Modelling Location Decisions: The role of R&D activities

Isabel Mota, António Brandão
CETE - Faculdade de Economia, Universidade do Porto

June 15, 2005

Abstract (Preliminary version)

In this essay, we intend to evaluate the importance of R&D activities for firms' decision about location. For that purpose, we make use of micro-level data for the Portuguese manufacturing sector and focus on the location choices made by new starting firms during 1992-2000 within 275 municipalities. We consider two samples: the first one includes the entire manufacturing sector, while the second one restricts for the R&D intensive industries. The set of explanatory variables includes technological variables, such as R&D expenditures and human capital stock, as well as other explanatory variables that are traditionally stressed by urban and regional theory, such as production costs, demand indicators and agglomeration economies (urbanization and localization economies). The model is based on the random utility maximization framework and proceeds through a Poisson and a Negative Binomial regression. From our results, we were able to conclude that for the total manufacturing sector, the main determinants for firms' location decision were the labor and land costs and agglomeration economies. However, when considering the R&D intensive sample, those traditional location determinants lose importance, whilst the R&D expenditures become relevant.

Keywords: Location; R&D; Poisson model; Negative Binomial model

JEL Classification: R12; C25

1 Introduction

Ever since Marshall (1920 [1890]), it is widely accepted that firms gain from their joint location because they benefit from economies on the transport of goods, people and ideas. However, if firms are rivals in the product market, geographical proximity makes competition between them fiercer and this acts as a centrifugal force. Obviously, the outcome of both centripetal and centrifugal forces depends on their relative strengths.

The purpose of this research is to evaluate if firms' decision about location is related with the location of R&D activities, which is an evidence of the existence of spillovers of these activities.

There are two strands of the literature that focus on the location choice as a strategic decision to access and absorb knowledge spillovers. The first tradition appeals to the concept of geographically mediated spillovers and includes a geographic dimension to the determinants of innovation. Usually, it draws on the production function approach and uses some measure of innovation as the dependent variable against a set of possible explanatory variables. The most prominent work on this tradition is due to Jaffe (1989), who modified Griliches (1979)' knowledge production function to include a geographical dimension, and since then a lot a scientific production emerged. Other methodologies involve paper trails, case studies or surveys.

In general, empirical studies on location and technological externalities stress the importance of spatial proximity for knowledge transmission. In effect, most studies conclude for the strong propensity for the clustering of innovation-related activities, although it can vary with the type of recipient's firm (**?**) or with firms' industrial sector (Audretsch and Feldman (1996)). There is also a profusion of works that aim at identifying the nature and origins of the spillovers effects, usually distinguishing between corporate R&D labs (internal or external to firms), university R&D labs and other public research organizations (Varga (2000) and Arundel and Geuna (2001)). Usually, the superiority of university R&D for the emergence of knowledge spillovers is revealed (Jaffe *et al.* (1993)), while its importance is recognized to vary with firms' industrial sector (Anselin *et al.* (2000)).

A second approach for the study of strategic location decision focus on the determinants of the location choice to start new firms. Usually, it appeals to econometric models with a discrete depend variable - the location of a new firm - which is put against a set of explanatory variables that aim at describe location determinants, among which we may include R&D activities. Most works focus on University R&D activities and conclude for its relevance as attracting the location of new firms, whereas its relevance may vary with firms' R&D intensity (e.g. Audretsch *et al.* (2003) and Woodward *et al.* (2003)).

According to Greenhut (1993), contributes for the location theory may be aggregated into three categories. The cost minimizing theory claims that the optimal location corresponds to the local where the production and transportation costs are minimum (e.g. Weber (1928 [1909]) and Isard (1956)). The spatial interaction theory postulates that the production costs are irrelevant and that the optimal location results from the determination of the optimal market area in a context of spatial competition (e.g. Hotelling (1929)). Finally, the profit maximization theory suggests that the optimal location depends both on the costs and revenues that derives from each location (e.g. Lösch (1954)).

In summary, the optimal location theory provides three categories of location determinants: production costs, market and agglomeration economies. Production costs usually includes land, labor and capital costs. On the demand side, the market size and its acessebility are usualy considered. Agglomeration economies reflects the benefits that results from the general development of an industry (Marshall (1920 [1890])). The now standard classification of Marshallian externalities is attributed to Hoover (1936) and distinguishes between *urbanisation economies*, which reflect the benefits from operating in large population centres with correspondingly large overall labour markets and large, diversified service sectors to interact with manufacturing, and *localisation economies* that reflect economies of intra-industry specialisation that allow a finer division of function among firms, labour market economies that reduce search costs for firm seeking workers with specific training and communication economies that can speed up adoption of innovations. Black and Henderson (1999) introduced Marshall's ideas that agglomeration economies are the result of positive spillovers between firms that share the same location (localisation economies). Henderson's view can also be related with the Marshall - Arrow - Romer (MAR) argument according to which the existence of increasing returns and learning by doing causes industries to agglomerate in particular areas (Glaeser *et al.* (1992)). However, according to Jacobs (1969), the most important knowledge transfers come from outside the core industry (urbanization economies). As a result, variety and diversity of geographically proximate industries rather than geographical specialization promotes innovation and growth.

