

# SPATIAL EFFECTS IN INDUSTRIAL LOCATION CHOICES: INDUSTRY CHARACTERISTICS AND URBAN ACCESSIBILITY

ANDRÉS ARTAL-TUR\*, JOSÉ MIGUEL NAVARRO-AZORÍN\*,  
MARÍA LUISA ALAMÁ-SABATER\*\* & ANTONIO JUAN BRIONES-PEÑALVER\*\*\*

\**Technical University of Cartagena, Department of Economics, C\ Real 3, 30201 Cartagena, Spain. E-mails: Andres.artal@upct.es; JMiguel.Navarro@upct.es*

\*\**Universitat Jaume I, Department of Economics, Campus Riu Sec E-12071, Castellón de la Plana, Spain. E-mail: alama@eco.uji.es*

\*\*\**Technical University of Cartagena, Department of Business Administration, C\ Real 3, 30201 Cartagena, Spain. E-mail: Aj.briones@upct.es*

Received: October 2011; accepted November 2012

---

## ABSTRACT

In this paper we study how neighbourhood-related spillovers affect location choices of manufacturing firms at a local level. A spatial Dirichlet-multinomial regression model is applied to 90,000 new establishments of the Spanish Mediterranean Axis. Empirical findings show that spatial spillovers play an important role, together with traditional explanatory factors, in driving decisions of companies. Their size and scope depends on two main issues, the specific characteristics of the manufacturing industry the firm belongs to, and the accessibility of the urban environment where the firm is located.

**Key words:** Location, spatial spillovers, industry characteristics, urban accessibility

---

## INTRODUCTION

External economies have played a central role in economic theory since the works of Alfred Marshall. Today, the study of how spatial effects determine location decisions of people and firms constitutes one of the most interesting branches of research in regional studies (Glaeser 2007). However, and despite the generalization of the theoretical concept of spillovers in economics, the empirical measurement of such variables is still an open issue (Burger *et al.* 2010). Departing from traditional location models, recent contributions building on spatial econometrics methods have introduced the role of neighbourhood-related effects in shaping firm's

decisions (Le Sage & Pace 2009). In the industrial location field, and following the pioneer work of Rosenthal & Strange (2003), some papers have start dealing with this type of spatial effects. In doing so, they focus on extending the conditional logit model relying on random utility maximisation problems (Autant-Bernard 2006; Woodward *et al.* 2006; Jofre-Monseny 2009; Alamá *et al.* 2010). Autant-Bernard (2006) studies how spatial spillovers influence the location of R&D laboratories in NUTS 2 regions of France. Woodward *et al.* (2006) analyse the choices of high-technology firms in US counties by employing spatially weighted explanatory variables. Jofre-Monseny (2009) focuses on the spatial scope of agglomeration economies

along the Spanish region of Catalonia. Finally, Alamá *et al.* (2010) also build on such spatial framework for explaining location decisions of manufacturing firms in the Spanish region of Murcia. All four contributions are seminal in this literature, and propose ways of progressively improving the way we cope with spatial spillovers in location choices.

The present paper continues developing such an approach, with a deeper insight on the factors determining location choices of firms in the presence of neighbourhood-related spatial effects. As main novelties, we first provide a quantitative measure of the spatial effects spilling over local boundaries. We define this measure in relative terms to the magnitude of traditional externalities arising at a local level. In this way, we can show the relevance of spatial effects in the process of location of firms. Second, we identify the main factors driving spatial effects in the presence of local shocks. In this sense, they appear to be related to the type of industry the firm belongs to, as well as to accessibility issues characterising the municipality where the firm is expected to locate. After this introduction, we present the analytic framework employed in the study. In the third section we discuss the choice of the set of explanatory variables, run out our empirical routines and discuss main results. Finally, the fourth section concludes.

## MODEL SETTING

This section introduces a location model based on the standard that the firm will choose the municipality with the highest expected profit among several alternatives.<sup>1</sup> From the point of view of a firm  $i$  which operates in a particular industry  $s$ , each municipality  $j$  in the set of  $J$  possible locations offers an expected profit of  $\pi_{ij}$  such that:

$$\pi_{ij} = X_j\beta + \delta WX_j\beta + \eta_j + \varepsilon_{ij}, \quad (1)$$

where the variables in  $X_j$  include characteristics of the municipality affecting the location decisions of firms in all industries (population, accessibility, availability of skilled labour force, etc.), and in a particular industry  $s$  (i.e., the specialisation index of the municipality in that industry);  $WX_j$  represents the spatially weighted

average of the characteristics of neighbouring municipalities also affecting the probability of a firm to be located in  $j$ ;<sup>2</sup>  $\eta_j$  is a random variable capturing unobserved factors specific to each location;<sup>3</sup> and finally,  $\varepsilon_{ij}$  is a random term for other unobservable factors which determine the expected profits of firm  $i$  from locating in municipality  $j$ .

The specification in equation (1) implies that the expected profit from establishing in a given location  $j$  not only depends on the advantages offered by this own location, but on those ones spilling over from its neighbourhood. The spatially lagged term in equation (1) is composed by two parts: the first one reflects the existing stock of endowments at nearby municipalities (pool of qualified workers, institutions, etc.), captured by the term  $WX_j\beta$ ; the second one represents the truly spatial effects spilling over from the neighbourhood to municipality  $j$ , captured by parameter  $\delta$ . Following this specification, the higher the value of  $\delta$ , the more important the spatial effects, while the closer the neighbouring localities, as defined in  $W$  matrix, the bigger these effects too.

As it is common in a discrete choice framework, the firm will choose to locate the plant in the (expected) most profitable location, the municipality  $j$  in this case. Hence, the probability of choosing location  $j$  accomplishes:

$$\Pr(\pi_{ij} > \pi_{ik}), \quad \text{for } j \neq k, \text{ and } j, k = 1, 2, \dots, J, \quad (2)$$

and it can be shown that if the error term  $\varepsilon_{ij}$  is independent and identically distributed (i.i.d.) according to a type I extreme value distribution, the probability that a firm chooses municipality  $j$  conditional on the  $\eta$  random effects is:

$$P_j = \frac{\exp\{X_j\beta + \delta WX_j\beta\}}{\sum_{k=1}^J \exp\{X_k\beta + \delta WX_k\beta\}}. \quad (3)$$

The main focus of the paper would be then twofold. First, we will wonder if spatial effects play a role for location choices of companies. And second, this being the case, we will concentrate in analysing the main factors influencing such spatial effects. Pursuing these objectives, and to define an operational definition of the spatial spillover measure, we consider the cross-derivative term:

$$\frac{dP_k}{dX_j} = \frac{\partial P_k}{\partial X_j} + \sum_{r \neq j} \frac{dWX_r}{dX_j} \frac{\partial P_k}{\partial WX_r}, \quad k \neq j, \quad (4)$$

which shows how a marginal change in any covariate of the municipality  $j$  affects the expected probability of the nearby municipality  $k$  in attracting new firms of a particular industry ( $P_k$ ). From equation (4) it can be distinguished two types of impacts on  $P_k$ : the first term is the direct impact of one shock in  $X_j$  on that probability; the second one summarises the indirect impact of the shock driven by changes in the neighbourhood characteristics (captured by  $WX_r$ ). In this way, the cross-derivative term in equation (4) allows us to define the appropriate measure of the spatial spillovers we want to identify, denoted as  $SSE_{j \rightarrow k}$ . From equations (1) and (3), it follows that:

$$SSE_{j \rightarrow k} = \delta P_k \left( w_{kj} - \sum_{r \neq j} w_{rj} P_r \right) \beta, \quad k \neq j. \quad (5)$$

