ESTUDIOS SOBRE LA ECONOMIA ESPAÑOLA

Determinants of Industrial Location. An Application for Catalan Municipalities

Josep Maria Arauzo Carod

EEE 138

May 2002

FEDEA

Fundación de Estudios de Economía Aplicada

http://www.fedea.es/hojas/publicado.html

DETERMINANTS OF INDUSTRIAL LOCATION.
AN APPLICATION FOR CATALAN MUNICIPALITIES

Josep Maria Arauzo Carod
Department of Economics
Facultat de Ciències Econòmiques i Empresarials (Universitat Rovira i Virgili)
Av. de la Universitat, 1; 43204 Reus (Catalonia, Spain)
Phone: + 0034 977 798 800, Fax: + 0034 977 798 810
E-mail: jmac@fcee.urv.es

Summary

The aim of this paper is to show the main determinants of industrial location processes. We used local data (for Catalan municipalities) so that we could portray phenomena such as agglomeration economies (which are normally studied at regional or national level) better. We used the conditional logit model and the Poisson model to identify location patterns of firms. Our main statistical source was the REI (Spanish Industrial Establishments Register), which has plant-level microdata for the creation and location of new industrial establishments. Our results show that the characteristics of a municipality strongly affect the location decisions of industrial establishments (these decisions were different in all the sectors we considered).

Key words: industrial location, cities, logit models

1. Introduction

This paper discusses the main determinants of firms' location decisions. We show that these are not the same for all kinds of firms. We explain the guidelines for location decisions and how the territory affects the performance of firms and their location patterns. We know that territories influence the economic activities that take place within them. These activities have specific characteristics (the intensity of land use, the use of energetic inputs, human capital requirements, etc.) requiring specific environments that are provided in different ways by different types of cities. In other words, the resources available in a large town are not the same as those available in a small one.

The main part of this paper involves a Poisson-type estimation that differentiates between the incoming firms in accordance with their characteristics and a conditional-logit-type estimation that groups the location alternatives (municipalities) in homogeneous groups. The data base we used is also extreme interesting because it works on the municipal scale —a level of territorial aggregation not often used in this kind of study because it is difficult to collect reliable data.

This paper is organised as follows: in section 2 we discuss the models used (a conditional logit model and a Poisson model); in section 3 we present the data and the econometric specification; in section 4 we show our main results; and in section 5 we provide a short conclusion. Appendices follow section 5.

2. A theoretical approach

The conditional logit model (McFadden, 1974) is often used in the literature on industrial location. The point of departure is that for any establishment that wants to enter a market, the profit functions depend (partially) on its future location. That is, the entering establishment will choose its location to maximise its profits, because profits are different in each location. The first applications of this type of model for studying industrial location were from Carlton (1979 and 1983) and Bartik (1985). More recent studies are those by Coughlin *et al*. (1991), Woodward (1992), Friedman *et al*. (1992), Baudewyns (1999) and Guimarães *et al*. (2000b).

We analysed industrial location from the location decisions taken between 1987 and 1996 inside Catalan municipalities. There is an interaction between firms and the territories in which these firms are located, and we believe that these interactions are greater at the local level than at the regional or at national levels. As there have been few studies at this level, we also aim to identify the territorial patterns acting on industrial location from a local point of view.

Our starting point is the fact that a firm passes through several stages before it locates in a certain territory. These stages may be chronological or simultaneous, but they occur at the same time when the firm is created/located. Differentiating the various stages of an industrial location process allows scope for many areas of research but in this paper we are going to concentrate on just one of them.

We have considered the following stages:

- *Deciding to enter the market*. This occurs when a possible business opportunity is detected and capital or human resources are available.
- *Choosing the activity and the levels of technology and organisation*. This decision is linked to the previous one (each activity usually implies a specific level of technology and a minimum efficient size).
- *Choosing the location*. At this final stage, firms assume that the areas in which they could locate offer different levels of profit. At this stage the task is to identify the sites that offer maximum profits.

Although we concentrate on the stage at which firms choose the territory, we in fact consider all the stages involved since all these stages usually coexist.

We have focused on what happens at the final stage: choosing the location. This starting point may involve a certain bias because we only considered location decisions that actually led to the location of industrial establishments.

But in this paper we consider both municipalities that receive one or more industrial establishments and those that receive no industrial establishment. From a territorial point of view, this means that out of 946 municipalities in Catalonia, only 720 have been chosen as a site of their activities by one or more industrial establishments (between 1987 and 1996). We are therefore analysing location decisions that affect only 76% of municipalities in Catalonia, although the resident population in these municipalities is 98.8%.