Recently, most studies on the location decision analyses the actual location of firms, instead of optimal location, as it allows to determine why a firm or a set of firms choose a specific location. In this essay, we depart from this approach and introduce a set of technological variables, such as R&D expenditures and human capital, as determinants of firms' location. Our main hypothesis is that the location and size of R&D activities affect firms decision about location. For that purpose, we will focus on the Portuguese experience and proceed through an econometric study based on the Random Utility Maximization framework.

The remainder of the paper is organized as follows. Next section is devoted to a brief explanation of the Random Utility Maximization framework. Next, we proceed with a detailed description of data and variables we considered in our study. Finally, we present some empirical results and concluding remarks.

2 Methodology

Research on firms' decision about location usually appeals to discrete-choice models that rely on the Random Utility Maximization framework of McFadden (1974). This methodology was first implemented on location choice by Carlton (1983) and Bartik (1985). Most subsequent research on spatial probability choice has relied on these approaches (for instance, Coughlin *et al.* (1991), Friedman *et al.* (1992), Guimarães *et al.* (2004) and Woodward *et al.* (2003))

In this framework, decision probabilities are modelled in a partial equilibrium setting where firms maximize profits subject to uncertainty that derives from unobservable characteristics. For our purposes, we will consider an economy with K industrial sectors ($k = 1, ..., K$). Assume that there are N investors ($i = 1, ..., N$) who independently select a location j from a set of J potential locations ($j = 1, ..., J$). The potential profit that a firm i assigns to each location j and each industrial sector k is:

$$\pi_{ijk} = \boldsymbol{\alpha}' x_j + \boldsymbol{\theta}' y_k + \boldsymbol{\beta}' z_{jk} + \varepsilon_{ijk}$$

where $\boldsymbol{\alpha}$, $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are vectors of unknown parameters, x_j is a vector of location specific variables, y_k is a vector of sector specific variables and z_{jk} is a vector of variables that change simultaneously with the sector and the location. ε_{ijk} is an identically and independently distributed random term with an Extreme Value Type I distribution.

For every spatial option, the investor will compare expected profits and choose alternative r if:

$$\pi_{irk} > \pi_{ijk}, \forall j \neq r$$

Due to the stochastic nature of the profit function, the probability of an investor i of the industrial sector k chooses the location j is:

$$P(j) = Prob(\pi_{irk} > \pi_{ijk})$$

Or, similarly;

$$P_{j|k} = \frac{\exp\left(\boldsymbol{\alpha}' x_j + \boldsymbol{\beta}' z_{jk}\right)}{\sum_{j=1}^{J} \exp\left(\boldsymbol{\alpha}' x_j + \boldsymbol{\beta}' z_{jk}\right)}$$

which expresses the conditional logit model formulation. Estimation is carried out by maximizing the log-likelihood:

$$log\ L = \sum_{k=1}^{K} \sum_{j=1}^{J} n_{kj}\ \log P_{j|k}$$

where n_{kj} denotes the number of investments carried out in sector k and region j. However, and according to Guimarães *et al.* (2003), the conditional logit model may be inadequate when we has to handle with a large number of spatial alternatives. For this reason, they proposed a tractable solution that adopts the Poisson regression and takes n_{jk} as the dependent variable and includes x_j and z_{jk} plus a set of dummy variables for each sector as explanatory variables. By this procedure, the coefficients of the conditional logit model can be equivalently estimated using a Poisson regression. The Poisson model is then adequate when the dependent variable is a count variable - the number of firms that choose location j - allowing to obviate problems related with nil observations.

3 Data and Variables

3.1 Dependent variable

We used *Quadros do Pessoal* (*MTE ((1991-2000))* to identify all plant births in each municipality (*concelho*) between 1992 and 2000. This statistical database is a yearly survey of the Portuguese Ministery of Employment for all companies operating in the country except family business without wage-earning employees that collects information at the firm and plant level since 1982, with a special emphasis on the characteristics of the labor force. By using a unique identifying number addicted to each firm and its establishments and employees, we were able to merge data about firms, plants and labor force. However, this identifying number was modified in 1991, leading us to limit our study to the period 1991 to 2000. Finally, we recur to the *Portuguese Classification of Economic Activities* at two digit level (CAE - 15 to 37) (INE (1994)) to restrict for the manufacturing sector. We also considered the *Code of the Administrative Division* (INE (1987)) to select the 275 municipalities (*concelhos)* [1] [2].