As shown in equation (5), the intensity of the spatial spillover between municipalities  $j$  and  $k$  depends on four key elements. First, on the value of parameter  $\delta$ , that measures the relevance in firms' profits of spatial spillovers. Second, on the strength of the neighbouring relationship between both municipalities, as defined by the elements of the spatial weight matrix  $W$ . Closer neighbours to location  $k$  would be characterised by higher values of the corresponding element of the  $k$ -th row of  $W$  and, consequently, the term in parenthesis in equation (5) would be also higher.<sup>4</sup> Third, the magnitude of the spatial spillover is proportional to the probability  $P_k$ . In this sense, the industrial structure of municipality  $k$  is also determining how it can benefit, or absorb, the emerging spatial effects. If the shock is industry-neutral this is not a matter of fact, but if the shock is industry-specific, the degree of specialisation of municipality  $k$  in that particular industry will contribute to determine the magnitude of the spatial effect. If we generalise this fact to the whole neighbourhood of municipality  $j$  benefiting from spatial spillovers, we can understand how the existence of clusters, highly specialised in particular industries, reinforce the magnitude of spatial spillovers. Both concepts of 'degree of specialisation of a neighbourhood',

and the previous one of a 'closer neighbourhood', would allow us to define our concept of accessibility of a municipality, that will be a centre piece of the paper. Fourth,  $\beta$  parameter captures the role of endowments in producing spatial spillovers. For computing our spatial effects, and in order of not introducing more complexity in the model, we assume that such parameter shows the same value in the case of municipality  $j$ , than for its nearby municipalities. That is, endowments share the same coefficients in locality  $j$  ( $X_j$ ) and for the (average) nearby municipalities ( $WX_j$ ). In theoretical terms this is a plausible assumption in this type of models, as shown in Anselin (2003), and from an empirical point of view it facilitates quantitative estimation of spatial effects for every industry in the study.<sup>5</sup>

Finally, since equation (5) is just measuring spatial spillovers arising to location  $k$  from a shock in municipality  $j$ , it is interesting to define summary measures of total spatial spillovers emerging from the model. Accordingly, we will integrate the terms  $SSE_{j \rightarrow k}$  over locations  $K - j$ , for a nearby threshold area. In doing so, we rely on the evidence showing that spatial effects dissipate after a certain distance (Arauzo-Carod 2008).<sup>6</sup>

**ESTIMATION RESULTS**

This section is devoted to quantify the spatial spillover effects affecting firms' location choices in the geographical area of analysis. To this end, we start by introducing the data set and the explanatory variables employed in the empirical exercise. Further, estimates of the spatially extended location framework are used to compute the summary indices of spatial spillovers generated by each municipality, then discussing their magnitude and scope.

**Geographical scope of data set: The Spanish Mediterranean Axis** – Our analysis draws on a database for the population of industrial firms newly established in the Spanish Mediterranean Axis (SMA, henceforth) provided by the Central Directory of Enterprises (DIRCE) of INE (National Statistics Institute). The SMA is defined as the territory of the Spanish Mediterranean seaboard stretching from the French frontier to the Straits of Gibraltar, that is,

Table 1. *Classification of economic activities.*

Sector	Economic activities	ISIC
Natural resources	Food products and beverages. Tobacco products	15,16
	Manufacture of paper and paper product	21
Labour intensive	Textile products	17
	Wearing apparel	18
	Dressing of leather	19
	Wood products (except furniture)	20
Product differentiation	Publishing and printing	22
	Manufacture of other non-metallic mineral products	26
	Manufacture of basic metals	27
	Manufacture of fabricated metal products, except machinery and equipment	28
	N.E.C. machinery and equipment	29
	N.E.C. electrical machinery apparatus	31
Scale economies	Manufacture of chemical industry	24
	Manufacture of rubber and plastics products	25
	Manufacture of motor vehicles and trailers	34
	Other transport equipment	35
Science	Office machinery, computing machinery	30
	Radio, TV and Communication equipment	32
	Medical, precise, optical instruments	33

between the regions of Catalonia and Andalusia. The geographical area making up the SMA accumulates 40 per cent (approximately 19 million inhabitants) of the population of Spain (3.8% of EU 27) in 20 per cent of the country surface (2.2% of EU 27), accounting for 40 per cent of the national GDP (3.7% of EU 27). It makes the area bigger in geographic surface, population and economic activity than many of the EU countries. The data set is defined at the geographical level of municipalities, this being the optimal framework for studying spatial effects (Arauzo-Carod *et al.* 2009). Besides, the SMA is chosen as our area of analysis for two main reasons: first, the familiarity with municipal data sources in Spain allows us to compile all necessary covariates at the local level, a pivotal issue of the study. And second, this territory shows important economic agglomerations, accumulating 50 per cent of the industrial employment of the country, with a remarkable growth of the number of firms (45%) and employment (50%) along the past ten years (IVIE 2010). In this way, it becomes an optimal field of study for our purposes. SMA is mainly characterised by a remarkable contrast between the active and densely populated

seaboard and the rather inhabited inland, with some discontinuities in the urban network. Two cities, Barcelona and Valencia, make the difference in terms of global connectivity, becoming well consolidated urban structures connected to the rest of Europe.

The data set comprises around 90,000 new firms (start-ups) located in 314 municipalities between 1998 and 2008. We assume that the role played by the factors determining the location choice of the firm may vary according to the characteristics of the industry that it belongs to. Consequently, we decided to split out the data set into five sub-samples, corresponding to five broadly defined sectors on the basis of the factors affecting the location process (OECD 1987). The five categories reached are those of natural resource-based NR (e.g. food industry), labour intensive LI (e.g. textiles), product differentiated PD (e.g. publishing and printing), scale-based industries SE (e.g. motor vehicles and trailers), and science-based industries R&D (e.g. computing machinery; medical and precision instruments). The classification scheme for industries is shown in Table 1. The spatial distribution of firms in the SMA at a municipal level is depicted in

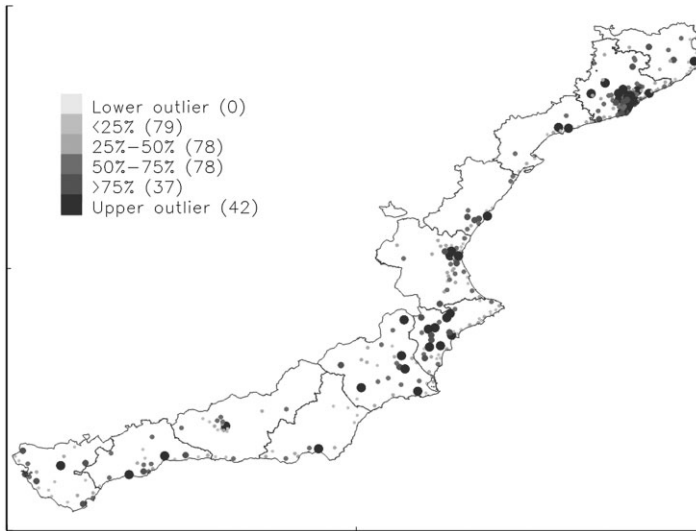


Figure 1. *Geographical distribution of firms in the dataset.*

Figure 1, showing big concentrations in the urban metropolitan areas of Barcelona and Valencia, and along the coastal corridor.