Any bias due to not considering those municipalities (226) where no industrial establishment is located over the period analysed disappears partially when we use a Poisson model. This count model shows how many times each location (municipality) is chosen by an establishment. Hence, when $y = 0$ (i.e. municipalities where no establishment is located) are relevant because values of independent variables in these locations explain why they have not been chosen by entering establishments[1] (also, $y > 0$ denotes that at least one firm has been located at this municipality).

As we have already mentioned, in this paper we have decided to work at a local level[2] and use both a conditional logit model and a Poisson model. Our basic argument is that space matters. Territory is not a neutral variable: it affects the people and firms located in it in different ways. From this hypothesis, the most suitable method is a conditional logit model that uses only location characteristics (alternatives) as explanatory variables. This is because of the desire to explain territorial aspects more than the characteristics or results of firms.

However, there are computational problems involved in using a conditional logit model when there are a lot of alternatives. In Catalonia, for instance, the number of alternatives is determined by new firms choosing between 946 municipalities. To avoid these problems, Guimarães *et al.* (2000b) used a Poisson model to obtain results as a conditional logit model more easily.

We use the Poisson model when the dependent variable is a count variable (in this paper, the number of times a firm locates in a municipality). The number of alternatives in a conditional logit model equals the number of observations in a Poisson model. This implies that increasing alternative locations when we analyse the phenomenon at a local level is not a major problem[3]. Another advantage of Poisson models is that nil observations[4] do not imply modelisation problems (unlike conditional logit models).

When using a Poisson model to analyse industrial location, Guimarães *et al*. (2000b) considered two situations:

1. Entrant decisions are formulated according to a vector of territorial variables shared by all entrants ($z_{ij} = z_j$).

2. Entrant decisions are formulated according to a vector of territorial variables shared by groups of entrants ($z_{ij} = z_{jg}$ for $g = 1, 2, ..., G$, where G denotes the number of groups).

In the first situation, site characteristics affect all the entrants equally and do not distinguish between them. Entering establishments are therefore homogeneous because we do not consider their specific characteristics. This is obviously a simplification but we think that at a first stage it is not a major problem.

In the second situation, different groups of entrants (G) are also affected in a different way by location characteristics. Grouping can be done according to criteria such as the date of location, the activity or size of the establishment, etc. In this way we try to identify the heterogeneity of entering establishments.

We consider that firms choose the municipality where they begin their activity to maximise profits (expected profits are not neutral to location). If we follow the method of Guimarães *et al.* (2000a), an industrial establishment i that attempts to be located in Catalonia must choose between 946 municipalities. When this establishment i locates at municipality j, it reaches a profit level π_{ij}, which includes a deterministic term (U_{ij}) and a stochastic term (ε_{ij}). Formally:

$$\pi_{ij} = U_{ij} + \varepsilon_{ij} \tag{2.1}$$

In this model, firm i will choose location j if:

$$\pi_{ij} > \pi_{ik}, \qquad \forall k, k \neq j \tag{2.2}$$

The stochastic nature of the profit function implies that the probability that location j is selected by firm i is:

$$P(Y_i = j) = \text{Prob}\ (\pi_{ij} > \pi_{ik}) \qquad \forall k, k \neq j \tag{2.3}$$

(Notice that Y_i is a random variable that shows the choice made).

In other words (Greene, 1998), each Y_i is a random variable with Poisson distribution and with λ_i parameter (related to regressors x_i):

$$P(Y_i = y_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} \qquad y_i = 0, 1, 2, ... \tag{2.4}$$

In which the most common representation of λ_i is:

$$\ln \lambda_i = \beta' x_i \tag{2.5}$$

And where the expected events are:

$$E[y_i|x_i] = Var[y_i|x_i] = \lambda_i = e^{\beta' x_i} \tag{2.6}$$

And:

$$\frac{\partial E[y_i|x_i]}{\partial x_i} = \lambda_i \beta \tag{2.7}$$

We can also use a conditional logit model in which we overcome the problems caused by the large number of choices (municipalities) by dividing them into homogeneous groups. Thus we assume that the characteristics of a town depend on its size and we group the 946 municipalities that the 17,718 entering establishments could choose according to population[5]:

- Less than 2,000 inhabitants;
- Between 2,000 and 10,000 inhabitants;
- Between 10,001 and 50,000 inhabitants;
- Between 50,001 and 100,000 inhabitants;
- Between 100,001 and 1,000.000 inhabitants; and
- More than 1,000.000 inhabitants.