A plant was identified as new if it was the first time it appeared in the merged data set[3]. We were able to identify 38 479 new plants between 1992 and 2000. The sectorial and spatial distribution of these newly created establishments is presented in tables 1 and 2. As we can observe, the most dynamic industrial branches correspond to the

[1] In the construction of our database, we had to account with a change of the *Portuguese Classification of Economic Activities* (1994) and a change of the *Code of the Administrative Division* (1998).

[2] We excluded the islands of Azores and Madeira, as the number of new plants of the manufacturing sector created during 1992-2000 was quite small.

[3] We exclude possible temporary exits/errors by comparing the age of the firm with the age of the oldest employee.

manufacturing of wearing apparel, dressing and dyeing of fur and the *manufacturing of fabricated metal products, except machinery and equipment.*

CAE - Rev. 2	Manufaturing	Plant births (1992-2000) Total	%
15	Manufacture of food products and beverages	4.164	10,82%
16	Manufacture of tobacco	0	0,00%
17	Manufacture of textile	2.522	6,56%
18	Manufacture of wearing apparel; dressing and dyeing of fur	7.595	19,74%
19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	2.156	5,60%
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of straw and plaiting materials	3.527	9,17%
21	Manufacture of paper and paper products	273	0,71%
22	Publishing, printing and reproduction of recorded media	2.230	5,80%
23	Manufacture of coke, refined petroleum products and nuclear fuel	4	0,01%
24	Manufacture of chemicals and chemical products	415	1,08%
25	Manufacture of rubber and plastics products	531	1,38%
26	Manufacture of other non-metallic mineral products	2.582	6,71%
27	Manufacture of basic metals	179	0,47%
28	Manufacture of fabricated metal products, except machinery and equipment	5.678	14,76%
29	Manufacture of machinery and equipment n.e.c.	1.503	3,91%
30	Manufacture of office, accounting and computing machinery	3	0,01%
31	Manufacture of electrical machinery and apparatus n.e.c.	393	1,02%
32	Manufacture of radio, television and communication equipment and apparatus	131	0,34%
33	Manufacture of medical, precision and optical instruments, watches and clocks	287	0,75%
34	Manufacture of motor vehicles, trailers and semi-trailers	171	0,45%
35	Manufacture of other transport equipment	205	0,53%
36	Manufacture of furniture; manufacturing n.e.c.	3.801	9,88%
37	Recycling	118	0,31%
Total		38.479	100,00%

Source: MTE (1991-2000), *Quadros do Pessoal*

Similarly, as we can observe in table 2, the most dynamic region (NUT2) is *Região Norte*, which account for more than 50% of total industrial plant births between 1991 and 2000.

NUT3	Designation	Total	%
10101	Minho-Lima	898	2.33%
10102	Cávado	3 131	8.14%
10103	Ave	5 290	13.75%
10104	Grande Porto	4 859	12.63%
10105	Tâmega	4 105	10.67%
10106	Entre Douro e Vouga	2 319	6.03%
10107	Douro	413	1.07%
10108	Alto Trás-os-Montes	515	1.34%
	Região Norte	*21 530*	*55.95%*
10201	Baixo Vouga	1 527	3.97%
10202	Baixo Mondego	707	1.84%
10203	Pinhal Litoral	1 382	3.59%
10204	Pinhal Interior Norte	475	1.23%
10205	Dão-Lafões	851	2.21%
10206	Pinhal Interior Sul	147	0.38%
10207	Serra da Estrela	156	0.41%
10208	Beira Interior Norte	283	0.74%
10209	Beira Interior Sul	211	0.55%
10210	Cova da Beira	276	0.72%
	Região Centro	*6 015*	*15.63%*
10301	Oeste	1 417	3.68%
10302	Grande Lisboa	4 154	10.80%
10303	Península de Setúbal	1 665	4.33%
10304	Médio Tejo	737	1.92%
10305	Lezíria do Tejo	754	1.96%
	Lisboa e Vale do Tejo	*8 727*	*22.68%*
10401	Alentejo Litoral	254	0.66%
10402	Alto Alentejo	324	0.84%
10403	Alentejo Central	569	1.48%
10404	Baixo Alentejo	291	0.76%
	Alentejo	*1 438*	*3.74%*
10501	Algarve	769	2.00%
	Algarve	*769*	*2.00%*
	Total	38 479	100.00%

Source: MTE (1991-2000), *Quadros do Pessoal*

3.2 Explanatory variables

Location determinants usually include both geographical and economic characteristics that affect firms' expected profits. On the supply side, we will consider both land, labor and capital costs. On the demand side, we will consider the market size and its acessibility. As it is typical in location theory, we will consider agglomeration economies that includes both localization and urbanization economies. Finally, to account for technological variables, we include both R&D expenditures and human capital.