**Model estimates** – The present investigation is not another location exercise for industrial establishments, just being addressed to quantify the role of spatial spillovers in such processes. Subsequently, the empirical model builds on a set of explanatory variables that have become standard in this type of studies (see e.g., Arauzo-Carod *et al.* 2010). Basically we move to ensure correct behaviour for the covariates, resulting in robust estimates of the spatial effects of the model. The definition and data sources of the covariates are summarised in Table 2.

Our dependent variable is defined as the location of new firms (start-ups) in a given municipality for the years 1998 to 2008. The set of explanatory variables starts by including a size control variable as the municipality area (AREA), and the altitude of each locality over the sea level (ELEVATION) to tackle with some accessibility issues. Infrastructures are accounted for by distances between municipalities to the nearest ports (PORTS) and airports (AIRPORTS), together with new kilometres of roads and highways built since the accession to the European Union (ROADS). Workers per squared kilometre in the urban core of the municipality act as a measure of urbanisation

economies (URBANISATION), while local population measure is expected to capture demand potential (POPULATION). We approach the human capital variable (EDUCATION) by employing the share of labour force with higher education (secondary and tertiary studies finished). The design of the Spanish infrastructure networks clearly favours urban centres at the provincial level, so locations close to administrative centres would benefit from better accessibility, this is addressed by the variable DISTANCE TO CENTRE. A set of knowledge-related variables in the model includes PATENTS, ICT INFRASTRUCTURES and R&D EXPENDITURES, all of them at the level of province, this being the more disaggregated geographical level we have access to in Spain. Finally, we also apply two traditional agglomeration economies of location models. First, we employ the inverse of the Herfindahl index for capturing the effects of a richer environment in the supply of economic activities (DIVERSITYINDEX). And second, we employ a specialisation index at a local level (MAR-type economies) for each industry  $s$  in the investigation (LOCATION QUOTIENT). In this way, all types of standard explanatory variables in location models are accounted for in our exercise, including geographical measures, infrastructures, agglomeration forces, technological factors, and other supply and demand variables.

Table 2. *Explanatory variables: definition and data sources.*

Variable	Definition	Source
AREA	Municipality area in squared kilometres (log) 1998.	Yearbook of Statistics, INE.
ELEVATION	Height of the municipality over the sea level (log) 1998.	Yearbook of Statistics, INE.
POPULATION	Number of inhabitants in the municipality (log) 1998.	Yearbook of Statistics, INE.
URBANISATION	Total workers per km <sup>2</sup> at the urban core of the municipality (log) 1998.	Own elaboration from Statistics Yearbook, La Caixa.
EDUCATION	Share of labour force with higher education in the municipality (log) 1998.	Own elaboration from Population Census data 1991, and Labour Force Survey EPA1998, INE.
DISTANCE TO CENTRE	Distance of the municipality to administrative centre, in kms. (log) 1998.	Own elaboration from Statistics Yearbook, La Caixa.
PORTS	Distance of the municipality to nearest Port, in kms. (log) 1998.	Own elaboration from Ministry of Fomento.
AIRPORTS	Distance of the municipality to nearest Airport, in kms. (log) 1998.	Own elaboration from Ministry of Fomento.
ROADS	Kms. of new highways + local roads in the municipality since 1986 (log) 1998.	Own elaboration from Ministry of Fomento.
PATENTS	Number of Patents per province (log) 1998.	Statistics of Patents, INE.
ICT INFRASTRUCTURES	ICT stock in million €/person per province (log) 1998.	COTEC Foundation.
R&D EXPENDITURES	R&D expenditures per province in million € (log) 1998.	R&D survey, INE.
DIVERSITY INDEX	Index of diversification computed as the inverse of the Herfindahl Index, defined as $H = \sum_s c_s^2$ where $c_s$ is the number of firms in industry $s$ ( $s=NR, LI, \dots, RD$ ) over total firms established in each municipality 1998.	Own elaboration from DIRCE data, INE.
LOCATION QUOTIENT	Location quotient for defined industrial branches 1998.	Own elaboration from DIRCE data, INE, and Statistics Yearbook, La Caixa.

Note: INE stands for the Spanish Statistics Institute.

We expect all these variables to show a positive coefficient in the regression, except for ELEVATION and infrastructure (distance-based) variables, where we expect negative signs. We also expect obtaining different estimates for every industrial subsample in the study.

Further, the model includes a cross-regressive spatial term ( $WX_j$ ) capturing the (spatially weighted average) covariates of the neighbourhood for a given municipality  $j$ . In order to compute this term, we need to define the spatial weight matrix  $W$ . It is done in terms

of the inverse of the Euclidean distances among municipalities, with a representative term:

$$w_{jk} = \begin{cases} d_{jk}^{-1} / \sum_{k \neq j} d_{jk}^{-1} & \text{if } d_{jk} \leq R \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where  $d_{jk}$  is the Euclidean distance between municipality  $j$  and municipality  $k$ , and  $R$  represents a threshold distance determining the range of action of spatial effects,<sup>7</sup> if present. Note that this definition of the  $W$  matrix implicitly



Table 3. Location choice of firms in the Spanish Mediterranean Axis (SMA): spatial Dirichlet-multinomial regression.