In this way we reduce the number of choices from 946 to just 6, and can now calculate the estimators If we consider that ε_{ij} is distributed independently and according to a Weibull distribution, the probability that a firm *i* is located at municipality *j* is:

$$P_{ij} = \frac{\exp(U_{ij})}{\sum_{k=1}^{6} \exp(U_{ik})} \qquad (2.11)$$

Now we consider that the deterministic term (U_{ij}) of profit function (2.1) is a linear combination of *m* variables linked to location *j*:

$$U_j = \beta_1 X_j^1 + \beta_2 X_j^2 + \ldots + \beta_m X_j^m \qquad (2.12)$$

We selected the variables in (2.12) according to statistical availability and economic relevance and because we considered them to affect a firm's profit level.

3. Data and econometric specification

Our data refer to the 946 municipalities in Catalonia[6]. These municipalities have an average population of 6,367 inhabitants (in 1986). They received 17,719 new industrial establishments in the 10 years we studied (from 1987 to 1996). The entrants mainly went to the larger municipalities: 88.7% of these new establishments are concentrated in cities with over 2,000 inhabitants.

Our information is from the REI (Spanish Industrial Establishments Register), which has plant-level microdata (plant-level) about the creation and location of new industrial establishments[7]. Newcomers were mainly small —7.8 employees per establishment.

In this paper we explain firms' location decisions according to the characteristics of the site (municipality). These are: population density (DENSI), urbanisation economies (URB), urbanisation dis-economies (DIS-E), industrial diversity (DIV), the level of commuting (COMM), the skill level of human capital (HC), the percentage of jobs in the industrial sector (IND), the percentage of jobs in the services sector (SER), the distance from the capital of the *comarca* (DISCAPI), the municipalities with at least 100,000 inhabitants (DIS100), and the accessibility to transport networks i.e. highways (HIG) and rail (RAIL). See Appendix 1 for a more detailed description of the variables.

We used three econometric specifications[8]:

1. A Poisson model in which the site vector of variables affects all entrants equally (Guimarães *et al.*, 2000b: situation 1).

2. A Poisson model in which the site vector of variables affects entrants differently according to their characteristics (here, the characteristics are the sector[9] and size of the establishments) (Guimarães *et al.*, 2000b: situation 2).

3. A conditional logit model in which sites are divided into 6 groups according to the population of the municipality (Bartik, 1985).

For the first specification (Poisson for all entrants) and the second specification (Poisson for entrants grouped):

$$\begin{aligned}\log\left[E(y|x_1, x_2, ..., x_k)\right] = \beta_0 &+ \beta_1 DENSI + \beta_2 CH + \beta_3 DIV + \beta_4 COMM + \\ &+ \beta_5 IND + \beta_6 SER + \beta_7 DISCAPI + \beta_8 DIS100 \qquad (2.13) \\ &+ \beta_9 \mathrm{URB} + \beta_{10} \mathrm{DESU} + \beta_{11} \mathrm{AUT} + \beta_{12} \mathrm{TREN}\end{aligned}$$

For the third specification (logit condition for groups of municipalities), we can write the deterministic term of profit function (2.1) at municipality *j* as:

$$U_j = \beta_1 CH_j + \beta_2 COMM_j + \beta_3 IND_j + \beta_4 SER_j + \beta_5 URB_j + \nu_j \qquad j = 1,...,6 \qquad (2.14)$$

4. Results

We present the results of the Poisson estimation first by assuming that the vector of variables acts equally over all entering establishments, second by grouping the entering establishments by sector and third by grouping entering establishments by size. We then discuss the results of a conditional logit estimation for alternatives (municipalities) in terms of the size of the city.

The signs of the coefficients in the Poisson model (Table 1) are as expected i.e. firms are more influenced by variables related to their activities than by variables related to amenities. These results are quite robust because they are very similar for the other specifications.

The entering establishments are affected positively by the level of economic activity in the municipalities: urbanisation economies act positively, i.e. the higher the concentration of jobs (irrespective of sector), the greater the attraction of the location. At the same time, however, if a municipality concentrates too much on an activity, there are negative externalities (urbanisation diseconomies) that restrict its ability to attract new establishments. The most attractive activity (which may act as location determinants) is the service industry, followed by the manufacturing industry.