Following Bartik (1985)'s approach that assumes that industrial and residential users compete for the same space, we used population density as a proxy for *land costs* (INE (1991 - 2000), INE (2001) and INE (2003b)). To account for *labour costs* in each concelho and manufacturing sector, we recur to the wages per working hour by plant[4], for each municipality and CAE (MTE ((1991-2000)). *Capital costs* are measured by the taxes over companies collected by municipalities, that included both *derrama*[5] and other taxes over firms (DGAA (1991 - 2001) and INE (1992-2001)). The *market size* is approximated by the Purchasing Power Index for each municipality [6] (INE (1993 - 2002)). Finally, we considered the market accessibility of each municipality as measured by the minor physical distance between each municipality to Porto or Lisbon, the most important cities in Portugal (INE (2003a)).

In what concerns agglomeration economies, literature usually distinguishes between urbanization economies, which are external to firms and industries but internal to a city, and localization economies, which are external to firms but internal to an industry. To account for *urbanization economies*, we considered the density of manufacturing and service plants (CAE D, G, H, I, J, K) per square kilometer in each municipally (MTE ((1991-2000) and INE (2003b)). To account for *localization economies*, we considered the share of manufacturing employment for each CAE - 2 digit in each municipally (MTE ((1991-2000)).

Finally, technological variables includes both human capital and R&D activities. To account for the *human capital* stock in each municipally, we computed the average years of schooling of the active population for each municipality, according to the methodology of Barro and Lee (1993):

$$H = \frac{1}{Pop}(D_1\left(h_{1c} + h_{2f}\right) + \frac{D_1}{2}h_{1i} + D_2\left(h_{2c} + h_{3f}\right) + \frac{D_2+D_1}{2}h_{2i} + D_3\left(h_{3c} + hsf\right) + \frac{D_3+D_2}{2}h_{3i} +$$

$$D_s\left(h_{sc} + h_{mi} + h_{Bf} + h_{Lf}\right) + \frac{D_s+D_3}{2}h_{si} + (D_s+1)h_{mc} + D_B h_{Bc} + \frac{D_B+D_s}{2}h_{Bi} +$$

$$D_L\left(h_{Lc} + h_{Mi} + h_{Mf} + h_{Df}\right) + \frac{D_L+D_s}{2}h_{Li} + \left(D_L+1\right)h_{Mc} + \left(D_L+2\right)h_{Dc} + \left(D_L+1\right)h_{Di})$$

[4] We considered the base wage and regular working hours.

[5] The municipal surcharge (*derrama*) is a local municipal tax that can be charged annually by municipal authorities up to maximum of 10% of the amount paid in corporate tax (IRC)

[6] The Purchasing Power Index (IPPC) intends to capture the purchasing power in each municipality. It is built by means of a model of factorial analysis and recurring to 20 variables.

where

h_j = Population with age 25 to 64 years with the education level j, where

$j = 1$ (1st cycle), 2 (2nd cycle or ISCED 1), 3 (3rd cycle or ISCED 2), s (Secondary or ISCED 3), m (ISCED 4), B (ISCED 5B), L(High school or ISCED 5A), M (Master) e D (PHD); c = Complete, i = Incomplete and f = Frequency (INE (1991-2001)).

D_i = Duration, in years, of each education level i: $D_1 = 4, D_2 = 6, D_3 = 9, D_s = 12, D_B = 15, D_L = 16, D_m = 13$.

Pop = Population with age 25 to 64 years (INE (1991-2001)).

After computing the human capital stock for each municipality in 1991 and 2001, and due to its structural nature, we calculated the annual average rate of growth between 1991 and 2001 and determined the human capital stock between 1991 and 2001.

With respect to the *R&D* variable, we recur to total R&D expenditures in each municipality per habitant for years 1995 to 2001 by using a biannual national inquiry (OCES (1995, 1997, 1999, 2001)). For the years with unavailable information, we used the average of the nearest years.