	NR	LI	PD	SE	RD
AREA	0.1422*** (0.0214)	0.0069 (0.0187)	0.0936*** (0.0177)	0.1048*** (0.0266)	0.0111** (0.0057)
ELEVATION	0.0544*** (0.0174)	-0.0029 (0.0159)	-0.0018 (0.0131)	-0.0647** (0.0190)	-0.0152** (0.0082)
POPULATION	0.7707*** (0.0269)	1.0066*** (0.0236)	0.8805*** (0.0209)	0.8220*** (0.0324)	0.9847*** (0.0343)
URBANIZATION	0.1936* (0.0996)	0.3210*** (0.0812)	0.3205*** (0.0695)	0.2613** (0.1335)	0.3932*** (0.1217)
EDUCATION	-0.0080 (0.0576)	-0.1856*** (0.0504)	0.3317*** (0.0469)	0.2508*** (0.0684)	0.3787*** (0.0774)
DISTANCE TO CENTRE	0.0383** (0.0198)	-0.0498*** (0.0186)	-0.0168 (0.0162)	-0.0048 (0.0260)	0.0011 (0.0255)
PORTS	0.0922** (0.0474)	0.0256 (0.0260)	0.1022** (0.0526)	0.1339** (0.0681)	0.1003 (0.1315)
AIRPORTS	0.0256 (0.0210)	0.0024 (0.0026)	0.0023 (0.0018)	0.0589* (0.0328)	0.0052* (0.0028)
ROADS	0.1282*** (0.0398)	0.0752* (0.0397)	0.1420** (0.0727)	0.1121*** (0.0343)	0.1225* (0.0647)
PATENTS	0.0238 (0.0433)	0.0263 (0.0336)	0.0941** (0.0497)	0.1439*** (0.0538)	0.1722*** (0.0889)
ICT INFRASTRUCTURES	0.1224** (0.0626)	0.0631* (0.0335)	0.1302 (0.2723)	0.1937** (0.0996)	0.2115** (0.1095)
R&D EXPENDITURES	0.0223* (0.0119)	0.0752* (0.0396)	0.1424** (0.0731)	0.1107*** (0.0333)	0.0921* (0.0490)
DIVERSITY INDEX	0.8437*** (0.1372)	1.6636*** (0.1440)	1.3723*** (0.1425)	0.9843*** (0.1754)	0.9822*** (0.1947)
LOCATION QUOTIENT	0.2065*** (0.0111)	0.3338*** (0.0098)	0.3762*** (0.0281)	0.3114*** (0.0160)	0.3695*** (0.0629)
$\delta$	0.2171** (0.1177)	0.1675*** (0.0547)	0.2441*** (0.0815)	0.4214*** (0.1271)	0.2972*** (0.0803)
LR test for overdispersion	786.2***	3180.8***	3634.6***	576.7***	121.8***
Overall significance	542.0***	679.0***	739.7***	562.2***	719.8***
Log-likelihood	-1076.1	-1397.21	-1464.4	-1015.64	-733.5

Notes: The dependent variable is location choice by new firms for 1998–2008 in every industrial branch defined. Standard errors in parenthesis.

\*\*\* significant at the 1% level; \*\* significant at the 5% level and \* significant at the 10% level

assumes that spatial effects tend to dissipate as the physical distance between locations increases, eventually disappearing beyond a threshold distance. It must be highlighted that the definition we consider for the *W* matrix constitutes a particular choice, ultimately determining the estimates of the  $\delta$  parameter of the empirical model. In the Appendix we include some additional results for alternative *W* matrices just for comparison purposes. In general we show no greater impact of such an issue on the resulting estimates of the spatial effects.

Table 3 presents the results for the five subsets of industries in the study. Evidence of

the presence of overdispersion in data is assessed by means of a likelihood ratio test, shown at the bottom of the table. These statistics provide strong empirical support to the Dirichlet-multinomial specification against the conditional logit model, so we follow such a modelling strategy.<sup>8</sup> In general, the estimated parameters show the expected signs, with their magnitudes (marginal effects) being in line with those of the literature.<sup>9</sup> Despite employing the spatial lag model specification, it could be possible to find some correlation in the residual term. By relying on a substantive approach, we assume that the spatial effects basically apply to

immediate neighbours, and so 'the proper spatial range of the explanatory variables is constrained to the location and its immediate neighbours (but not beyond)' (Anselin 2003, p. 161). In the same line of reasoning, a recent work of Robertson *et al.* (2009) applying Monte Carlo simulation conclude that the inclusion of spatially lagged explanatory variables is the most effective procedure to capture spatial effects on discrete choice models. Moreover, we have checked for remaining spatial correlation problems in the residuals of the model, and following Pinkse and Slade (1998), we have applied Moran's *I* tests to the Pearson generalised residuals of the Dirichlet-multinomial regressions. Our results show that estimated coefficients are not biased by the presence of spatial correlation in residuals. Remaining details regarding this procedure are included in the Appendix.

Geographical variables such as AREA and ELEVATION appear to be more relevant for natural resources (NR) activities. Total surface of the municipality is also an important factor for scale economies (SE). Accessibility to transport infrastructures seems to be important for location of NR, PD and SE industries for PORTS (goods transport), SE and RD for AIRPORTS (people transport), and all industries for ROADS (people and goods transport). In general, coefficients for infrastructural factors appear in average higher than those of geographical measures. DISTANCE TO CENTRE, as a connectivity measure, appear to be important for NR and LI industries, both activities with remarkable content of logistics in their production and distribution activities.

URBANISATION economies, measured by total (density of) workers in urban surface at the municipal level appear to be a driving factor of firm's choices for all subsectors, with special emphasis for LI, PD and RD, emphasising the importance of labour pooling effects in the location of new firms. EDUCATION is also relevant, with major emphasis for RD, PD and SE industries. Despite the good behaviour of the human capital variable, we are aware of the relevance of counting on data about occupational tasks of workers in order to improve our focus on this issue (Florida 2002).

Technology and knowledge-related factors, measured by PATENTS, ICT INFRAST, and

R&D EXPEND, show interesting results, particularly for RD, SE and PD industries. Curiously, R&D expenditures show higher estimated values for coefficients of PD and SE industries, although also appearing significant in the case of RD as one would expect. Looking at the results for RD industries, ICT infrastructures appear to be the most important factor pursued by this type of companies in our sample, followed by Patents, and R&D expenditures. Such results are in line with other existing studies for EU countries, where capital inputs (R&D expenditures) are in second place of relevance in fostering knowledge economies in comparison to labour inputs (knowledge workers and creativity) (Raspe & Van Oort 2008). In the case of the SMA, development of infrastructures for knowledge is still an important task to be pursued, because of the imbalances shown in this regard between some leading urban areas, as Barcelona and Valencia cities, and less populated provinces located in the south.

Traditional localisation (LOCATION QUOTIENT) and diversity (DIVERSITY INDEX) economies seem to play a salient role in attracting new firms too. Diversity of industrial activities appears very important for location of any type of industry, showing the highest coefficients of all covariates in the model. LI and PD industries, those with higher content of inter-industry linkages in their final output show the highest coefficients, a result that is confirming similar findings on well renamed contributions in the (co) location literature (Ellison *et al.* 2010). Moreover, localisation economies (LOCATION QUOTIENT) show robust significance for all industries, and important sizes of coefficients too. NR companies seem to be the less affected by these two agglomeration forces, although still playing a role.