Table 1

Industrial location determinants

(Poisson model)

	Coefficients
CONS	2.04709 (.08152)*
DENSI	-4.36e-06 (4.48e-06)
URB	.00073 (.00003)*
DIS-E	-4.55e-08 (3.92e-09)*
DIV	-1.24900 (.06187)*
COMM	-.17776 (.00592)*
HC	-.01188 (.00472)*
IND	.03854 (.00063)*
SER	.04790 (.00079)*
DISCAPI	-.00003 (1.46e-06)*
DIS100	-.00002 (5.45e-07)*
HIG	.32226 (.02763)*
RAIL	-.36243 (.04492)*
χ^2	54699.39*
Log likelihood	-6077.6065

(*) Significance at 5%; standard error in brackets

Entering firms prefer a diversified industrial environment: the greater the industrial diversity in a municipality, the greater the probability that it will receive new establishments. The opposite is true for the skill level of human capital: the higher the skill level in a municipality, the lower the probability that it will receive new firms. There are two possible reasons for this. Firstly, the industrial mix in Catalonia is biased towards activities with medium and low technological levels (which implies that the skill level of human capital is not a strong location factor)[10]. Secondly, there is a growing mismatch between the location of one's job and the location of one's residence (which implies that firms are not located in the municipality where the workforce lives — they could move to neighbouring municipalities, for instance).

The level of commuting complements the previous statement. Effectively, those municipalities that daily require the most commuters (measured in terms of kilometres travelled by residents) have a lowest probability of being chosen by new industrial establishments. This confirms the mismatch between the location of one's job and the location of one's residence. These results are as expected if we assume that residents who commute great distances do so because their environment offers them amenities that are more important than commuting costs.

Also, entering establishments do not only look for a particular economic environment: they also try to locate close to administrative centres. We have measured this in terms of the distance from the municipality to the capital of the *comarca* (DISCAPI)[11] or to the nearest municipality with at least 100,000 inhabitants (DIS100).

If we consider transport infrastructure, we can see that there is another territorial mismatch between rail network and economic activity. Highways are more useful to firms than trains. The rail network in Catalonia (not that of the coastline) runs through several regions that were pioneers of the process of industrialisation begun in the 19th Century, but which now suffer from a lack of competitiveness and a less dynamic industry.

A previous estimation showed data in an aggregated way and did not consider that entrants may have different environmental requirements according to characteristics such as size or sectorial activity. Although the results are not much different, when we disaggregate the entrants we find certain important clarifications.

Table 2

Industrial location determinants

(Poisson model by groups: OECD classification)

	Natural resources	Labour intensive	Sectors Economies of scale	Differentiated products	R+D intensive
CONS	0.23548	0.49372	-0.28873	1.30968	-3.06108
	(0.14395)	(0.11506)*	(0.21150)	(0.10860)*	(0.53367)*
DENSI	-0.00002	-0.00001	-0.00002	1.25e-06	0.00004
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00003)
URB	0.00080	0.00098	0.00055	0.00057	0.00076
	(0.00010)*	(0.00006)*	(0.00010)*	(0.00005)*	(0.00024)*
DIS-E	-4.60e-08	-8.75e-08	-7.67e-09	-2.25e-08	-3.59e-08
	(1.12e-08)*	(6.83e-09)*	(1.11e-08)	(6.20e-09)*	(2.80e-08)
DIV	-1.16786	0.09312	-3.15487	-2.52402	-2.45627
	(0.15160)*	(0.09684)	(0.21058)*	(0.11031)*	(0.54023)*
COMM	-0.14465	-0.24604	-0.13957	-0.14462	-0.10596
	(0.01375)*	(0.01103)*	(0.01674)*	(0.00934)*	(0.04199)*
HC	0.01057	0.00670	-0.01736	-0.03330	0.04603
	(0.01149)	(0.00863)	(0.01291)	(0.00766)*	(0.03116)
IND	0.02575	0.03517	0.05116	0.04373	0.04294
	(0.00144)*	(0.00111)*	(0.00206)*	(0.00105)*	(0.00540)*
SER	0.04229	0.05019	0.05506	0.04833	0.05363
	(0.00177)*	(0.00141)*	(0.00252)*	(0.00133)*	(0.00610)*
DISCAPI	-0.00001	-0.00003	-0.00003	-0.00004	-0.00003
	(0.00000)*	(0.00000)*	(0.00000)*	(0.00000)*	(0.00001)*
DIS100	-1.42e-05	-1.89e-05	-2.28e-05	-2.16e-05	-6.95e-06
	(1.01e-06)*	(7.61e-07)*	(1.32e-06)*	(7.24e-07)*	(2.83e-06)*
χ^2	6189.21*	15619.66*	8553.13*	23622.25*	2044.55*
Log likelihood	-1716.262	-2965.0302	-1461.7821	-3156.7097	-454.04332

(*) Significance at 5%; standard error in brackets

First of all, we took into account the activities of the entrants (Table 2). We grouped the entrants into five sectors according to OECD classification (natural resources: labour intensive: economies of scale: differentiated products and R+D intensive).