Next table summarizes information about explanatory variables:

Variable	Designation	Indicator	Expected sign	Data Source
Labour cost	WH	Base wage plus regular working hours, per working hour, by *concelho* and CAE	Negative	Ministério do Trabalho e do Emprego (MTE), *Quadros do Pessoal* (1991-2000)
Localization economies	ELOCNPS	Share of manufacturing employment for each CAE - 2 digit, by *concelho*	Positive	MTE, *Quadros do Pessoal* (1991-2000)
Urbanization economies	EURBN	Density of manufacturing and service plants (CAE D, G, H, I, J, K) per square kilometer, by *concelho*	Positive	MTE, *Quadros do Pessoal* (1991-2000); INE - *REFTER* (2003)
Land cost	DP	Population density, by *concelho*	Negative	Instituto Nacional de Estatística (INE), *Estimativas da População Residente* (1991 - 2000), *Censos* (2001), *Referenciação Territorial* (*REFTER) (*2003)
Capital cost	DT	Derrama plus other taxes over firms, by total firms in each *concelho*	Negative	Direcção-Geral das Autarquias Locais (DGAA), *Finanças Municipais* (1991 - 2001); INE, *Ficheiro Central de Empresas e Estabelecimentos (FUE) (*1992-2001)
Market size	M	Purchasing Power Index (IPPC), by concelho	Positive	INE, *Estudo sobre o Poder de Compra Concelhio* (1993 - 2002)
Human capital	H	Average years of shooling of the active population, by *concelho*	Positive	INE, *Censos* (1991, 2001)
Research & Development	DID	R&D expenditures per capita, by concelho	Positive	Observatório da Ciência e Ensino Superior (OCES), *Inquérito ao Potencial Científico e Teconlógico* (IPCTN) (1995, 1997, 1999, 2001); INE, *Estimativas da População Residente* (1991 - 2000); *Censos* (2001)
Market accessibility	DPL	Minimum distance to Porto/Lisbon, by *concelho*	Negative	INE, *Base Geográfica de Referenciação da Informação* (BGRI), 2003)

Table 3: Explanatory variables

In order to evaluate the impact of R&D activities on firms' decision about location, we considered two data-bases. The first one includes the entire manufacturing sector, while the second one restricts for those industrial branches that cumulatively account for 50% of R&D Total Employment (OCES (1995, 1997, 1999, 2001)): CAE 15, 17, 23, 24, 29, 30, 31, 32, 33, 34.

4 Empirical Results

We modeled the location choice of new manufacturing plants within 275 municipalities through a Poisson estimation. In table 4 we present the main results from our econometric analysis, considering both total manufacturing sector and R&D intensive sample. Several regression models were developed, which performed quite well, as we can observe from the chi-square statistitics for the likelihood ratio test of overall significance.

The regression (1) respects to a standard Poisson regression. As we can observe, all variables are highly significant and with the expected sign, except for the market size and the human capital. When we include dummy variables by CAE and CAE/NUT3 (regressions 2 and 3, respectively), both market size and human capital remain not significant or without the expected sign. Additionally, the capital cost becomes not significant. When considering the R&D intensive sample (regressions 4 to 6), we find a similar result: the human capital variable has not the expected sign. On the contrary, the market size has now the expected sign (or it is not significant), whilst market acessibility is not significant.

In both samples, the most significant determinants for firms' decision about location are the labour and land costs, market accessibility and the agglomeration economies, particularly, the urbanization ones. Additionally, the R&D variable as a minor impact on attracting firms, but larger when talking about the R&D intensive sector. In fact, we estimated that, everything else constant, a 1 percent increase in R&D expenditures led to an 0.083 percent increase in the number of new plant births in the total manufacturing sample, while the same elasticity is equal to 0.098 if we consider the R&D intensive sample.

	Total manufacturing plants			R&D intensive manufacturing plants		
	Simple	**With dummy (Year* CAE)**	**With dummies (Year *CAE, Nut3)**	**Simple**	**With dummy (Year* CAE)**	**With dummies (Year *CAE, Nut3)**
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	birth	birth	birth	birth	birth	birth
Ln WH	-1.852	-1.169	-1.206	-1.687	-1.018	-0.910
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln ELOCNPS	0.711	0.724	0.725	0.646	0.606	0.528
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln EURBN	0.988	0.959	1.184	0.875	0.783	1.038
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln DP	-0.621	-0.504	-1.126	-0.453	-0.323	-0.876
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln DT	-0.024	0.014	0.008	-0.039	-0.006	-0.009
	(0.000)	(0.128)	(0.437)	(0.023)	(0.742)	(0.664)
Ln M	-0.310	-0.173	0.020	0.052	0.148	0.405
	(0.008)	(0.000)	(0.594)	(0.350)	(0.030)	(0.000)
Ln H	-0.536	-1.474	-0.345	-0.835	-1.370	-0.853
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)
Ln DID	0.083	0.096	0.050	0.098	0.102	0.065
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln DPL	-0.255	-0.165	-0.221	-0.023	0.106	0.010
	(0.000)	(0.000)	(0.000)	(0.407)	(0.707)	(0.839)
Constant	19.883	15.655	16.188	15.860	11.579	11.160
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of obs.	11871	11871	11871	4493	4493	4493
Log likelihood	-24719.196	-21627.306	-19295.056	-6791.113	-6252.070	-5686.146
LR test	37652.300	43836.080	48500.580	7400.110	8478.200	9610.040
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Goodness of fit	32916.680	26732.900	22068.400	8233.034	7154.947	6023.100
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are in parentisis