At the end of Table 3 we include results for the  $\delta$  parameter, showing them to be significant in the five groups of industries for the SMA area. Higher values of the parameter are shown for SE industries, and for RD industries, although all industries present relevant and very significant estimates for this variable. Such results appear to confirm our first working hypothesis, that spatial spillovers enter the profit functions of companies, and they account for neighbourhood-related external economies when choosing where to locate their



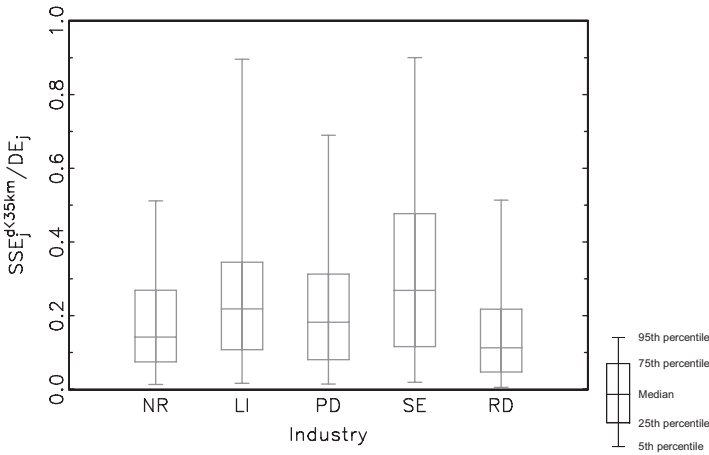


Figure 2. Distribution of spatial spillovers from each municipality.

establishments. In the following subsection we get deeper evidence on the issue.

**Analysing ‘spatial spillovers’ arising in the location model** – In this subsection we concentrate in measuring and describing spatial effects arising in the model. As a measure of the ability of every municipality  $j$  to generate spatial spillovers we compute the following relative index:

$$SSE_j^{\bar{D}} = \frac{1}{DE_j} \sum_{k \neq j; d_{jk} < \bar{D}} SSE_{j \rightarrow k} \quad (7)$$

Note that by normalising for the direct (locally bounded) effects of the shock ( $DE_j = dP_j/dX_j$ ), the index is a standardised (scale free) measure not depending on  $\beta$  parameter. The optimal threshold distance ( $DE_j = dP_j/dX_j$ ) employed to compute the index is 35 kilometres, which coincides approximately with the minimum distance such that every municipality in the sample has at least one neighbour (Arauzo-Carod 2008; Burger *et al.* 2010).<sup>10</sup>

Figure 2 shows the distribution of the ratio of spatial spillovers to direct effects ( $SSE_j^{\bar{D}}$ ) for all 314 municipalities in the sample. The ratio indicates by how much a shock in the municipality  $j$  changes the probability of a firm to locate in the surrounding area, relative to that of the own locality  $j$ , in the presence of spatial effects. We must note again that, first, the magnitude of the ratio  $SSE_j^{\bar{D}}$  relies on the relevance that a firm confers to neighbourhood

characteristics in their expected profit function, captured by the parameter  $\delta$ . *Ceteris paribus*, spatial spillovers will be more relevant for those industries with the highest  $\delta$ . And second, given that  $\delta$  parameter remains by definition constant for all municipalities in every industry, the size of the ratio also depends on the accessibility features defining every location. As previously discussed, the concept of accessibility relies on the degree of specialisation of the nearby locations in the particular industry  $s$ , as well as on the degree of closeness existing between location  $j$  and its neighbourhood.

Some conclusions emerge from results in Figure 2. In the first place, we observe that the ratios exhibit the highest median values for scale economies-based industries (SE), followed by labour-intensive (LI) and product differentiation (PD) ones, although, in general, all industries rely on those externalities in their location choices. In average, spatial effects increase 10 per cent to 30 per cent the probability of firms to choose nearby locations, relative to locally bounded traditional effects. In this way, they appear to be important determinants of location choices of firms, at least as important as other traditional explanatory variables shown in Table 3. Second, Figure 3 shows the distribution of the geographical scope where those spillovers arrive, now for the single measure of SSE effects introduced in equation (5). We define the mean scope of spatial spill-

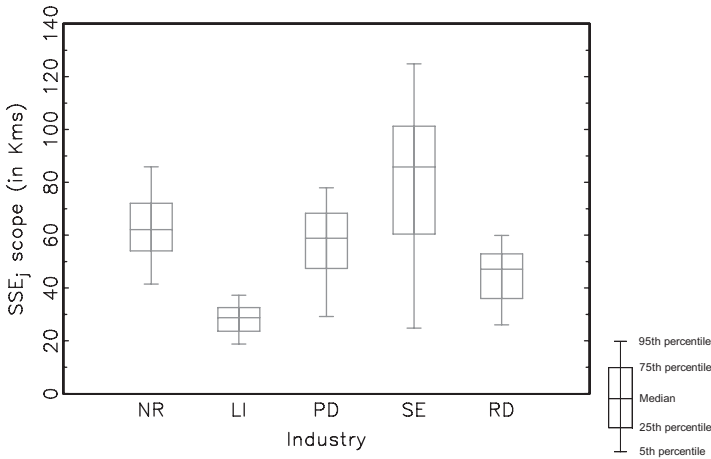


Figure 3. Distribution of the scope of spatial spillovers from each municipality.

Table 4. Results of the model for some selected municipalities in the sample.

Municipality	Accessibility	SSE (35 km)	SSE/DE (35 km)	Scope (in kms)
<b>SE industries</b>				
Oropesa	0.00152558	0.00025039	1.57714568	117.51
Torrent	0.00154239	0.00014749	5.15348307	93.44
Valencia	0.00297049	0.00068643	0.01873212	48.15
Barcelona	0.00576377	0.00143511	0.01630041	39.11
<b>RD industries</b>				
Oropesa	0.00148981	8.9631E-05	1.30324673	60.58
Torrent	0.00083412	4.6451E-05	3.75724381	52.32
Valencia	0.00385449	0.00012331	0.00166361	33.46
Barcelona	0.00489824	0.00032799	0.00199189	31.59

Notes: Accessibility for the municipality  $j$  is computed as  $\sum_{r \neq j} w_{rj} P_r$ , where  $w_{rj}$  are elements of the  $j$ -th column of the spatial weight matrix  $W$  and  $P_r$  is the probability of location of new firms  $r$  in the municipality  $r$ .

overs as the average distance in which a given municipality exerts positive spatial effects over other locations. As shown, the scope of spatial spillovers varies among sectors, from a mean of 28 kilometres in the case of LI sectors to 81 kilometres of SE industries, with higher values on the upper tail of the distribution.

Despite these general findings for the area of analysis, we present detailed results in Table 4 for some selected localities, in order to gain better insights on the role of spatial externalities in driving location processes. The table includes the results for accessibility indexes, SSE, the ratio  $SSE_j^D$ , and geographical scope of SSE for two central localities, namely the cities

of Valencia and Barcelona.<sup>11</sup> We also include data for two typical intermediate municipalities, with some degree of industrial specialisation and good connectivity with central cities in their provinces (15–25 kilometres), namely Oropesa in the province of Castellón, and Torrent in the province of Valencia. Results for more isolated localities are not included because they do not show relevant values in terms of the variables analysed. As main trends in the table, we see that urban centres show the highest accessibility values, and the highest SSE spilling over their neighbourhoods.<sup>12</sup> As expected, accessibility of locations becomes an issue in terms of generation of (absolute)

spatial effects, as well as the industry of reference chosen. However, central and more accessible localities present more limited values for  $SSE_j^D$ , and shorter scopes of their SSE in kilometres; that is, the spatial effects do not show higher values in the comparison with the volume of locally bounded traditional externalities, and they dissipate more rapidly in space, perhaps because they do not need to reach longer distances given higher concentration of specialised clusters in their neighbourhood and good accessibility issues. For the set of intermediate municipalities, Oropesa and Torrent, spatial effects are shown to be lower in absolute value, but so relevant in relative terms to direct effects of traditional externalities (1.3 to 5 times). They also reach a longer scope in kilometres compared to central places (up to 90–110 kilometres).