The coefficients for commuting were negative for all estimates, but the most negative coefficient was in the labour-intensive sector. This could be also explained by the fact that the more daily exits there are from a municipality, the smaller is the probability that it will receive an industrial establishment. The least negative coefficient was in the R+D- intensive sector. This may be because a municipality with a high level of commuting has a better quality of life (certain amenities, for instance), and the kind of environment that is suitable for residential purposes may outweigh the inconvenience of having to commute to work every day. If we assume that R+D-intensive firms usually locate in non-rundown areas (because this adds value to their activities), the lower level of commuting is easy to understand.

The R+D-intensive sector and the economies-of-scale sector are the only ones for which urbanisation dis-economies have no significance i.e. for firms in this sector a high concentration of jobs in a municipality does not make it less attractive. This is because these firms look for sites in which there is a large number of dynamic firms so that they can obtain advantages from knowledge spillovers.

At the same time, the labour-intensive sector is the only one for which lack of industrial diversity does not make a municipality less attractive. One explanation for this could be that of the 'nursery cities' (Duranton and Puga, 2001): activities in this sector do not continuously introduce innovations, so they do not need a diversified and dynamic industrial environment from which they can obtain (or copy) innovations. For firms in this sector, a specialised environment is usually more profitable.

Thirdly, we considered the size of firms entering a municipality (Table 3). Our results are similar to those of previous studies[12]: the location patterns of small firms are mainly linked to current population and job distribution but those of large firms are less affected by these variables. We consider the hypothesis that small firms are more receptive to variables that are related to the personal background of the entrepreneur (as these variables are not available, they are not used in this estimate), while large firms are more receptive to objective (economic) variables.

The level of commuting and the skill level of human capital clearly show that the size of industrial establishments affects their behaviour. Small firms, for instance, are less likely than large firms to be located in municipalities where the level of commuting is high. One explanation for this could be the role of edge (mainly residential) cities built near large industrial firms. Also, small firms, i.e. those with less than 50 employees, react negatively to the skill level of human capital but this variable is not significant for large firms (in small firms the percentage of workers with a university degree is lower than in large firms). We can also see that the role of the other variables is similar to the previous estimations.

Table 3

Industrial location determinants

(Poisson model: size of the firm)

	Size of the firm			
	Less than 10 workers	**Between 10 and 49 workers**	**Between 50 and 99 workers**	**More than 99 workers**
CONS	1.80160 (0.06988)*	-0.01150 (0.18560)	-3.33068 (0.98924)*	-7.15958 (1.66565)*
DENSI	-3.84e-06 (4.97e-06)	-1.13e-05 (1.05e-05)*	1.08e-05 (5.21e-05)	-4.54e-06 (8.8e-05)
URB	0.00074 (0.00004)*	0.00070 (0.00008)*	0.00022 (0.00040)	-0.00011 (0.00062)
DIS-E	-4.73e-08 (4.34e-09)*	-3.77e-08 (9.20e-09)*	1.08e-05 (5.21e-05)	7.78e-08 (6.72e-08)
DIV	-1.25077 (0.06731)*	-1.75070 (0.16307)*	-2.00189 (0.80860)*	-3.33114 (1.16949)*
COMM	-0.17925 (0.00642)*	-0.12473 (0.01485)*	-0.20824 (0.07545)*	0.00387 (0.08767)
HC	-0.00901 (0.00521)**	-0.02918 (0.01181)*	-0.04429 (0.05930)	0.07246 (0.05450)
IND	0.03638 (0.00068)*	0.04899 (0.00178)*	0.05756 (0.00959)*	0.09108 (0.01684)*
SER	0.04861 (0.00085)*	0.04799 (0.00235)*	0.06307 (0.01200)*	0.07780 (0.02003)*
DISCAPI	-0.00003 (0.00000)*	-0.00004 (0.00000)*	-0.00003 (0.00002)**	-0.00004 (0.00002)**
DIS100	-1.79e-05 (4.62e-07)*	-3.02e-05 (1.34e-06)*	-3.7e-05 (6.80e-06)*	-1.88e-05 (7.64e-06)*
χ^2	43623.89*	10747.78*	443.60*	263.24*
Log likelihood	-5251.7834	-1742.2045	-222.21656	-160.08129

(*) Significance at 5%, (**) significance at 10%; standard error in brackets.

Fourthly, we used a conditional logit estimation in which we assumed that, from a firm's point of view the size of the town indicates its characteristics, and divided the sites (municipalities) into six groups according to population[13] (Table 4).

