. See Mompó and Monfort (1989) for a description of REI.

[27] During the nineties regional governments started managing REI delegations, and data about new establishments are neither provided in a timely fashion for all the regions nor in a friendly format to be processed.

The BB Joint Count Test[28] for spatial autocorrelation or spatial dependence reflects whether binary variables are clustered or randomly distributed in space. The BB Join Count Test is defined as follows:

$$BB = (1/2)\sum_i \sum_h w_{ih} LOC_i LOC_h \qquad (10)$$

where *LOC* is a binary variable, which is set to 1 when a manufacturing establishment is created over a period of time, and *LOC* is set to 0 otherwise. W_{ih} is the *i-th* element of a spatial weights matrix *W*, which reflects whether municipalities *i* and *h* share a common border, that is, they are neighbours. Thus BB reflects the number of times a municipality where there has been manufacturing births is contiguous to another municipality where there has been manufacturing births. A positive and significant z-value for this statistic indicates positive autocorrelation, that is, for a given manufacturing industry establishments births are more spatially clustered than might be caused purely by chance (Anselin 1992).

Using a measure of spatial autocorrelation for a binary variable seems sensible, since we are interested on whether the location decision is implemented or not. However, it could be argued that in our case, the measure could produce misleading results, since *LOC* is a binary variable which does not account for the number of establishments created. The BB statistic will be the same whether there is one or many new establishments created in the municipality.

In order to avoid this criticism, we will also apply Moran's I statistic[29], which is defined as follows:

$$I = N / S_0 \sum_i \sum_h w_{ih}(x_i - \mu)(x_h - \mu) / \sum_i (x_i - \mu)^2 \qquad (11)$$

where N is the number of observations; w_{ih} is as defined above; x_i and x_h are the number of new establishments of a given manufacturing activity which have been set up in municipalities *i* and *h* respectively; and S_0 is a scaling constant, $S_0 = \sum_i \sum_h w_{ih}$. A positive and significant z-value for this statistic indicates

[28] As Anselin (1992) points out binary variables take on only the values 1 and 0, areal units with observations 1 are often referred to as coloured Black. Black-Black (BB) join counts is the number of times a join, coloured area, is contiguous to another Black unit. See also Cliff and Ord (1980) for technical details.

[29] See Cliff and Ord (1980) or Anselin (1988) for more information about Moran's I statistic.

positive spatial autocorrelation, that is, municipalities which have been chosen as locations for the new entries in a given manufacturing activity tend to be close to each other.

If BB Join Count statistic and Moran's I statistic show there is spatial autocorrelation in location decisions and in the creation of new manufacturing establishments respectively, it does not necessarily mean that this spatial co-location is due to Marshallian agglomeration economies, since firms may cluster because of history, random events, natural advantages etc., as noted in the introduction. So, if the location decisions and the establishments births are spatially autocorrelated we will apply Moran's I statistic to the location quotient of the 11 manufacturing industries considered. The location quotient, LQ_{ij}, represents advantages of geographical specialization, traditional localization economies, Marshallian externalities or MAR's type agglomeration economies. If the location quotient, or municipality specialization in a given industry, is autocorrelated in space, then location decisions and establishment births may be autocorrelated in space in order to get the advantages derived from a specialized environment.

3. Results

3.1. Exploratory analysis results

In this section we provide results on the spatial statistics tools applied to the location decisions (Table 1), on the creation of new manufacturing establishments (Table 2), and on the manufacturing industry specialization in the Spanish municipalities (Table 3)[30]. These analyses correspond to the 1980-1998 period and involve 11 manufacturing industries[31].

As can be seen in Table 1, which shows the BB Join Count Test on the location decisions in Spanish municipalities, location decisions are spatially autocorrelated in all the manufacturing industries considered, except for

[30] As most of spatial statistics are significant, in order to reduce the length of tables 1 and 2, we will only show results for 1980, 1985, 1990, 1995, 1998, and for the years in which some of the spatial statistics are not significant.

[31] As our dataset comprises 19 years and 11 manufacturing industries, due to length limitations, the descriptive analysis only include the number of municipalities in which there was creation of manufacturing establishments, the number of manufacturing establishments created per year, and the maximum of establishments created in a given year (see appendix 1).

computer and office equipment and electric and electronic equipment industries in 1980 and 1981. That is, municipalities which have been chosen for the location of manufacturing establishments of a given industry tend to share a common border with other municipalities where there are manufacturing births for that industry, in a fashion greater than could be caused purely by chance.

Looking at the number of births for every manufacturing industry in Table 2, results are very similar. Thus, both positive location decisions for a given industry and a given year, and the number of manufacturing births, are autocorrelated in space. These spatial patterns may be due to Marshallian agglomeration economies or to other reasons, as stated at the beginning of the introduction. In order to support the evidence for Marshallian agglomeration economies, Moran's I statistic is applied to the level of municipality specialization in every manufacturing industry considered, which is measured through the location quotient, defined in expression 9. As shown in Table 3, except for the food industry, which is widely spread across the Spanish territory, specialized municipalities in a given industry tend to be neighbors. So, since municipality specialization in a given industry is autocorrelated in space, and so are location decisions and new manufacturing births, we may not reject that the benefits of locating in specialized municipalities are behind these spatial patterns.

Table 1 BB Join Count test (1980-1998)

Industry												
	Food & tobacco		Clothes & leather		Wood & furniture		Printing & paper		Chemistry		Other non metallic minerals	
Year	BB	z-val	BB	z-val	BB	z-val	BB	z-val	BB	z-val	BB	z-val
1980	350	23.8	48	27.3	86	34.4	10	22.5	52	41.7	25	18.4
1985	617	26.5	164	48.9	173	44.2	31	23.5	126	58.9	43	31.0
1990	543	28.1	156	42.7	217	45.0	71	37.8	108	46.4	87	36.6
1995	622	35.4	109	47.3	172	43.9	60	39.5	86	46.1	47	33.5
1998	506	48.1	95	46.8	82	41.0	47	46.0	87	63.2	45	32.0

Industry										
	First transformation of metals		Machinery		Computer. office equipment etc		Electric & electronic equipment		Transport equipment	
Year	BB	z-val	BB	z-val	BB	z-val	BB	z-val	BB	z-val
1980	117	34.2	9	11.0	0	-0.1	0	-0.3	4	12.2
1981	94	29.5	10	15.2	0	-0.1	1	2.4	5	13.6
1985	198	49.6	67	38.2	5	16.3	12	19.0	8	16.4
1990	255	53.9	74	41.4	5	15.3	18	27.0	39	32.6
1995	227	51.8	60	36.8	3	10.4	23	33.7	18	25.4
1998	152	59.5	66	49.1	6	26.7	14	22.3	8	13.9