In terms of policy, spatial effects appear to reinforce existing clusters in central locations, although at a small scale than locally bounded traditional effects. In the case of intermediate industrial locations, spatial spillovers appear to be a remarkable force for attracting new firms in comparison to direct effects. Given that intermediate locations present good connectivity with central places too, we observe all types of externalities, traditional and spatial ones, reinforcing each other in fostering growth of central urban agglomerations. On balance, it leads to a typical centre-periphery pattern of regional development, with central places becoming greater, inland locations becoming more isolated, and intermediate developed cities reinforcing central places. Given this result, regional policies would have to continue focusing on balancing that pattern of development at a national and EU scale, to counteract market forces. However, regional policies would also need to be aware of the existing trade-off between the higher absolute value of spatial effects arising in central places, and the higher relative value of SSE/DE ratio observed for intermediate locations, in order to obtain the higher returns to private and public investments at a regional scale.

## CONCLUSIONS

In this paper we get deeper insight on the role played by spatial effects in shaping location

choices of firms. To this end, we have proposed a definition of neighbourhood-related spillovers for this framework of analysis, obtaining empirical evidence on the issue. The main findings of the investigation support the hypothesis that those externalities matter for manufacturing firms. Beyond the characteristics of each potential location, companies also take into account the features of the neighbouring area when building their establishments. Spatial spillovers clearly enter the (expected) profit function of the firm when choosing a new place; they vary by type of industry, and are closely related to accessibility features of the municipality where the firm is expecting to locate. In general, spatial effects account for 10 per cent to 30 per cent of locally bounded (direct) traditional effects arising in our sample. This is a remarkable value for any local government wanting to attract new firms to their territory.

Further results also point out that central places in the sample generate the most important spatial effects in absolute terms, while for municipalities with an intermediate level of development greater values are estimated for relative spatial effects (in terms of direct ones). Given the good accessibility of central and intermediate places, and the lower capacity of attraction of distant inland places, externalities naturally lead to a centre-periphery pattern of development. In this way, regional policies should, at least in the medium run, apply some cohesion actions to balance these forces. However, our results also indicate that any new investment in central and intermediate cities generates the most important spatial spillovers, and consequently the highest returns of the investment.

## APPENDIX A. SENSITIVITY OF SPATIAL SPILLOVERS MEASURE TO THE DEFINITION OF SPATIAL WEIGHT MATRIX

As argued in the paper, the choice of appropriate spatial weights ( $W$  matrix) is a key issue of spatial models as it assumes *a priori* a structure of spatial dependence, which may or may not correspond closely to the reality one wants to explore. In general, there are multiple possible choices, and the literature tells us little about adequate foundations for them. Typically, the spatial weights are defined as functions of

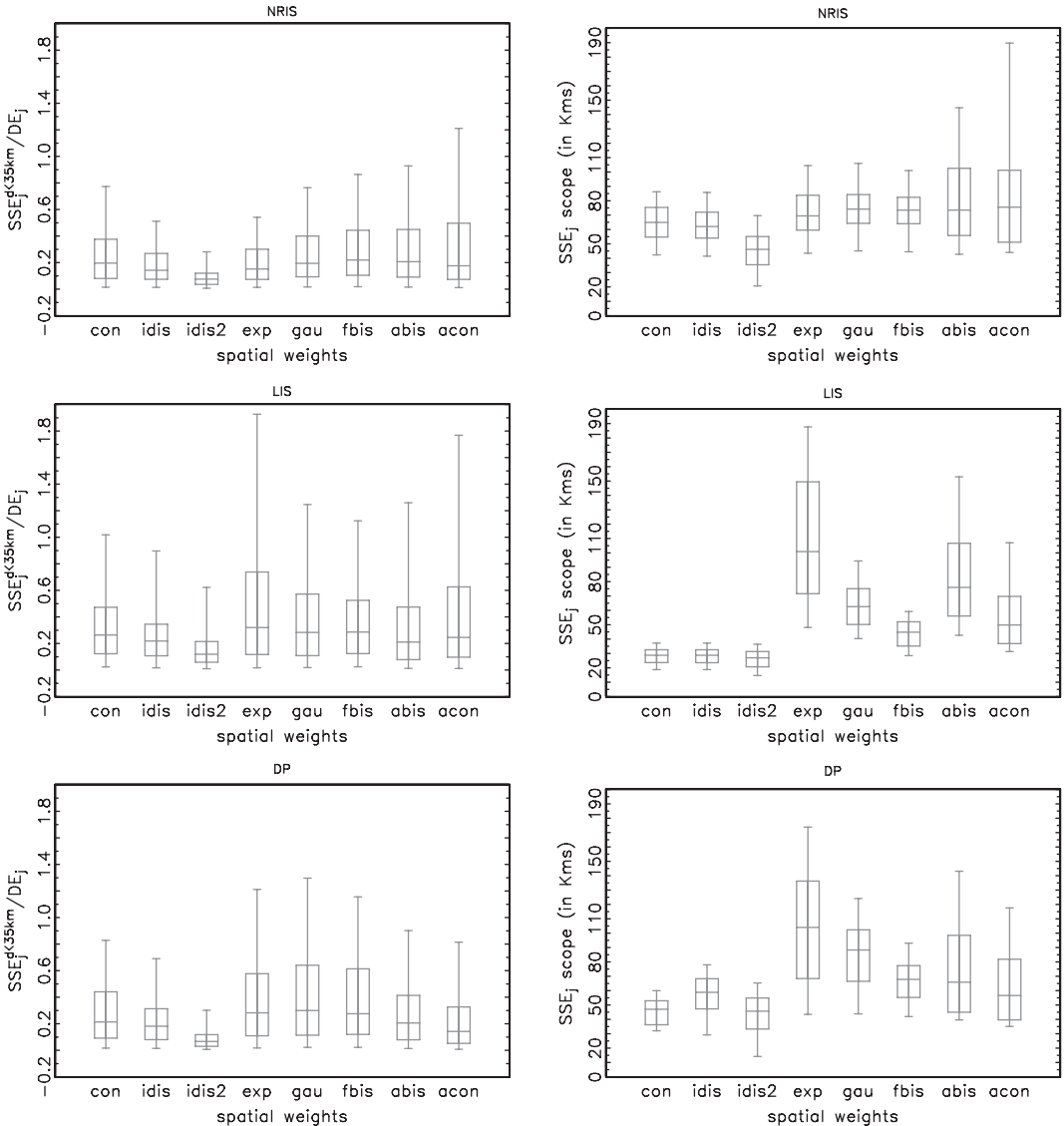


Figure A1. Distribution of spatial spillovers and geographical scope for alternative definitions of the spatial weight matrix.

economic or geographic distances, as the (inverse) distance decay functions (see e.g. Arauzo *et al.* 2009). Accordingly, we follow such an approach in specifying our  $W$  matrix in order to not introduce more complexity in the analysis, and concentrate on measurement issues of spatial spillovers in the model.