Table 4: Location determinants of new manufacturing plants (1991-2000): Poisson Regression

However, the adjustment quality tests indicates that the Poisson regression is not adequate to our data base, suggesting that we should try the negative binomial model. Tables 5 and 6 resumes main results from our estimation. In both cases, we ran a simple Negative Binomial with or without dummies. Then, we performed a Negative Binomial for panel data with random and fixed effects by municipality (*concelho*).

When taking under consideration total new starting firms, and focusing on the panel data estimation, we find that the most significant determinants for firms' decision about location are the the traditional ones: labour and land costs, agglomeration economies and market acessibility. Market size and human capital are only significant and with the expected sign when the dummy for NUT3 is not included. On the contrary, R&D expenditures are statisticaly significant and with the expected sign if the dummy for NUT3 is included, although its elasticity is low. In fact, we estimate that, everything else constant, a 1 percent increase in R&D expenditures led to an 0.047 percent increase in the number of new plant births, while the same elasticity for labor cost is 0.90, approximately.

	Simple	With dummy (Year* CAE)	With dummies (Year *CAE, Nut3)	With random effects by "concelho" and dummy by Year*CAE	With random effects by "concelho" and dummies by Year*CAE and Nut3	With fixed effects by "concelho" and dummy by Year*CAE	With fixed effects by "concelho" and dummies by Year*CAE and Nut3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	birth	birth	birth	birth	birth	birth	birth
Ln WH	-1.742	-1.048	-1.083	-0.919	-0.919	-0.897	-0.894
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln ELOCNPS	0.633	0.563	0.567	0.576	0.549	0.559	0.532
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln EURBN	0.898	0.939	1.069	0.556	0.865	0.402	0.771
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln DP	-0.597	-0.570	-1.035	-0.572	-0.741	-0.672	-0.701
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln DT	-0.024	-0.016	-0.008	-0.025	-0.032	-0.034	-0.035
	(0.157)	(0.304)	(0.625)	(0.265)	(0.154)	(0.146)	(0.137)
Ln M	0.418	0.332	0.270	0.187	-0.052	0.193	-0.075
	(0.000)	(0.000)	(0.000)	(0.055)	(0.563)	(0.058)	(0.404)
Ln H	-0.413	-1.032	-0.041	2.064	0.391	2.832	0.125
	(0.003)	(0.000)	(0.834)	(0.000)	(0.273)	(0.000)	(0.744)
Ln DID	0.328	0.052	0.050	0.021	0.047	0.023	0.046
	(0.001)	(0.000)	(0.000)	(0.112)	(0.001)	(0.094)	(0.001)
Ln DPL	-0.208	-0.207	-0.239	-0.196	-0.428	-0.297	-0.715
	(0.000)	(0.000)	(0.000)	(0.009)	(0.004)	(0.001)	(0.000)
Constant	15.124	12.183	13.499	7.749	13.201	7.554	14.675
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of obs.	11871	11871	11871	11871	11871	11862	11862
Log likelihood	-17846.100	-16887.516	-16379.294	-15975.025	-15878.196	-15161.454	-15029.572
LR test	6204.640	8121.810	9138.250				
	(0.000)	(0.000)	(0.000)				
Goodness of fit							
Wald test				7824.920	8091.110	7738.280	8106.470
				(0.000)	(0.000)	(0.000)	(0.000)
Hausman tet							2.26
							(1.000)
Likelihood-ratio	14000.000	9479.580	5831.520	4317.980	3049.260		

Table 5: Location determinants of total new manufacturing plants (1991-2000): Negative Binomial Regression

When taking under consideration the R&D intensive starting firms, and focusing on the panel data estimation, we can observe that market size, human capital and capital cost are not significant. Additionally, the R&D variable is significant if a random effects model is performed. The Hausman test allow us to conclude that this specification is the correct one. So, we estimate that, everything else constant, a 1 percent increase in R&D expenditures led to an 0.050 percent increase in the number of new plant births, while the same elasticity for labor cost is now 0.57. We may then conclude that traditional location determinants lose importance when a R&D intensive sample is considered (except for urbanization economies), while the R&D variable gains importance.