Table 2 Morans's I statistic (1980-1998)												
Industry												
	Food & tobacco		Clothes & leather		Wood & furniture		Printing & paper		Chemistry		Other non metallic minerals	
Year	I	z	I	z-val	I	z-val	I	z-val	I	z-val	I	z-val
1980	0.1	21.9	0.2	26.2	0.2	25.1	0.1	14.5	0.2	27.9	0.1	15.2
1985	0.1	14.6	0.1	11.3	0.3	48.6	0.1	18.1	0.3	43.5	0.3	43.1
1990	0.1	18.4	0.2	24.0	0.2	31.9	0.3	39.4	0.3	43.8	0.2	36.4
1995	0.2	26.4	0.2	27.5	0.3	38.6	0.3	38.7	0.4	55.4	0.3	41.9
1998	0.1	20.1	0.1	20.2	0.3	39.5	0.2	36.1	0.3	39.3	0.3	38.6

Industry										
	First transformation of metals		Machinery		Computer. office equipment etc		Electric & electronic equipment		Transport equipment	
Year	I	z-val	I	z-val	I	z-val	I	z-val	I	z-val
1980	0.2	24.8	0.1	8.5	0.0	-0.1	0.0	-0.4	0.1	12.8
1981	0.1	20.3	0.0	6.8	0.0	-0.2	0.0	0.2	0.1	11.4
1984	0.4	53.3	0.2	29.7	0.0	0.1	0.1	14.8	0.1	19.3
1985	0.3	48.1	0.3	47.5	0.0	1.5	0.2	35.3	0.1	7.5
1990	0.3	46.6	0.2	29.9	0.0	5.0	0.2	30.0	0.3	37.2
1995	0.4	53.6	0.2	35.6	0.0	7.1	0.2	32.7	0.1	20.5
1998	0.4	56.5	0.3	42.3	0.1	17.5	0.1	20.5	0.1	11.9

Table 3: Moran's I statistic on municipality specialization (location quotient)		
Industry	Moran's I	z value
Food, beverages and tobacco	0.009	1.397
Clothes and leather	0.260	38.522
Wood and furniture	0.253	37.426
Printing and Paper	0.174	25.835
Chemistry	0.085	12.536
Other non metallic minerals	0.189	27.945
First transf. of metals	0.221	32.680
Machinery	0.115	17.018
Computer, office equipment, etc	0.015	2.183
Electric and electr. Equipment	0.114	16.831
Transport equipment	0.128	19.033

3.2. Econometric results

In this section As noted in section 2.1, three models are estimated for each manufacturing industry: a standard Probit with spatially lagged explanatory variables, (PLEV), a Bayesian spatial autoregressive probit, (SARP), and Bayesian spatial error probit with spatially lagged explanatory variables, (SEMP). The SARP and SEMPs Bayesian models both allow for heteroskedasticity. Spatially lagged explanatory variables in PLEV models are built with first order contiguity *SWM*. As PLEV models results suggest spatial effects do exist in location decisions, we extend the geographical scope of these effects. The Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) was used to select the *SWM* specification.

This criterion is commonly used in Bayesian analyses with competing models (LeSage et al. 2011), and is based on the model likelihood. The DIC provides a measure of fit, which adjusts for the complexity of a model. Formally, the DIC is defined as:

$$DIC = \bar{D}(\boldsymbol{\theta}) + p_D \qquad (12)$$

Where D(θ) = -2LL(θ), or negative two times the log likelihood, and

$$p_D = \bar{D}(\boldsymbol{\theta}) - D(\bar{\boldsymbol{\theta}}) \qquad (13)$$

where $D(\bar{\boldsymbol{\theta}})$ is the deviance calculated using the mean of the parameters $\bar{\boldsymbol{\theta}}$ obtained from the MCMC draws, and the average deviance ($\bar{D}$) is computed by taking the average of the deviance over the MCMC draws (Spiegelhalter et al., 2002). As can be seen in Table 4, multiple SWMs were examined, including nearest neighbors, NN, inverse distance, InvDist, and inverse distance squared, InvDistSQ. The 20 NN SWM and the InvDistSQ SWM for 10 kilometres had the lowest DIC score for SEMP and SARP models respectively (with the difference in DICs much greater than 7 in each case), providing strong evidence for the superiority of these models (LeSage et al. 2011). Note that DIC in SARP models is lower to the one in SEMP models.

To test for convergence of the MCMC routines, Raftery-Lewis convergence diagnostics (LeSage and Pace, 2009) were used. Results indicate that convergence was achieved in fewer than 4,000 draws for all models, with the majority converging at around 2,000 draws.

The results of the econometric models are summarized in Tables 5-10. All non- spatially lagged explanatory variables, except for *LQ* in the Food and Tobacco industry in SARP and SEMP models, are significant and show the expected sign across all three models. According to these results we cannot reject that population skills, manufacturing specialization (localization economies), market potential, and diversity (urbanization or Jacobs external economies) play an important role in location processes. Results for Food industry in spatial probit models are consistent with the lack of significance of Moran's I for the location quotient in Table 3.

These results differ to a certain extent from the evidence shown in previous studies, such as Viladecans-Marsal (2004), where urbanization economies influence location in most sectors, but specialization only plays a minor role.

Looking at the spatially lagged explanatory variables in the PLEV models in Table 5, which account for the sources of agglomeration economies in neighboring municipalities, *WLQ* and *WDI*, are always significant and show the expected sign except for *WLQ* in Food and in First transformation of metals. However, *WLQ* is highly significant all the other industries, which could reflect the positive effect of neighbouring municipalities due to Marshallian agglomeration economies. As noted in section 3, the insignificant Food results may be due to the fact that this industry is highly spread across Spain[32].

The high significance of the spatially lagged diversity indicator, *WDI,* stresses the key role of inter-industrial linkages at an interurban level. As was suggested at the beginning of this paper and in the comments on *WLQ* and *WDI* indicator they also support evidence on the geographical scope of agglomeration economies.

A striking result is the lack of significance of the spatially lagged Human Capital indicator, *WHC,* in most manufacturing activities. It could mean that commuting is not very important in Spain as a whole (excluding the biggest cities) or that the commuters are not very skilled, but that its effect is also represented in *WLQ* since a qualified labour market is also a source of agglomeration economies.

The spatially lagged potential market indicator, *WMP,* is not significant in most manufacturing activities. Therefore, decision-makers seem to focus primarily on their internal market.

Moving on to the full spatial models in Tables 6 and 7, note that the spatial error and lag parameters, λ and ρ, are significant in all models except computers and office equipment (SARP, and SEMP models) and electric and electronic equipment and transport equipment (SEMP models). Computers and office equipment is a manufacturing industry highly clustered

[32] If we could disaggregate the Food industry, results would probably differ.

in certain areas, and not very widespread in Spain. This agrees with the findings of the BB Joint Count test. Also, if we use ρ as a measure of the spatial dependence present in the SARP model, computers and office equipment has the lowest coefficient at 0.09. It also has the lowest λ coefficient in Table 6. The strongest spatial dependence is shown in food industry, since spatial autoregressive coefficient ρ is 0.57, which is consistent with the fact that this industry is highly spread across Spain. As λ and ρ are highly significant most manufacturing industries analyzed we cannot reject that location decisions in neighboring municipalities matter in industrial location decisions.