We grouped the sites because logit estimations have computational constraints when the number of alternatives (in this case, the number of municipalities) is high. This reduces the number of alternatives from 942 to 6, so it is easier to estimate the model. With this simplification we can lose information[14], so we aggregated different alternatives to create groups that were highly homogeneous.

Table 4
Industrial location determinants
(conditional logit model)

	Coefficients
URB	-.00539 (.00018)*
COMM	-4.53404 (.14492)*
HC	2.7833 (.08731)*
IND	-.22437 (.00952)*
SER	.05762 (.00243)*
χ^2	3672.55*
Log likelihood	-29892.201

(*) Significance at 5%; standard error in brackets

The results show the probability that a location chosen by an industrial establishment is affected negatively by the density of jobs per km^2 (URB). Municipalities with a greater intensity of land use for economic purposes create negative externalities that make them less attractive to potential entrants (increases in land prices, congestion, etc.).

At the same time, the probability that territories with a higher level of commuting (exits) will receive new industrial establishments is low. The explanation for this is similar to that of previous estimations: high daily exit levels indicate a mainly residential specialisation. This implies that the equipment and infrastructure in the municipality is designed for residential purposes and does not necessarily correspond with the requirements of industrial establishments.

On the other hand, the higher the skill level of human capital in a municipality, the greater the probability that it will receive industrial firms. This is not consistent, however, with territorial patterns in which more skilled and better paid individuals tend to locate further away from their place of work in search of a less dense environment and a better quality of life.

Sector variables reveal something we have not seen in previous estimations i.e. industrial relocations from traditional industrial municipalities (usually town centres) to other locations (usually on the outskirts of towns) where the service sector carries a greater weight.

5. Conclusions

One of the main conclusions of this paper is that entering establishments are more influenced by the territorial distribution of economic activity than by the territorial distribution of population. That is, the concentration of jobs in a municipality acts positively on entrant firms, while the concentration of population does not have any power to attract new firms. These results raise an important issue, namely that the location patterns of firms and populations are different.

From a sectorial point of view, the weight of industry and services in a municipality is also a locational factor: the greater the weight of services in a municipality, the higher the probability that new firms will enter. Industrial diversity also favours industrial location, but this is not consistent for all the econometric specifications.

The geographical location of municipalities has an affect on the location of firms. As the previous estimations show, the greater the distance between a municipality and the capital of the *comarca* in which it is located (or town with over 100,000 inhabitants), the lower the probability that it will receive new industrial establishments. Also, the URB and DIS-E variables show that firms prefer sites with a high concentration of economic activity (because there is greater competitiveness) than sites with access to public and private infrastructure or to labour markets. However, if this concentration is too great and firms have to endure agglomeration diseconomies (congestion, soaring of land prices, etc.), these locations are less attractive to new firms.

Acknowledgements

This research was partially funded by the CICYT (SEC2000–0882–C02–02). I would like to acknowledge the helpful and supportive comments of Agustí Segarra and Miguel Manjón and suggestions of one anonymous referee. I am also grateful to the Centre for Spatial and Real State Economics (CSpREE) of the University of Reading, where this paper was finished during my stay as a research visitor. Any errors are, of course, my own.

References

ARAUZO, J. M. and MANJÓN, M. (2001): "Firm Size and Geographical Aggregation: An Empirical Appraisal in Industrial Location", *Working Paper 5-2001*, Departament d'Economia (Universitat Rovira i Virgili).

BARTIK, T. J. (1985): "Business Location Decisions in the U.S.: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States", *Journal of Business and Economic Statistics*, 3 (1), 14-22.

BAUDEWYNS, D. (1999): "La localisation intra-urbaine des firmes: une estimation logit multinomiale", *Revue d'Économie Régionale et Urbaine*, 5, 915-930.

CARLTON, D. (1983): "The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables", *Review of Economics and Statistics*, 65, 440–449.

CARLTON, D. (1979): "Why new firms locate where they do: An econometric model, in W. Wheaton (ed.), *Interregional Movements and Regional Growth*, The Urban Institute, Washington.

COSTA, M. T.; SEGARRA, A. and VILADECANS, E. (2000): "Business dynamics and territorial flexibility", Paper presented at *40th European Regional Science Association Congress* (Barcelona).

COUGHLIN, C. C.; TERZA, J. V. and ARROMDEE, V. (1991): "State characteristics and the location of foreign direct investment within the United States", *The Review of Economics and Statistics*, 73 (4), 675-683.