Yet, it was not the aim of the paper to present a thoroughly research on the issue, it appears

of interest, as some reviewers well pointed out, to note that our estimates should be interpreted as conditioned on the particular choice of the spatial weights we have made. In order to illustrate the influence of the definition of the spatial weights in the estimates of the spatial spillovers Figure A1 shows the distribution of such effects (and their geographical scope) for a collection of eight alternative spatial weights

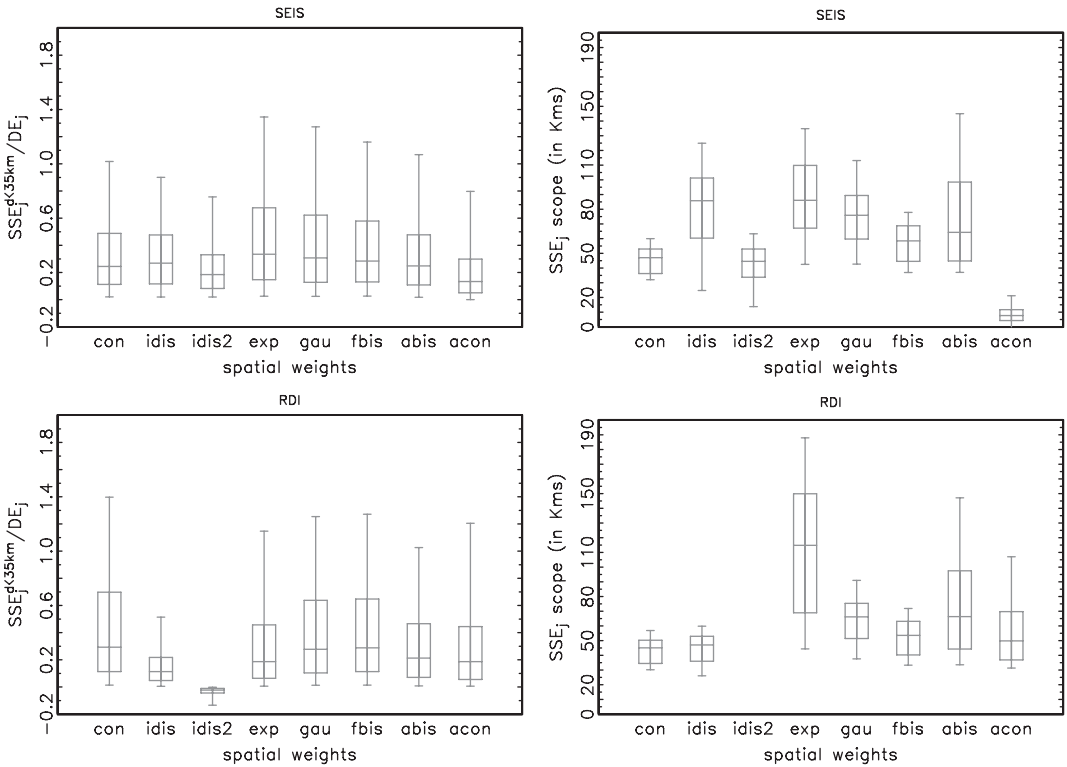


Figure A1. *Continued*

definitions (contiguity, inverse distance, inverse square distance, exponential decay, Gaussian, fixed bisquare, adaptive bisquare, and adaptive contiguity). The analysis suggests that our choice of the inverse distance based  $W$  matrix does not affect the size and scope of spatial spillover measures in a sharp way. Moreover, it seems that our results tend to underestimate the spatial effects regarding other  $W$  definitions as, for example, those of exponential decay, Gaussian, and bisquare. It is possible to employ a likelihood or information criteria (AIC, BIC) approach in choosing the more appropriate  $W$  matrix, but when it was attempted we obtained a different  $W$  definition for every industry, which does not make so much sense as a full methodology for treating our entire data set, so we decide to follow the mainstream literature in driving our approach on the issue. Anyway, all these estimates show the relevance of accounting for such spillovers in the location models, what was the main aim of the paper.

**APPENDIX B. COMPUTATION OF MORAN'S I TESTS FOR SPATIAL CORRELATION PROBLEMS IN THE RESIDUALS OF THE MODEL**

We have checked for remaining spatial correlation problems in the residuals of the estimated models. To this end, in the spirit of Pinkse and Slade (1998), we have applied a Moran's  $I$  test to the Pearson generalised residuals of the Dirichlet-multinomial regressions. Given that the asymptotic distribution of the test statistic is unknown, we take a numerical approach, which can be summarised as follows:

1. By using the Dirichlet-multinomial parameter estimates  $\hat{\theta} \equiv (\hat{\beta}, \hat{\delta})$ , we construct the generalised residuals

$$\hat{u}_j = \frac{n_j - \hat{P}_j \sum_j n_j}{\sqrt{\hat{\phi} \hat{P}_j (1 - \hat{P}_j) \sum_j n_j}}$$

where  $n_j$  represents

the number of new firms established

Table B1. Moran's *I* statistic for the generalised residuals of the Dirichlet-multinomial.

	NR	LI	PD	SE	RD
Moran's/p-value	0.0046 (0.995)	0.0116 (0.995)	0.0022 (0.995)	-0.0011 (0.995)	0.0035 (0.995)

in municipality  $j$ , and  $\phi =$

$$\frac{\alpha^{-1} \sum_j \exp\{X_j\beta + \delta WX_j\beta\} + \sum_j n_j}{\alpha^{-1} \sum_j \exp\{X_j\beta + \delta WX_j\beta\} + 1}.$$

After that, we compute the Moran's *I* statistics.

2. We draw a new random sample such that the probability of a new firm to localise in municipality  $j$  equals  $\hat{P}_j$  (for  $j = 1, 2, \dots, J$ ).
3. On the basis of the new data, we re-estimate the model to obtain new estimates of  $\hat{\theta}_r$ .
4. We repeat  $R$ -times steps (i) to (iii) with  $\hat{\theta}_r$  replaced by  $\hat{\theta}_r$ .

After completing  $R$  replications, the p-value of the test is estimated as the fraction of the  $R$  simulated statistic greater than the statistics computed for the original model's estimates. In our case, the estimated Moran's *I* statistic values (see Table B.1) seem to be not significant, that is, the null hypothesis of not remaining spatial correlation in the residuals of the model cannot be rejected (we use  $R = 200$ ). Subsequently, we are confident that findings reported in the paper are not biased by the presence of spatial correlation in residuals. We include a footnote on the paper for not adding more complexity to modelling issues.

**Acknowledgements**

Andrés Artal-Tur acknowledges financial support from the Spanish Ministry of Science and Innovation (Project ECO2011-27169). José M. Navarro-Azorín thanks Fundación Séneca-Región de Murcia (Project 08736/PHCS/08), and Luisa Alamá-Sabater the Spanish Ministry of Public Works (Project MIFI 3229/2008 and Project P21/08) for financial contributions to this research. Antonio J. Briones acknowledges financial support of Ministry of Foreign Affairs, AECID Project AP/043583/11.