The coefficient estimates from Tables 6 and 7[33] are not easily compared to Table 6, since the impact of both the coefficient and its lag must be accounted for in the latter. Although some of the non-spatial (Table 5) coefficients are within the credible intervals for the spatial results—such as *LQ* for all estimates except machinery—there are many others that do not fall within the interval.

As stated in section 2.1 as location processes may seem more localized, our SEMP models include spatially lagged explanatory variables (Table 6). Results on these variables do not differ much from the ones in PLEV models. *WHC* is not significant or present a negative sign in most industries; WLQ is significant in all industries but food; and *WMP* and *WDI* are significant and show the expected sign in all industries.

As shown in Table 4, according to DIC criteria SAR models get a better fit than SEM ones. Effect estimates for these models are shown in Tables 8-10. As expected, direct effects, Table 9, are larger than indirect effects, Table 10, in all industries. All explanatory variables are significant, but *LQ* in food industry. Location decisions of each municipality seem more influenced by changes in human capital (*HC*) and industrial diversity (*DI*).

The indirect effect or spatial spillovers impact on neighbour municipalities of each explanatory variable is shown in Table 10. These results are mostly consistent with most of the ones in spatially lagged variables in PLEV and SEM models. However, human capital is significant and shows the expected sign in most industries. Changes in neighbouring human capital and in industrial

[33] The coefficient estimates from the SARP models (table 7) cannot be interpreted as representing how changes in the explanatory variables affect location decisions. In order to do so, direct and indirect effects have to be estimated (tables 9-10). See Lesage et al (2011) for more information.

diversity seem to have larger impact on location decision than the ones in municipality product and industrial specialization.

These results highlight the importance of properly controlling for spatial dependence. Although past papers have used specifications similar to Table 5, that kind of model does not fully control for the error structure of spatial dependence. Although Viladecans-Marsal (2004) provides empirical evidence on the geographical scope of agglomeration economies in the biggest Spanish cities, her results differ, since agglomeration effects only spill over beyond the administrative borders in three of the six industries analyzed[34]

[34] We must bear in mind that these studies were not carried out using the same methodology and do not use exactly the same dataset, thus full comparison is not possible.

Table 4 DIC for SEM and SAR models											
	Food and tobacco	Clothes and leather	Wood and furniture	Printing and paper	Chemistry	Other non metallic minerals	First transf. or metals	Machinery	Computers and office equipment	Electric and electronic equipment	Transport equipment
SEM Models											
InvDist10000	25109	28527	26237	31286	30041	28058	26042	30312	38578	34342	33279
InvDist15000	24426	28421	26094	31233	29955	27982	25808	30231	38575	34321	33253
InvDist10000SQ	25210	28549	26264	31295	30064	28077	26073	30330	38613	34356	33290
InvDist15000SQ	24975	28510	26207	31281	30036	28046	25961	30304	38586	34357	33287
NNSWM10	24133	28383	26069	31186	29875	27956	25748	30209	38554	34293	33232
NNSWM15	23328	28230	25906	31145	29735	27893	25446	30143	38565	34240	33204
NNSWM20	22777	28022	25718	31080	29535	27803	25145	30069	38552	34211	33168
SAR Models											
InvDist10000	20353	20401	20299	20450	20433	20341	20335	20462	20671	20587	20535
InvDist15000	20435	20504	20345	20525	20585	20409	20397	20558	20676	20604	20569
InvDist10000SQ	17197	17075	17154	17027	17046	17050	17183	17069	16936	16993	16962
InvDist15000SQ	17256	17093	17182	17042	17062	17074	17229	17083	16935	16998	16966
NNSWM10	20399	20540	20341	20576	20655	20416	20391	20589	20673	20632	20598
NNSWM15	20310	20591	20352	20657	20761	20466	20370	20674	20705	20719	20649
NNSWM20	20191	20611	20334	20730	20818	20489	20336	20721	20729	20799	20716

Table 5: Standard Probit with Spatially lagged explanatory variables results (*SWM* = first order contiguity spatial weights matrix)

	Food and tobacco	Clothes and leather	Wood and furniture	Printing and paper	Chemistry	Other non metallic minerals	First transf. or metals	Machinery	Computers and office equipment	Electric and electronic equipment	Transport equipment
Constant	-2.422†	-3.275†	-3.034†	-3.705†	-3.72†	-3.022†	-2.976†	-3.861†	-4.374†	-4.254†	-3.701†
	(0.09)	(0.116)	(0.103)	(0.154)	(0.136)	(0.106)	(0.102)	(0.134)	(0.248)	(0.175)	(0.155)
Human Capital (HC)	1.76†	1.269†	1.738†	1.89†	1.919†	2.034†	1.365†	1.827†	2.577†	1.828†	2.672†
	(0.258)	(0.332)	(0.279)	(0.3962)	(0.353)	(0.306)	(0.282)	(0.365)	(0.545)	(0.457)	(0.396)
Loc Quotient (LQ)	0.002l‡	0.177†	0.058†	0.09l†	0.114†	0.056†	0.04†	0.038†	0.03†	0.027†	0.194†
	(0.001)	(0.012)	(0.007)	(0.014)	(0.014)	(0.005)	(0.008)	(0.007)	(0.01)	(0.004)	(0.023)
MunGDP (MP)	0.018†	0.02†	0.048†	0.054†	0.017†	0.017†	0.06†	0.035†	0.01†	0.016†	0.01†
	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)
Diversity Index (DI)	1.08†	1.202†	1.448†	0.956†	1.147†	1.146†	1.446†	1.043†	0.457†	0.902†	0.897†
	(0.06)	(0.079)	(0.074)	(0.092)	(0.083)	(0.073)	(0.075)	(0.086)	(0.113)	(0.097)	(0.088)
WHC	-1.92†	-1.005‡	0.416	-0.379	-0.33	-1.725†	-0.248	0.468	-0.371	1.068‡	-0.989‡
	(0.359)	(0.441)	(0.378)	(0.541)	(0.483)	(0.414)	(0.385)	(0.492)	(0.82)	(0.61)	(0.549)
WLQ	0.001	0.14†	0.059†	0.212†	0.277†	0.097†	0.026	0.037†	0.219†	0.072†	0.562†
	(0.001)	(0.023)	(0.0192)	(0.051)	(0.045)	(0.017)	(0.019)	(0.01)	(0.048)	(0.0277)	(0.084)
WMP	0.002‡	-0.001	0.003*	0.001	-0.001	0.000	0.005†	-0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
WDI	1.168†	0.989†	0.653†	0.837†	1.033†	0.913†	0.832†	0.968†	1.164†	0.821†	0.671†
	(0.091)	(0.114)	(0.099)	(0.137)	(0.121)	(0.104)	(0.097)	(0.123)	(0.211)	(0.159)	(0.141)
AIC+	0.801	0.525	0.722	0.321	0.415	0.556	0.723	0.383	0.120	0.215	0.268
McFadden R^2	0.243	0.347	0.326	0.456	0.367	0.298	0.341	0.383	0.450	0.413	0.372