DURANTON, G. and PUGA, D. (2001): "Nursery Cities: Urban diversity, process innovation, and the life cycle of products", *American Economic Review*, 91(5), 1454-1477.

FIGUEIREDO, O. and GUIMARÃES, P. (1999): "Start-ups Domestic Location and the Entrepreneur's Geographical Origin", *Working Paper 91*, Facultade de Economia, Universidade do Porto.

FRIEDMAN, J.; GERLOWSKI, D. A. and SILBERMAN, J. (1992): "What attracts foreign multinational corporations? Evidence from branch plant location in the United States", *Journal of Regional Science*, 32 (4), 403-418.

GREENE, W. H. (1998): *Econometric Analysis*, Prentice Hall, New Jersey.

GUIMARÃES, P., FIGUEIREDO, O. and WOODWARD, D. (2000a): "Agglomeration and the Location of Foreign Direct Investment in Portugal", *Journal of Urban Economics*, 47, 115-135.

GUIMARÃES, P., FIGUEIREDO, O. and WOODWARD, D. (2000b): "A Tractable Approach to the Firm Location Decision Problem", *Working Paper NIMA 2/2000*, Universidade do Minho.

HENDERSON, V.; KUNCORO, A. and TURNER, M. (1995): "Industrial development in cities", *Journal of Political Economy*, 103 (5), 1067-1090.

McFADDEN, D. (1974): "Analysis of qualitative choice behaviour", in P. Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, 105–142.

MEESTER, W. (2000): "Locational self-preference of firms", Paper presented at *40th European Congress of the Regional Science Association* (Barcelona).

MOMPÓ, A. and MONFORT, V. (1989): "El Registro Industrial como fuente estadística regional: el caso de la Comunidad Valenciana", *Economía Industrial*, 268, 129-140.

STAM, E. and SCHUTJENS, V. (2000): "Locational Behaviour of Young Firms: a Life Course Perspective", Paper presented at *40th European Congress of the Regional Science Association* (Barcelona).

WOODWARD, D. (1992), 'Locational determinants of Japanese manufacturing start-ups in the United States', *Southern Economic Journal*, 58, 690–708.

Appendices

Appendix 1

Explanatory variables: definition and sources

Variable	Definition	Source
DENSI	Population density (1986)	Catalan Statistical Institute
COMM	*Commuting* (km travelled daily per employee) (1986)[15]	Our own variable using data from the Catalan Statistical Institute and the Catalan Cartographical Institute
IND	% of jobs in the industrial sector (1986)	Catalan Statistical Institute
SER	% of jobs in the services sector (1986)	Catalan Statistical Institute
HC	% of population with a university degree (1986)	Catalan Statistical Institute
DIV	HHI Index of the diversity of industrial jobs (1986)[16]	Catalan Statistical Institute
DISCAPI	Distance (in km) to the capital of the *comarca*	Catalan Cartographical Institute
DIS100	Distance (in km) to the nearest city with at least 100,000 inhabitants	Catalan Cartographical Institute
URB	Urbanisation economies (jobs / km^2)[17]	Catalan Statistical Institute
DIS-E	Urbanisation dis-economies (URB^2)[18]	Catalan Statistical Institute
RAIL	Accessibility to a rail network in the same *comarca*[19]	Catalan Government
HIG	Accessibility to a highway network in the same *comarca*[20]	Catalan Government

Appendix 2

CNAE-93 industrial classification (sectors 12 to 36)

Code	Name
12	Mining of uranium and thorium ores
13	Mining of metal ores
14	Other mining and quarrying
15	Manufacture of food products and beverages
16	Manufacture of tobacco products
17	Manufacture of textiles
18	Manufacture of leather clothes
19	Tanning and dressing of leather
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
21	Manufacture of pulp, paper and paper products
22	Publishing, printing and reproduction of recorded media
23	Manufacture of coke, refined petroleum products and nuclear fuel
24	Manufacture of chemicals and chemical products
25	Manufacture of rubber and plastic products
26	Manufacture of other non-metallic mineral products
27	Manufacture of basic metals
28	Manufacture of fabricated metal products, except machinery and equipment
29	Manufacture of machinery and equipment n.e.c.
30	Manufacture of office machinery and computers
31	Manufacture of electrical machinery and apparatus n.e.c.
32	Manufacture of radio, television and communication equipment and apparatus
33	Manufacture of medical, precision and optical instruments, watches and clocks
34	Manufacture of motor vehicles, trailers and semi-trailers
35	Manufacture of other transport equipment
36	Manufacture of furniture; manufacturing n.e.c.

Source: INE (Spanish Statistical Institute).