	Simple	With dummy (Year* CAE)	With dummies (Year *CAE, Nut3)	With random effects by "concelho" and dummy by Year*CAE	With random effects by "concelho" and dummies by Year*CAE and Nut3	With fixed effects by "concelho" and dummy by Year*CAE	With fixed effects by "concelho" and dummies by Year*CAE and Nut3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	birth	birth	birth	birth	birth	birth	birth
Ln WH	-1.627	-0.821	-0.782	-0.607	-0.574	-0.549	-0.515
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln ELOCNPS	0.589	0.479	0.434	0.470	0.461	0.452	0.415
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln EURBN	0.783	0.815	0.965	0.883	1.130	0.666	0.977
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln DP	-0.425	-0.428	-0.876	-0.646	-0.905	-0.844	-0.277
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)	(0.522)
Ln DT	-0.037	-0.010	-0.011	-0.021	-0.009	-0.051	-0.029
	(0.176)	(0.705)	(0.689)	(0.535)	(0.800)	(0.202)	(0.466)
Ln M	0.530	0.360	0.350	0.221	0.085	0.171	-0.032
	(0.000)	(0.004)	(0.005)	(0.160)	(0.583)	(0.319)	(0.837)
Ln H	-0.742	-1.076	-0.436	0.008	-0.531	1.419	-1.557
	(0.001)	(0.000)	(0.176)	(0.986)	(0.322)	(0.025)	(0.046)
Ln DID	0.051	0.072	0.068	0.043	0.050	0.027	0.038
	(0.002)	(0.000)	(0.000)	(0.042)	(0.020)	(0.257)	(0.119)
Ln DPL	-0.047	-0.062	-0.137	-0.228	-0.306	-0.718	-3.442
	(0.319)	(0.164)	(0.049)	(0.030)	(0.109)	(0.001)	(0.000)
Constant	12.871	9.424	10.552	9.873	12.033	10.715	23.324
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of obs.	4493	4493	4493	4493	4493	4457	4457
Log likelihood	-5827.261	-5539.303	-5336.238	-5224.753	-5198.353	-4638.338	-4579.020
LR test	1975.700	2551.610	2957.740				
	(0.000)	(0.000)	(0.000)				
Wald test				2815.330	2830.450	2761.920	2954.950
				(0.000)	(0.000)	(0.000)	(0.000)
Hausman tet							0.001
							(1.000)
Likelihood-ratio test of alpha=0				1065.850	543.150	---	---
				(0.000)	(0.000)	---	---

Notes: p-values are in parentisis

Table 6: Location determinants of new R&D intensive manufacturing plants (1991-2000): Negative Binomial Regression

5 Concluding remarks

In this article, we exploit the importance of traditional and technological determinants for firms' decision about location. As it was expected, localization economies and, in particular, urbanization economies reveal to be extremely important for firms' decision about location, which accords with existing literature. Additionally, cost variables were also statistically significant. Particularly, labour costs have one of the highest elasticities, whilst it reduces when taking under consideration a R&D intensive sample.

On the contrary, capital cost was never statistically significant, which can be justified by the nature of the variable: in Portugal, there are no significant differences in the cost of capital within municipalities. On the demand side, market accessibility is relevant for firms' decision about location, while market size loses relevance, particularly in the R&D intensive sample. Again, the undersized of most municipalities justify the lack of importance of this variable, whilst the proximity with the most important cities in Portugal reveals to be extremely relevant for location decisions.

Our main goal was to study the importance of R&D activities for the location decision of firms. As we observed, the elasticity of plant births with respect to R&D expenditures was quite small, whilst it becomes higher when taking under consideration the R&D intensive sample. On the contrary, the human capital variable has an irregular behavior, ranging from statistical insignificance to unexpected sign. This could be justified by the nature of industrial sector (not intensive in human capital) but also by the aggregate measure of human capital stock, which do not allow to evaluate the importance of some specific skills (e.g. engineers) for the location of firms.

As a preliminary version, we expect to improve this research in two ways: first, by experimenting different samples with uneven R&D intensities; second, by recurring to gravitational variables (e.g. market size).

References

Anselin, L., A. Varga and Z. J. Acs (2000), "Geographical spillovers and University Research: A spatial econometric perspective", *Growth and Change*, vol. 31, pp. 501–515.

Arundel, A. and A. Geuna (2001), "Does proximity matter for knowledge transfer from public institutes and universities to firms?", Working paper 73, SPRU.