Note: † indicates significance at the 0.01 level, ‡ indicates significance at the 0.05 level, and * indicates significance at the 0.10 level. +Akaike Information Criteria

Table 6: Spatial Error Probit Model (*SWM* = Nearest Neighbor 20)											
	Food and tobacco	Clothes and leather	Wood and furniture	Printing and paper	Chemistry	Other non metallic minerals	First transf. of metals	Machinery	Computers and office equipment	Electric and electronic equipment	Transport equip.
Constant	-2.835169†	-2.699286†	-3.118686†	-2.764437†	-2.871096†	-2.644153†	-3.232750†	-2.948293†	-2.168787†	-2.612012†	-2.402713†
SD	0.147848	0.118785	0.112933	0.127772	0.122833	0.112668	0.116003	0.117828	0.132195	0.125337	0.133556
Cred. Intervals (5/95)	-2.55/3.131	-2.469/-2.935	-2.900/-3.343	-2.516/-3.015	-2.636/-3.116	-2.423/-2.867	-3.004/-3.469	-2.725/-3.185	-1.925/-2.426	-2.372/-2.861	-2.138/-2.658
Human Capital (HC)	2.164033†	1.281672†	2.269255†	1.840696†	1.358541†	1.774066†	2.018827†	1.660995†	0.751790††	0.984306†	1.286685†
SD	0.303856	0.318749	0.297647	0.316516	0.345169	0.317856	0.300685	0.322098	0.369525	0.352084	0.342816
Cred. Intervals (5/95)	2.765/1.574	1.927/0.661	2.838/1.685	2.429/1.189	2.043/0.670	2.416/1.167	2.614/1.436	2.279/1.045	1.477/0.016	1.680/0.300	1.956/0.630
Loc Quotient (LQ)	0.001210	0.148753†	0.046339†	0.106920†	0.112517†	0.059625†	0.028765†	0.015705†	0.028912†	0.036499†	0.191570†
SD	0.001963	0.014931	0.008391	0.014355	0.018383	0.008316	0.008341	0.009449	0.011055	0.011204	0.027033
Cred. Intervals (5/95)	0.005/-0.003	0.178/0.120	0.063/0.031	0.134/0.078	0.150/0.080	0.076/0.043	0.046/0.014	0.040/0.004	0.049/0.007	0.060/0.016	0.247/0.140
MunGdp (MP)	0.001171†	0.001725†	0.001068††	0.001861†	0.001686†	0.001650†	0.000955††	0.001893†	0.002339†	0.002173†	0.001905†
SD	0.000691	0.000649	0.000615	0.000653	0.000693	0.000614	0.000618	0.000691	0.000665	0.000682	0.000638
Cred. Intervals (5/95)	0.003/0.000	0.003/0.001	0.002/0.000	0.003/0.001	0.003/0.001	0.003/0.001	0.002/0.000	0.003/0.001	0.004/0.001	0.004/0.001	0.003/0.001
Diversity Index (DI)	1.080585†	0.817369†	1.212116†	0.663196†	0.693520†	0.875924†	1.276097†	0.684944†	0.190156	0.422637†	0.454384†
SD	0.068741	0.068674	0.070346	0.071332	0.073235	0.071237	0.068913	0.074187	0.075891	0.075420	0.073614
Cred. Intervals (5/95)	1.219/0.946	0.951/0.682	1.343/1.070	0.800/0.524	0.831/0.544	1.014/0.740	1.411/1.147	0.836/0.546	0.338/0.043	0.568/0.273	0.597/0.310
WHC	-1.724061†	-0.766394*	-0.316770	-0.567331	0.091160	-1.23905††	0.110027	0.295245	-0.293630	0.483772	-0.542221
SD	0.571026	0.508802	0.458829	0.537009	0.520873	0.485846	0.489989	0.517336	0.594210	0.544021	0.584928
Cred. Intervals (5/95)	-0.612/-2.836	0.213/-1.759	0.600/-1.186	0.488/-1.607	1.130/-0.909	-0.329/-2.252	1.064/-0.837	1.305/-0.718	0.858/-1.469	1.562/-0.554	0.555/-1.684
WLQ	-0.004883	0.161111†	0.050510††	0.118822††	0.287357†	0.125260†	0.022629	0.032317†	0.114215††	0.050802††	0.531085†
SD	0.008787	0.031929	0.026965	0.055270	0.051676	0.024786	0.025025	0.010952	0.058445	0.032928	0.113720
Cred. Intervals (5/95)	0.012/-0.022	0.223/0.098	0.101/-0.004	0.224/0.012	0.387/0.182	0.174/0.077	0.073/-0.026	0.054/0.011	0.226/-0.003	0.111/-0.016	0.746/0.307
WMP	0.005286†	0.004993†	0.004072†	0.009241†	0.005070†	0.005135†	0.005275†	0.010056†	0.006533	0.007893†	0.004749†
SD	0.002018	0.001464	0.001698	0.001785	0.001588	0.001433	0.001897	0.001821	0.001300	0.001508	0.001295
Cred. Intervals (5/95)	0.009/0.002	0.008/0.002	0.008/0.001	0.013/0.006	0.008/0.002	0.008/0.002	0.009/0.002	0.014/0.007	0.009/0.004	0.011/0.005	0.007/0.002
WDI	1.466340†	0.698519†	1.042281†	0.533717†	0.627629†	0.670162†	1.237858†	0.575571†	0.048656†	0.222354*	0.182067*
SD	0.147475	0.134707	0.127669	0.135235	0.130129	0.127595	0.124485	0.127073	0.140660	0.135146	0.138013
Cred. Intervals (5/95)	1.762/1.165	0.969/0.432	1.293/0.797	0.797/0.270	0.884/0.374	0.923/0.420	1.481/0.987	0.824/0.327	0.319/-0.234	0.482/-0.043	0.459/-0.082
Lambda	0.510504†	0.133581†	0.170872†	0.059237*	0.104226†	0.089117†	0.215727†	0.070538*	0.015669	0.045640	0.046051
SD	0.029400	0.046406	0.046609	0.044965	0.043221	0.044742	0.046319	0.048543	0.055380	0.050852	0.042251
Note: † indicates significance at the 0.01 level, ‡ indicates significance at the 0.05 level, and * indicates significance at the 0.10 level											