Appendix 3

Classification of the manufacturing activities according to the elements that influence their competitiveness (OECD)

CNAE-93	Natural resources intensive sectors
15	Manufacture of food products and beverages
16	Manufacture of tobacco products
21	Manufacture of pulp, paper and paper products

CNAE-93	Labour intensive sectors
17	Manufacture of textiles
18	Manufacture of leather clothes
19	Tanning and dressing of leather
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
36	Manufacture of furniture; manufacturing n.e.c.

CNAE-93	Scale economies intensive sectors
24	Manufacture of chemicals and chemical products
25	Manufacture of rubber and plastic products
34	Manufacture of motor vehicles, trailers and semi-trailers
35	Manufacture of other transport equipment

CNAE-93	Sectors with differentiated products
22	Publishing, printing and reproduction of recorded media
26	Manufacture of other non-metallic mineral products
27	Manufacture of basic metals
28	Manufacture of fabricated metal products, except machinery and equipment
29	Manufacture of machinery and equipment n.e.c.
31	Manufacture of electrical machinery and apparatus n.e.c.

CNAE-93	R+D intensive sectors
30	Manufacture of office machinery and computers
32	Manufacture of radio, television and communication equipment and apparatus
33	Manufacture of medical, precision and optical instruments, watches and clocks

Source: INE (Spanish Statistical Institute).

Notes

[1] However, with this argument, there is a problem of how to choose the samples because an undetermined number of firms were not able to locate, so we did not count them. All of these are counted as zero.
[2] See Arauzo and Manjón (2001) for a discussion on territorial dis–aggregation (cities, counties, regions, etc.) when defining the data set and how this choice affects results.
[3] Obviously, working at a local level involves more observations than when working at regional or national level.
[4] In this context, nil observations are those from municipalities with no industrial location over the period covered. In this case the dependent variable counts how many times an industrial establishment locates in each municipality (this ranges between 0 and *n*, and *n* denotes the number of entering establishments).

[5] In his paper on the location decisions of firms in the USA, Bartik (1985) used a similar solution and grouped sites inside each state. Another solution is to make a multivariate analysis by dividing municipalities into homogeneous groups according to the variables used.
[6] Catalonia is an autonomous region of Spain with about 6 million inhabitants (15% of the Spanish population) and an area of 31895 km^2. It contributes 19% of Spanish GDP. Catalan municipalities have an average area of 19.6 km^2. The capital of Catalonia is the city of Barcelona.
[7] See Mompó and Monfort (1989) for further information about the REI.
[8] The variables used in the 3 estimations are roughly the same.
[9] We made the sectorial grouping using the OECD classification according to the main competitiveness of each sector and constructed the groups from the Spanish classification CNAE-93 (see appendices).
[10] We must therefore stress that this first estimation was made with aggregate data. The following estimations will clarify these initial results.
[11] The territorial division of Catalonia is based on municipalities and *comarques*. *Comarques* are made up of municipalities. There are 41 *comarques* in Catalonia with an average area of 781 km^2 and an average population of 145,000.
[12] See Costa *et al*. (2000) for Spain, Meester (2000) for Germany and the Netherlands, Figueiredo and Guimarães (1999) for Portugal and Stam and Schutjens (2000) also for the Netherlands.
[13] We constructed the following groups: municipalities with less than 2,000 inhabitants, municipalities with between 2,000 and 10,000 inhabitants, municipalities with between 10,001 and 50,000 inhabitants, municipalities with between 50,001 and 100,000 inhabitants, municipalities with between 100,001 and 1,000,000 inhabitants and municipalities with more than 1,000,000 inhabitants (only Barcelona).
[14] One of the disadvantages, for instance, is that some variables linked to territory cannot be used e.g the capitals of the *comarques*, whether the municipality is on the coast, the distance to the capital of the *comarca* or to the nearest municipality with over 100,000 inhabitants, and rail and motorway networks.
[15] Here we used the mean distance travelled by workers in each municipality.
[16] We used a Hirshmann-Herfindahl index (DIV) of diversity of industrial jobs. This is a non-diversity index, in which values range from 0 (diversity) to 1 (non-diversity).
[17] Urbanisation economies are static externalities that show the efficiency gains of firms located in the same place. This variable is measured by the density of jobs (jobs / km^2).
[18] Urbanisation dis-economies show decreases in the efficiency of firms due to an excessive concentration of jobs in the same place. This variable is measured by the square of the urbanisation economies variable (URB^2).
[19] RAIL shows whether the municipality is located in a *comarca* that has a rail network.
[20] HIG shows whether the municipality is located in a *comarca* that has a highway network.