**Notes**

1. The theoretical model builds on the original contribution for location analysis of Guimarães *et al.* (2004) and Woodward *et al.* (2006).
2. Defined as  $WX_j = \sum_{k \neq j} w_{jk} X_k$ , with weights  $W_{jk}$  inversely related to geographical distances among location  $j$  and neighbouring  $k$  locations, specified by using a spatial weight matrix.
3. It is assumed that the  $\eta_s$  are gamma distributed with parameters  $(\alpha^{-1} \lambda_j, \alpha^{-1} \lambda_j)$ , where  $\lambda_j \equiv \exp\{X_j\beta + \delta WX_j\beta\}$ , and  $\alpha > 0$ .
4. Given the definition of SSE we employ, we can observe both positive and negative effects. The term  $W_{hj}$  is inversely related to geographical distance between locations  $k$  and  $j$ , while  $\sum_{r \neq j} w_{rj} P_r$  is a weighted average of elements of the terms in the  $j$ -th column of the  $W$  matrix. Thus, if the term  $W_{hj}$  is greater than the weighted average, the spatial spillover would be positive, and negative otherwise.
5. The final measure employed for spatial effects is defined in relative terms to direct effects in the following sections, and do not rely on the value of the parameters in  $\beta$ , so this is not a pivotal assumption of the paper as we will see.
6. Note that our model integrates characteristics of all previous spatially extended location models in the literature, including spatially weighted explanatory variables (Autant-Bernard 2006; Woodward *et al.* 2006; Alamá *et al.* 2010) and a threshold distance for spillovers (Jofre-Monseny 2009). Additionally, we account for random effects in the specification of the model (Guimarães *et al.* 2004; Woodward *et al.* 2006). Beyond, we propose a statistical definition of the spatial spillover measure (in relative terms to locally bounded effects), and discuss the size and geographical scope of the spatial effects arising in the model.
7. In determining the value of  $R$ , we select the value that maximises the log-likelihood of the model through a grid search procedure.



8. We acknowledge the referees for calling our attention to this point.
9. We do not extend so much in the results of the location model, because our main focus is on the spatial terms of the extended model. However, following Arauzo-Carod *et al.* (2010) as a comprehensive review of the location literature, we find comparable results in terms of the employed explanatory factors, as well as for the magnitude of coefficients (marginal effects) obtained in previous studies. We elaborate some more on this issue in the text, deserving a greater attention to some interesting variables of the model.
10. Arauzo-Carod & Manjón-Antolín, (2009) pointed to 40–60 kms as the optimal threshold for analysing location choices in Catalonia region.
11. Note that these two cities not occupy a median position inside our sample, given their central role as urban centres for SMA (see Figure 2). However, we report results for these two particular cases in order to gain broader insights in policy terms.
12. Values in Table 4 appear to be small for absolute SSE effects, a result explained by the specific characteristics of the sample we employ. The great number of municipalities included in the study (314), determines the small value of spatial weights in  $W$  matrix for every municipality, and then the single value of probabilities computed. In this way, we show them just for comparison purposes between central and intermediate localities. In what regards the ratio SSE/DE, it can offer a closer idea of how spatial effects behave empirically, given its relative nature.

## REFERENCES

- ALAMÁ L., A. ARTAL & J.M. NAVARRO (2010), Industrial Location, Spatial Discrete Choice Models and the Need to Account for Neighbourhood Effects. *Annals of Regional Science* 47, pp. 393–418.
- ANSELIN, L. (2003), Spatial Externalities, Spatial Multipliers, and Spatial Econometrics. *International Regional Science Review* 26, pp. 153–166.
- ARAUZO-CAROD, J.M. (2008), Industrial Location at a Local Level: Comments on the Territorial Level of the Analysis. *Journal of Economic & Social Geography*, 92, pp. 193–208.
- ARAUZO-CAROD, J.M., D. LIVIANO-SOLIS & M. MANJÓN-ANTOLÍN (2010), Empirical Studies Industrial Location: An Assessment of their Methods and Results. *Journal of Regional Science* 50, pp. 685–711.
- ARAUZO-CAROD, J.M. & M. MANJÓN-ANTOLÍN (2009), *Optimal Spatial Aggregation in the Determinants of Industrial Location*. WP 2072/42866, Universitat Rovira i Virgili, Department of Economics.
- AUTANT-BERNARD, C. (2006), Where do Firms Choose to Locate their R&D? A Spatial Conditional Logit Analysis on French Data. *European Planning Studies* 14, pp. 1187–1208.
- BURGER, M.J., F.G. VAN OORT & G.A. VAN DER KNAAP (2010), A Treatise on the Geographical Scale of Agglomeration Externalities and the Modifiable Areal Unit Problem. *Scienze Regionali* 9, pp. 19–40.
- ELLISON, G., E.L. GLAESER & W.R. KERR (2010), What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100, pp. 1195–1213.
- FLORIDA, R. (2002), *The Rise of the Creative Class*. New York: Basic Books.
- GLAESER, E.L. (2007), *Agglomeration Economics*. Chicago, IL: NBER-University of Chicago Press.
- GUIMARÃES, P., O. FIGUEIREDO & D. WOODWARD (2004), Industrial Location Modeling: Extending the Random Utility Framework. *Journal of Regional Science* 44, pp. 1–20.
- IVIE (2010), *El Desarrollo del Arco Mediterráneo Español: Trayectoria y Perspectivas*. Alicante: Valencian Institute of Economic Research IVIE.
- JOFRE-MONSENY, J. (2009), The Scope of Agglomeration Economies: Evidence from Catalonia. *Papers in Regional Science* 88, pp. 575–591.
- LE SAGE, J. & R.K. PACE (2009), *Introduction to Spatial Econometrics*. Boca Raton, FL: Taylor & Francis Group.
- OECD (1987). *Structural Adjustment and Economic Performance*. Paris: OECD.
- PINKSE, J. & M. SLADE (1998), Contracting in Space: An Application of Spatial Statistics to Discrete-choice Models. *Journal of Econometrics*, 85, pp. 125–154.
- RASPE, O. & F.G. VAN OORT (2008), Localized Knowledge Externalities in Regional Economic Development and Firm Growth. *Journal of Regional Analysis and Policy* 38, pp. 100–116.
- ROBERTSON, R.D., G.C. NELSON & A. DE PINTO (2009), Investigating the Predictive Capabilities

- of Discrete Choice Models in the Presence of Spatial Effects. *Papers in Regional Science*, 88, pp. 367–389.
- ROSENTHAL, S.S. & W.C. STRANGE (2003), Geography, Industrial Organization and Agglomeration. *The Review of Economics and Statistics*, 85, pp. 377–393.
- WOODWARD, D.D., O. FIGUEIREDO & P. GUIMARÃES (2006), Beyond the Silicon Valley: University R&D and High-technology Location. *Journal of Urban Economics* 60, pp. 15–32.