Table 7: Spatial Autoregressive Probit Model

	Food and tobacco	Clothes and leather	Wood and furniture	Printing and paper	Chemistry	Other non metallic minerals	First transf. or metals	Machinery	Computers and office equipment	Electric and electronic equipment	Transport equipment
Constant	-1.6122†	-2.2200†	-2.4902†	-2.6662†	-2.4177†	-2.4131†	-2.3702†	-2.5687†	-2.6065†	-2.6367†	-2.6617†
SD	(0.0857)	(0.1071)	(0.1057)	(0.1332)	(0.1240)	(0.1173)	(0.0979)	(0.1298)	(0.1698)	(0.1591)	(0.1486)
Credible Interval (5)	-1.7545	-2.3966	-2.6664	-2.8879	-2.6234	-2.6089	-2.5315	-2.7875	-2.8933	-2.8953	-2.9075
Credible Interval (95)	-1.4722	-2.0440	-2.3194	-2.4487	-2.2142	-2.2223	-2.2094	-2.3623	-2.3300	-2.3747	-2.4213
Human Capital	1.0752†	0.9638†	2.1177†	2.0409†	1.8413†	1.3728†	2.0075†	2.1496†	1.2581†	1.6830†	1.8459†
SD	(0.2295)	(0.2577)	(0.2527)	(0.2934)	(0.2744)	(0.2716)	(0.2429)	(0.2935)	(0.3770)	(0.3348)	(0.3196)
Credible Interval (5)	0.7070	0.5353	1.7100	1.5638	1.3814	0.9268	1.6102	1.6689	0.6371	1.1208	1.3228
Credible Interval (95)	1.4572	1.3827	2.5433	2.5264	2.2946	1.8194	2.4096	2.6293	1.8794	2.2294	2.3660
Loc Quotient	0.0019	0.1922†	0.0604†	0.1159†	0.1420†	0.0885†	0.0386†	0.0183†	0.0331†	0.0801†	0.2255†
SD	(0.0021)	(0.0139)	(0.0088)	(0.0164)	(0.0219)	(0.0085)	(0.0094)	(0.0096)	(0.0123)	(0.0177)	(0.0301)
Credible Interval (5)	-0.0015	0.1695	0.0460	0.0894	0.1085	0.0746	0.0243	0.0069	0.0123	0.0518	0.1760
Credible Interval (95)	0.0052	0.2153	0.0751	0.1433	0.1795	0.1023	0.0545	0.0356	0.0525	0.1102	0.2750
MunGDP	0.0056†	0.0069†	0.0042†	0.0111†	0.0064†	0.0072†	0.0036†	0.0105†	0.0089†	0.0093†	0.0076†
SD	(0.0021)	(0.0017)	(0.0022)	(0.0020)	(0.0017)	(0.0018)	(0.0022)	(0.0019)	(0.0015)	(0.0017)	(0.0015)
Credible Interval (5)	0.0019	0.0043	0.0008	0.0080	0.0037	0.0042	0.0006	0.0075	0.0066	0.0067	0.0052
Credible Interval (95)	0.0090	0.0097	0.0078	0.0145	0.0094	0.0103	0.0075	0.0137	0.0115	0.0123	0.0103
Diversity Index	1.2344†	1.2006†	1.6592†	1.0554†	1.1138†	1.2985†	1.7150†	1.0586†	0.2775†	0.6803†	0.7123†
SD	(0.0685)	(0.0773)	(0.0795)	(0.0791)	(0.0786)	(0.0786)	(0.0775)	(0.0767)	(0.0897)	(0.0863)	(0.0797)
Credible Interval (5)	1.1231	1.0756	1.5304	0.9244	0.9849	1.1699	1.5894	0.9343	0.1323	0.5359	0.5819
Credible Interval (95)	1.3502	1.3289	1.7913	1.1871	1.2454	1.4308	1.8444	1.1841	0.4269	0.8212	0.8444
Rho	0.5668†	0.3641†	0.3571†	0.2788†	0.3766†	0.3006†	0.4068†	0.3215†	0.0903‡	0.2075†	0.1940†
SD	(0.0212)	(0.0271)	(0.0244)	(0.0325)	(0.0286)	(0.0304)	(0.0227)	(0.0315)	(0.0470)	(0.0405)	(0.0408)

Note: † indicates significance at the 0.01 level, ‡ indicates significance at the 0.05 level, and * indicates significance at the 0.10 level.

Table 8 Total Effects									
		Lower 0.05	Posterior Mean	Upper 0.95			Lower 0.05	Posterior Mean	Upper 0.95
Food and tobacco	Human Capital	0.2786	0.4099†	0.5347	First transf. of metals	Human Capital	0.5030	0.6316†	0.7506
	Loc Quotient	-0.0007	0.0006	0.0018		Loc Quotient	0.0059	0.0106†	0.0159
	MunGDP	0.0010	0.0019†	0.0028		MunGDP	0.0001	0.0012*	0.0024
	Diversity Index	0.3876	0.4237†	0.4571		Diversity Index	0.4604	0.4995†	0.5387
Clothes and leather	Human Capital	0.1209	0.2154†	0.3124	Machinery	Human Capital	0.2579	0.3424†	0.4291
	Loc Quotient	0.0323	0.0374†	0.0424		Loc Quotient	0.0007	0.0022‡	0.0049
	MunGDP	0.0008	0.0014†	0.0021		MunGDP	0.0012	0.0017†	0.0022
	Diversity Index	0.2049	0.2332†	0.2613		Diversity Index	0.1235	0.1466†	0.1694
Wood and furniture	Human Capital	0.4928	0.6202†	0.7491	Computers and office equipment	Human Capital	0.0370	0.0878†	0.1310
	Loc Quotient	0.0112	0.0158†	0.0198		Loc Quotient	0.0003	0.0024*	0.0043
	MunGDP	0.0002	0.0014†	0.0026		MunGDP	0.0005	0.0007†	0.0009
	Diversity Index	0.4178	0.4556†	0.4918		Diversity Index	0.0054	0.0196†	0.0332
Printing and paper	Human Capital	0.2227	0.3027†	0.3832	Electric and electronic equipment	Human Capital	0.0886	0.1552†	0.2231
	Loc Quotient	0.0114	0.0157†	0.0203		Loc Quotient	0.0040	0.0075†	0.0107
	MunGDP	0.0012	0.0016†	0.0022		MunGDP	0.0007	0.0009†	0.0013
	Diversity Index	0.1126	0.1352†	0.1578		Diversity Index	0.0456	0.0620†	0.0792
Chemistry	Human Capital	0.2350	0.3286†	0.4146	Transport equipment	Human Capital	0.0975	0.1636†	0.2265
	Loc Quotient	0.0164	0.0228†	0.0293		Loc Quotient	0.0041	0.0073†	0.0108
	MunGDP	0.0007	0.0011†	0.0016		MunGDP	0.0006	0.0009†	0.0012
	Diversity Index	0.1399	0.1642†	0.1888		Diversity Index	0.0442	0.0601†	0.0783
Other non metallic minerals	Human Capital	0.1699	0.2740†	0.3721	Note: † indicates significance at the 0.01 level, ‡ indicates significance at the 0.05 level, and * indicates significance at the 0.10 level.				
	Loc Quotient	0.0131	0.0159†	0.0187					
	MunGDP	0.0009	0.0015†	0.0023					
	Diversity Index	0.2121	0.2402†	0.2678					