Audretsch, D. B. and M. P. Feldman (1996), "R&D spillovers and the Geography of innovation and production", *American Economic Review*, vol. 86, 3, pp. 630–640.

Audretsch, D. B., E. E. Lehmann and S. Warning (2003), "University spillovers: Strategic location and new firm performance", Discussion Paper 3837, CORE.

Barro, R. and J.-W. Lee (1993), “International comparisions of educational attainment”, *Journal of Monetary Economics*, vol. 32, pp. 363–394.

Bartik, T. J. (1985), “Business Location Decisions in the US: Estimates of the Effects of Unionizaon, Taxes and Other Characteristics of States”, *Journal of Business and Economic Statistics*, vol. 3, 1, pp. 14–22.

Black, D. and V. Henderson (1999), “A Theory of Urban Growth”, *Journal of Political Economy*, vol. 107, 2, pp. 252–284.

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

Coughlin, C. C., J. V. Terza and V. Arromdee (1991), “State characteristics and the location of foreign direct investment within the United States”, *The Review of Economics and Statistics*, vol. 73, 4, pp. 675–683.

DGAA (1991 - 2001), “Finanças Municipais”, .

Friedman, J., D. A. Gerlowski and J. Silberman (1992), “What attracts foreign multinational corporations? Evidence from branch plant location in the United States”, *Journal of Regional Science*, vol. 32, 4, pp. 403–418.

Glaeser, E., H. Kallal, J. Scheikman and A. Schleifer (1992), “Growth in Cities”, *Journal of Political Economy*, vol. 100, 6, pp. 1126–51.

Greenhut, M. (1993), “Location Theory”, in *Encyclopedia of Economics*, D. Greenwald, editor, New York: McGraw Hill.

Griliches, Z. (1979), “Issues in Assessing the Contribution of R&D to Productivity”, *Bell Journal of Economics*, vol. 10, pp. 92–116.

Guimarães, P., O. Figueiredo and D. Woodward (2003), “A tractable approach to the firm location decision problem”, *The Review of Economic and Statistics*, vol. 85, 1, pp. 201–204.

Guimarães, P., O. Figueiredo and D. Woodward (2004), “Industrial Location Modeling: Extending the Random Utility Framework”, *Journal of Regional Science*, vol. 44, 1, pp. 1–20.

Hoover, E. M. (1936), *Location Theory and the Shoe and Leather Industries*, Cambridge (MA): Harvard University Press.

Hotelling, H. (1929), "Stability in competition", *Economic Journal*, vol. 39, pp. 41–57.

INE (1987), "Code of the Admnistrative Divison", .

INE (1991 - 2000), "Estimativas da População Residente", .

INE (1991-2001), "Censos", .

INE (1992-2001), "Ficheiro Central de Empresas e Estabelecimentos", .

INE (1993 - 2002), "Estudo sobre o Poder de Compra Concelhio", .

INE (1994), "Portuguese Classification of Economic Activities - Revision 2 (CAE-REV. 2)", .

INE (2001), "Censos", .

INE (2003a), "Base Geográfica de Referenciação da Informação (BGRI)", .

INE (2003b), "Referenciação Territorial (REFTER)", .

Isard, W. (1956), *Location and Space-Economy*, New York: M.I.T Press.

Jacobs, J. (1969), *The Economy of Cities*, New York: Random House.

Jaffe, A. B. (1989), "Real effects of academic research", *American Economic Review*, vol. 79, 5, pp. 957–970.

Jaffe, A. B., M. Trajtenberg and R. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations", *Quarterly Journal of Economics*, vol. 108, 3, pp. 577–598.

Lösch, A. (1954), *The Economics of Location*, New Haven: Yale University Press.

Marshall, A. (1920 [1890]), *Principles of Economics*, London (UK): Macmillan, eighteen ed.

McFadden, D. (1974), "Conditional logit analysis of qualitative choice behavior", in *Frontiers in Econometrics*, P. Zarembka, editor, chap. 4, Academic Press.

MTE ((1991-2000)), "Quadros do Pessoal", .

OCES (1995, 1997, 1999, 2001), "Inquérito ao Potencial Científico e Tecnológico", .

Varga, A. (2000), "Local academic knowledge transfers and the concentration of economic activity", *Journal of Regional Science*, vol. 40, 2, pp. 289–309.

Weber, A. (1928 [1909]), *Theory of the Location of Industry (trans. Uber den Standort des Industrien)*, Chicago: University of Chicago Press.

Woodward, D., O. Figueiredo and P. Guimarães (2003), "Beyond the Silicon Valey: University R&D and High-Technology Location", *Faculdade de Economia, Universidade do Porto*, , 133, p. Working Paper.