Tabel 9 Direct Effects

		Lower 0.05	Posterior Mean	Upper 0.95			Lower 0.05	Posterior Mean	Upper 0.95
Food and tobacco	Human Capital	0.2283	0.3347†	0.4369	First transf. of metals	Human Capital	0.4267	0.5386†	0.6423
	Loc Quotient	-0.0006	0.0005	0.0014		Loc Quotient	0.0051	0.0091†	0.0135
	MunGDP	0.0008	0.0016†	0.0023		MunGDP	0.0001	0.0010*	0.0021
	Diversity Index	0.3137	0.3461†	0.3762		Diversity Index	0.3903	0.4260†	0.4577
Clothes and leather	Human Capital	0.1091	0.1949†	0.2840	Machinery	Human Capital	0.2381	0.3158†	0.3969
	Loc Quotient	0.0290	0.0339†	0.0388		Loc Quotient	0.0007	0.0020*	0.0045
	MunGDP	0.0007	0.0013†	0.0019		MunGDP	0.0011	0.0015†	0.0020
	Diversity Index	0.1836	0.2110†	0.2392		Diversity Index	0.1120	0.1352†	0.1584
Wood and furniture	Human Capital	0.4276	0.5389†	0.6528	Computers and office equipment	Human Capital	0.0359	0.0866†	0.1306
	Loc Quotient	0.0098	0.0138†	0.0172		Loc Quotient	0.0003	0.0024‡	0.0042
	MunGDP	0.0002	0.0012†	0.0023		MunGDP	0.0005	0.0007†	0.0009
	Diversity Index	0.3618	0.3958†	0.4319		Diversity Index	0.0053	0.0194†	0.0325
Printing and paper	Human Capital	0.2023	0.2845†	0.3620	Electric and electronic equipment	Human Capital	0.0848	0.1502†	0.2173
	Loc Quotient	0.0108	0.0147†	0.0191		Loc Quotient	0.0039	0.0073†	0.0104
	MunGDP	0.0011	0.0015†	0.0021		MunGDP	0.0006	0.0009†	0.0012
	Diversity Index	0.1035	0.1270†	0.1486		Diversity Index	0.0435	0.0600†	0.0775
Chemistry	Human Capital	0.2178	0.3042†	0.3880	Transport equipment	Human Capital	0.0963	0.1582†	0.2207
	Loc Quotient	0.0151	0.0211†	0.0274		Loc Quotient	0.0039	0.0071†	0.0104
	MunGDP	0.0007	0.0010†	0.0015		MunGDP	0.0006	0.0009†	0.0012
	Diversity Index	0.1281	0.1521†	0.1766		Diversity Index	0.0420	0.0581†	0.0769
Other non metallic minerals	Human Capital	0.1571	0.2526†	0.3439	Note: † indicates significance at the 0.01 level, ‡ indicates significance at the 0.05 level, and * indicates significance at the 0.10 level.				
	Loc Quotient	0.0119	0.0147†	0.0173					
	MunGDP	0.0008	0.0014†	0.0021					
	Diversity Index	0.1950	0.2215†	0.2491					

Table 10 Indirect Effects

		Lower 0.05	Posterior Mean	Upper 0.95			Lower 0.05	Posterior Mean	Upper 0.95
Food and tobacco	Human Capital	0.0484	0.0752†	0.1039	First transf. of metals	Human Capital	0.0677	0.0930†	0.1205
	Loc Quotient	-0.0001	0.0001	0.0003		Loc Quotient	0.0008	0.0016†	0.0024
	MunGDP	0.0002	0.0004†	0.0005		MunGDP	0.0000	0.0002*	0.0004
	Diversity Index	0.0636	0.0777†	0.0922		Diversity Index	0.0572	0.0735†	0.0902
Clothes and leather	Human Capital	0.0105	0.0205†	0.0326	Machinery	Human Capital	0.0153	0.0265†	0.0402
	Loc Quotient	0.0024	0.0036†	0.0049		Loc Quotient	0.0001	0.0002*	0.0004
	MunGDP	0.0001	0.0001†	0.0002		MunGDP	0.0001	0.0001†	0.0002
	Diversity Index	0.0148	0.0222†	0.0299		Diversity Index	0.0066	0.0113†	0.0162
Wood and furniture	Human Capital	0.0574	0.0814†	0.1085	Computers and office equipment	Human Capital	-0.0018	0.0012	0.0044
	Loc Quotient	0.0013	0.0021†	0.0029		Loc Quotient	-0.0001	0.0000	0.0001
	MunGDP	0.0000	0.0002‡	0.0003		MunGDP	0.0000	0.0000	0.0000
	Diversity Index	0.0442	0.0597†	0.0750		Diversity Index	-0.0004	0.0003	0.0010
Printing and paper	Human Capital	0.0077	0.0182†	0.0290	Electric and electronic equipment	Human Capital	-0.0007	0.0050*	0.0109
	Loc Quotient	0.0004	0.0009†	0.0016		Loc Quotient	0.0000	0.0002*	0.0005
	MunGDP	0.0000	0.0001†	0.0002		MunGDP	0.0000	0.0000*	0.0001
	Diversity Index	0.0036	0.0081†	0.0126		Diversity Index	-0.0003	0.0020*	0.0040
Chemistry	Human Capital	0.0127	0.0243†	0.0376	Transport equipment	Human Capital	-0.0002	0.0054*	0.0117
	Loc Quotient	0.0009	0.0017†	0.0026		Loc Quotient	0.0000	0.0002*	0.0005
	MunGDP	0.0000	0.0001†	0.0001		MunGDP	0.0000	0.0000*	0.0001
	Diversity Index	0.0065	0.0121†	0.0179		Diversity Index	-0.0001	0.0020*	0.0040
Other non metallic minerals	Human Capital	0.0103	0.0214†	0.0334					
	Loc Quotient	0.0007	0.0012†	0.0018					
	MunGDP	0.0001	0.0001†	0.0002					
	Diversity Index	0.0112	0.0188†	0.0263					

Note: † indicates significance at the 0.01 level, ‡ indicates significance at the 0.05 level, and * indicates significance at the 0.10 level.

4 Conclusions

This paper is focused on this geographical scope of agglomeration economies in Spain, using data from municipalities. Specifically, on the role of the neighbouring municipalities characteristics in location decisions. Exploratory analysis has shown that for every manufacturing industry considered births are spatially autocorrelated, no matter that we test positive location decisions or the number of births. That is, municipalities which have been chosen as location for births in a given industry tend to be neighbours of municipalities which have also been chosen as location for the same manufacturing industry. Spatial exploratory analysis on the municipality specialization suggests that spatial behavior may be due to the existence of Marshallian agglomeration economies that expand beyond the municipality borders, because the location quotient is also spatially autocorrelated for every manufacturing industry. Therefore, the geographical scope of agglomeration economies may play a role in location decision.

In order to test the role of the geographical scope of agglomeration economies in industrial location decisions confirmatory analysis was carried out. A simple location model was outlined and estimated using Spatial Econometrics and Spatial Statistics techniques. Spatial variables are highly significant for most industries, so we cannot reject that the characteristics of neighbouring municipalities matter in industrial location decisions. That is, what happens in a municipality depends not only on what happens inside that municipality, but also depends on what happens in its neighbouring area. Interurban agglomeration economies due to industrial diversity seem to play a larger role in the location decision of neighbouring municipalities than the one of interurban agglomeration economies due to industrial specialization.

Policy makers of countries with a highly decentralized regional system, such is Spain, should bear in mind that these agglomeration economies can extend to or come from neighbouring areas which belong to other regions. Therefore, interregional coordination is needed before implementing local or regional location incentives. This might be an important argument to justify the industrial policy has a regional definition, avoiding either the national basis less efficient (Aghion et al, 2009) and the municipal basis. In fact, most of the variables determining localization (population skill, manufacturing specialization, market potential and diversity) are mainly affected by policies of regional scope.

Future research should check the kilometric extent of agglomeration economies for every industry. Longer in time and more disaggregated industrial datasets (3 or higher digit level) are needed to analyze both the industrial, the temporal and the geographical scopes of agglomeration economies properly.

Finally, spatial autocorrelation should be taken into consideration when estimating location models, since spatial dependence invalidates the use of traditional estimation techniques.

References

Acs ZJ, Audretsch DB, Feldman MP (1992) Real effects of academic research: Comment. Amer Econ Rev 82(1): 363-367

Aghion, P.; Bulanger,J. y Cohen, E.(2011) Rethinking industrial policy", Bruegel, issue, 2011/04, june.

Alañón-Pardo A (2002) Estimación del valor añadido per cápita de los municipios españoles en 1991 mediante técnicas de econometría espacial. Ekonomiaz 51: 172-194

Anselin L (1988) Spatial econometrics, Methods and models. Kluwer Academic, Dordrecht

Anselin L (1992) SpaceStat Tutorial, a book for using SpaceStat in the analysis of spatial data. University of Illinois, Urbana-Champaign

Anselin L (2002) Under the Hood. Issues in the Specification and Interpretation of Spatial Regression Models. Agr Econ 17: 247-267

Arauzo-Carod J (2002) Determinants of Industrial Location. An application for Catalan Municipalities. Working Paper Nº 138, Estudios sobre la Economía Española. FEDEA. Madrid

Arauzo-Carod JM, Liviano D, Manjón M (2010) Empirical Studies in Industrial Location: An Assessment of their Methods and Results. J Reg Sci 50(3): 685-711

Arita T, McCann P (2000) Industrial alliances and firm location behaviour: some evidence from the US semiconductor industry. Appl Econ 32: 1391-1403

Arrow, KJ (1962) Economic welfare and the allocation of resources for invention. In Nelson RR (ed) The Rate and Direction of Inventive Activity, Princeton-University Press, Princeton, pp 609- 626

Audretsch D, Feldman M (1996) R&D Spillovers and the geography of innovation and production. Amer Econ Rev 86: 630-640

Autant-Bernard C (2006) Where Do Firms Choose to Locate their R&D? A Spatial Conditional Logit Analysis on French Data. European Plan Stud 14 (9): 1187-1208

Bailey TC, Gatrell AC (1995) Interactive Spatial Data Analysis, 2 nd edn. Longman, London

Bartlesman EJ, Caballero RJ, Lyons RK (1994) Customer and supplier-driven externalities. Amer Econ Rev 84(4): 1075-1084

Baumgartner J R (1988) Physicians' services and the division of labor across local markets. J. Polit. Economy 96(5): 948-982

Carlton D (1983) The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. Rev Econ Statist 65: 440–449

Cliff A, Ord K (1980) Spatial processes: models & applications. Pion, London

Diamond CA, Simon CJ (1990) Industrial Specialization and the Returns to Labor. J Lab Econ 8(2): 175-201

Costa DL, Kahn ME (2001) Power Couples. Quart J Econ. 116:1287-1315

Duranton G, Puga D (2000) Diversity and specialisation in cities, why, where and when does it matter? Urban Stud. 37: 533-555

Duranton G, Puga D (2005) From sectoral to functional urban specialization. J Urban Econ 57(2): 343-370

Ellison G, Glaeser E (1997) Geographic Concentration in US Manufacturing Industries, A Dartboard Approach. J Pol Economy 105: 889-927

Ellison G, Glaeser EL, Kerr WR (2010) What causes industry agglomeration? Evidence from coagglomeration patterns. Amer Econ Rev 100 (3): 1195-1213

Figueiredo O, Guimarães P, Woodward D (2002) Home-field advantage: location decisions of Portuguese entrepreneurs. J Urban Econ 52: 341-361

Fischer MM, Griffith DA (2008) Modeling spatial autocorrelation in spatial interaction data: an application to patent citation data in the European Union. J Reg Sci 48 (5): 969–989

Fleming M (2004) Techniques for estimating spatially dependent discrete choice models. Iin Anselin L, Florax R, Rey S (eds) Advances in Spatial Econometrics, Methodology, Tools and Applications, Springer, Berlin, pp 145-167

Glaeser E, Kallal H, Scheinkman J, Shleifer A (1992) Growth in cities. J Polit Economy 100: 1126-1152

Guimarães P, Figueiredo O, Woodward D (2000) Agglomeration and the Location of Foreign Direct Investment in Portugal. J Urban Econ 47: 115-135

Guimarães P, Figueiredo O, Woodward D (2004) Industrial location modeling: extending the random utility framework. J Reg Sci 44: 1-20

Hayter R (1997) The dynamics of industrial location, The factory, the firm and the production system, Wiley, New York

Henderson, JV (1997) Externalities and Industrial Development. J Urban Econ 42: 449-470
Henderson, JV (2003) Marshall's Scale Economies. J Urban Econ 53: 1-28

Holl A (2004a) Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain. Reg Sci Urban Econ 34: 341-364

Holl A (2004b) Transport infrastructure, agglomeration economies, and firm birth. Empirical evidence from Portugal. J Reg Sci 44: 693-712

Holmes, T (1999) Localization of industry and vertical disintegration. Rev Econ Statist 81: 314–25
Hoover E (1948) The location of Economic Activity. McGraw-Hill, New York

Jacobs J (1969) The Economy of Cities. Vintage, New York

Jaffee, AB (1989) Real effects of academic research. Amer Econ Rev 79(5): 957-70

Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. Quart J Econ 108: 577-598

Jofré-Montseny, J (2009) The scope of agglomeration economies: Evidence from Catalonia. Pap Reg Sci 88(3): 575-590

Kaiser M, Cressie N (1997) Modeling Poisson variables with positive spatial dependence. Statist Probab Lett 35: 423-432

Krugman P (1991a) Geography and trade. MIT Press, Cambridge (Mass)

Krugman P (1991b) Increasing returns and Economic Geography. J Polit Economy 99: 484-99
LeSage, JP, R. Pace RK, Lam N, Campanella R, Liu X. (2011) New Orleans business recovery in the aftermath of Hurricane Katrina. Journal of the Royal Statistical Society: Series A 174 (4): 1007-1027.

Lesage JP, Fischer MM, Scherngell T (2007) Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects. Pap Reg Sci 86 (3): 393-422

Lesage J (1997) Bayesian Estimation of Spatial Autoregressive Models. Int Reg Sci Rev 20: 19-35
Lesage J (2000) Bayesian estimation of limited dependent variable spatial autoregressive models. Geographical Analysis 32: 19–35

Marshall A (1890) Principles of Economics, ed of 1920. MacMillan, London

McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In Zarembka P (ed) Frontiers in Econometrics. Academic Press, New York, pp 105-142

Mcmillen D (1995) Spatial effects in Probit models: a Monte Carlo investigation. In Anselin L, Florax R (eds) New Directions in Spatial Econometrics. Springer, Berlin, pp 189–228

Miller HJ (2004) Tobler's First Law and Spatial Analysis. Annals of the Association of American Geographers 94(2): 284-289

Mompó A, Monfort V (1989) El Registro Industrial como fuente estadística regional: el caso de la Comunidad Valenciana. Econ Ind 268: 129-140

Moretti E (2000) Estimating the social return to higher education, NBER Working Paper No 9018

Ottaviano G, Puga D (1998) Agglomeration in the Global Economy: A survey of the New Economic Geography. World Economy 21(6): 707-731

Pablo-Martí F, Muñoz-Yebra C (2009) Localización empresarial y economías de aglomeración: el debate en torno a la agregación espacial. Investigaciones Regionales 15: 139 -166

Pinkse J, Slade M (1988) Contracting in Space: An Application of Spatial Statistics to Discrete-Choice Models. J Econometrics 85: 125-154

Romer P (1986) Increasing Returns and Long Run Growth. J Polit Economy 94: 1002-1037

Rosenthal S, Strange W (2003) Geography, industrial organization and agglomeration. Rev Econ Stat 85: 377-393

Rosenthal S, Strange W (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson J, Thisse JF (eds) Handbook of Urban and Regional Economics, vol. 4. Elsevier, Amsterdam, pp 2119-2171

Simmie J (1998) Reasons for the Development of 'Islands of Innovation': Evidence from Hertfordshire. Urban Stud 35: 1261-1289

Smith T, Lesage J (2002) A Bayesian Probit Model with Spatial Dependencies. Webpage of Smith T at the University of Pennsylvania, School of Engineering and Applied Science, http://www.seas.upenn.edu/~tesmith/sprobit.pdf. Accessed 15 July 2011

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64: 583–639. doi: 10.1111/1467-9868.00353

Suárez-Villa L, Alrod W (1998) Operational Strategy, R&D and Intrametropolitan Clustering in a Polycentric Structure. The advanced Electronics Industries of the Los Angeles Basin. Urban Stud 34: 1343-80

Tobler WR (1979) A computer movie simulating urban growth in the Detroit region. Econ Geogr 46: 234–40

Unwin DJ (1996) GIS, spatial analysis and spatial statistics. Prog Hum Geography 20(4): 540-551

Van Soest D, Gerking S, Van Oort F (2006) Spatial impacts of agglomeration externalities. J Reg Sci 46: 881-899

Viladecans-Marsal E (2001) La concentración territorial de las empresas industriales: un estudio sobre el tamaño de las empresas y su proximidad geográfica. Pap Econ Española 89-90: 308-321

Viladecans-Marsal E (2003) Economías externas y localización del empleo industrial. Revista Econ Aplicada 31: 5-32

Viladecans-Marsal E (2004) Agglomeration economies and industrial location: city-level evidence. J Econ Geography 4: 562-582

Appendix 1 Descriptive Statistics

Table 11 Manufacturing Establishments creation (1980-1998) (1/2)

		1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Food,	Mun	528	556	562	692	687	767	760	771	745	726	678	669	577	600	646	663	624	387	458
drinks &	Esta	890	895	935	1328	1290	1490	1355	1465	1331	1458	1221	1336	1105	1133	1383	1331	1191	698	997
tobbacco	Max	44	37	40	61	80	72	38	54	41	106	40	78	57	58	90	84	59	30	51
-Clothes &	Mun	323	289	349	475	401	457	480	589	578	532	509	508	429	334	308	351	323	241	193
leather	Esta	656	605	856	1151	1039	1462	1597	1938	1693	1484	1250	1225	1013	731	762	826	807	644	544
	Max	39	33	73	69	53	182	203	166	94	116	70	60	61	40	42	47	46	50	51
Wood &	Mun	573	561	585	716	655	733	724	784	810	750	794	790	712	623	530	655	564	349	264
furniture	Esta	1093	1095	1110	1475	1372	1469	1588	1708	1734	1808	1701	1639	1761	1186	943	1216	1019	639	497
	Max	40	44	48	50	45	48	57	56	61	85	49	54	331	49	36	25	23	24	17
Paper &	Mun	120	129	161	185	164	192	210	237	251	280	274	275	256	242	189	251	225	164	115
printing	Esta	228	233	393	405	371	434	514	607	590	798	577	586	517	433	438	519	424	334	248
	Max	55	30	79	58	59	58	86	89	90	169	55	64	46	37	64	50	49	27	34
Chemistry	Mun	251	256	265	317	305	329	315	352	315	345	332	323	301	267	222	282	277	201	144
	Esta	414	406	457	587	557	665	597	719	600	619	614	545	476	434	368	453	440	345	229
	Max	13	20	19	36	29	27	37	41	18	23	17	15	16	23	24	10	10	13	12
Other non	Mun	289	265	221	298	231	270	283	350	338	366	370	423	330	311	257	305	282	188	143
metallic	Esta	415	374	303	420	346	402	409	538	509	569	575	633	449	393	338	433	403	259	206
minerals	Max	22	14	9	14	17	21	12	16	20	17	20	24	10	6	8	12	9	7	8
First transf.	Mun	702	686	623	733	641	695	712	750	830	795	814	855	775	672	596	698	623	372	280
Metals	Esta	1282	1266	1402	1599	1329	1486	1643	1851	1969	2036	1946	2004	1624	1298	1142	1489	1233	805	617
	Max	42	53	71	73	67	47	50	62	59	78	65	55	25	24	22	31	20	19	24

Mun: number of municipalities in which manufacturing establishments were created; Esta: establishments created in all municipalities; Max: maximum number of establishments created in a municipality

Table 12 Basic Summary Statistics for Location quotients (LQ) 1990											
	Food, drinks ..	Clothes & leather	Wood & furniture	Paper & printing	Chemistry	Other non metallic min.	First transf. of Metals	Machinery	Computer & office equip.	Electric & Electron. equip.	Transport equip.
Mean	2.10	0.69	1.33	0.24	0.26	0.79	1.09	0.78	0.16	0.21	0.12
Maximum	723.00	53.45	63.03	31.65	58.36	113.51	163.86	601.72	55.69	185.87	10.87
Std. Dev.	11.72	1.74	2.47	1.29	1.45	2.76	3.28	7.32	1.79	2.71	0.70

Table 13 Basic Summary Statistics for Common Regression Variables			
	Human Capital Index (HC)	Diversity Index (DI)	Municipality Product (MP)
Mean	0.28	0.04	6140.80
Maximum	0.94	6.48	5457229.69
Std. Dev.	0.09	0.08	78403.48