



## The determinants of location choice: Single plants versus multi-plants\*

Isabel Mota<sup>1</sup>, António Brandão<sup>1</sup>

<sup>1</sup>Faculdade de Economia, Fundação da Universidade do Porto and CEF.UP, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal (e-mail: imota@fep.up.pt, abrandao@fep.up.pt)

Received: 23 September 2009 / Accepted: 8 August 2011

**Abstract.** This paper adopts a count data model to explore the distinction between single plant and multi-plant location choices. It is hypothesized that start-up location decisions would be determined by supply variables (land, labour and capital costs, workforce and technological characteristics); demand variables (market size and market accessibility) and agglomeration economies. We use plant data and focus on location choices within Portuguese municipalities. Our research shows that new multi-plants are particularly sensitive to urbanization economies, land costs and the size of the local market, while new single plants are more responsive to labour costs, both localization and urbanization economies and accessibility to main markets.

**JEL classification:** R30, C25, L60

**Key words:** Location, count model, single plants, multi-plants

### 1 Introduction

Understanding the determinants of business location choice is the subject of a large body of literature, comprising both theoretical and empirical research. In fact, alongside renowned contributions to optimal location theory that claim the importance of cost factors, demand variables and agglomeration economies in location decisions, several empirical studies have examined the relative significance of different factors in the business site selection process (e.g., Carlton 1983; Bartik 1985; Coughlin et al. 1991). However, literature is scarce concerning firms' characteristics that might influence location decision. Specifically, the level of structural complexity of a firm might influence its access to information and degree of risk aversion and therefore, its location choice.

---

\* We wish to thank J. Varejão and E. Strobl for their helpful suggestions, the participants of the *Workshop in regional agglomeration, growth and multilevel governance* and three anonymous referees for their comments. Furthermore, we gratefully acknowledge the Portuguese Ministry of Employment – Statistics Department, the National Institute of Statistics and the Portuguese Ministry of Science, Technology and Higher Education for access to the data used. CEF.UP is funded by Fundação para a Ciência e a Tecnologia.

Early research on location choice assumed that firms were identical in features such as size, technological and sectorial characteristics (e.g., Bartik 1985; Hansen 1987; Schmenner et al. 1987). Usually, these studies evidenced the relevance of the agglomeration economies and cost variables for the location choice of new firms. Following these initial studies, some authors have stressed the importance of the capital structure by focusing on the location decisions made by foreign firms (e.g., Bartik 1985; Coughlin et al. 1991; Friedman et al. 1992; Woodward 1992; Head and Mayer 2004; Basile et al. 2009). They found that agglomeration economies (particularly, urbanization economies), market accessibility and the existence of institutional support for foreign investments were relevant, while the importance of labour and land costs was inconclusive.

Recently, some authors have claimed that the location determinants affect firms' location choice differently according to their characteristics, namely firm's size (e.g., Arauzo-Carod and Manjón-Antolin 2004), its industrial activity (e.g., Arauzo-Carod 2005), or technological intensity (e.g., Barrios et al. 2006), and the entrepreneur's preference for the home base (e.g., Figueiredo et al. 2002). Additionally, researchers claim that new and relocated firms seem to have different location profiles (Holl 2004a; Lee 2006), while regional competition within and across nations seems to influence firms' location (Basile et al. 2009). Although some general tendencies were still in evidence (e.g., the relevance of agglomeration economies for location choice), these studies showed that some location determinants have a quite different influence on location choice according to firms' characteristics.

The purpose of this study is to evaluate the importance of location determinants for firms' decisions about location, but by assuming that firms behave differently according to their structure. Particularly, we aim at evaluating whether location determinants affect single plants and multi-plants differently. To the best of our knowledge, this topic has not been addressed in the literature but deserves particular attention as the spatial context where new single plants and new multi-plants make their location choices is dissimilar.<sup>1</sup> In fact, while the location problem of new single plants and new multi-plants is the same – the search for the highest potential profit providing site – the relative importance of each location determinant may vary across both types of plants, for several reasons. First, new single plants and new multi-plants are born at different phases in the life cycle of firms. Therefore, new multi-plants might take advantage of a more mature entrepreneurship than new single plants. Second, the birth of a multi-plant is frequently motivated by capacity needs or market expansion,<sup>2</sup> and so they might benefit from multi-plant economies. Finally, location decisions are taken on the grounds of incomplete information and uncertainty. Hence, we might expect that multi-plants have access to more information about sites when they make their location choices; in addition, and according to Levinson (1996), multi-plants benefit from economies of scale in conducting site searches. Furthermore, multi-plants are less risk-averse than single plants as they act in several markets and, therefore, the idiosyncratic risk is null. Lastly, multi-plants may be more footloose than single plants as, according to Levinson (1996) and Figueiredo et al. (2002), single plants are more likely to open where the entrepreneur lives.

---

<sup>1</sup> Even without specifically focusing on the location choices of single plants and multi-plants, Levinson (1996) studied the effects of state environmental regulation on establishment location choices and recognized that branch plants of large firms appear to be more sensitive to local and environmental conditions than all plants in general; he then concentrated on the location choices of branch plants of large firms.

<sup>2</sup> This argument is also valid in the relocation of plants, that is, when a firm shuts down a plant and starts over in a new space (Lee 2006). According to Pellenbarg et al. (2002), the relocation of plants might be motivated by changes in markets, preferences in consumers, environmental regulations and technological progress, while Holl (2004a) calls attention to the capacity needs and physical space constraints, demand growth and market expansion, and restructuring to respond to new market opportunities as the factors that induce the relocation of firms.

So, the search for the optimal location is costly, but it is lower for new multi-plants. Additionally, this distinction is also relevant from the standpoint of regional politicians, as the effort devoted to attract multi-plant's branches or to promote local single plants' entrepreneurs is rather different.

Using micro-level data on Portuguese manufacturing plants, this paper examines the location choices made by new single plants and multi-plants within 275 municipalities between 1992 and 2007. We apply a discrete-choice model that relies on the random utility maximization framework (McFadden 1974), but proceeds through a count model where the probability of a new plant locating in a municipality relates to a set of potential location factors. As in past studies about firms' location, the set of explanatory variables includes variables that are traditionally stressed by urban and regional theory, such as production costs (land, labour and capital costs), demand indicators and agglomeration economies (urbanization and localization economies), as well as technological variables, such as the R&D expenditures.

From our results, we conclude that although the main location determinants are the agglomeration economies and both land and labour costs, there are some noteworthy differences concerning multi-plants and single plants location choices. In fact, we observed that new multi-plants are particularly sensitive to urbanization economies, land costs and local market, while new single plants are more responsive to labour costs, both localization and urbanization economies and market accessibility.

The rest of the paper is organized as follows. Section 2 presents the model specification, followed by a discussion of data in Section 3. Section 4 provides the main estimation results and the concluding remarks are presented in Section 5.

## 2 The model

Research on firms' decisions 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 popularized in articles by Bartik (1985), Coughlin et al. (1991), Figueiredo et al. (2002) and Head et al. (1999), among others.

Under the conditional logit formulation, firms maximize profits subject to uncertainty that derives from unobservable characteristics when choosing location. However, the conditional logit model assumes that the odds of choosing an alternative are a function of its attributes but are independent of other alternatives, which is known as the independence of irrelevant alternatives (IIA) assumption. This proposition may be implausible in location choice, as adjacent locations may have similar characteristics, which make them interdependent. Moreover, if the IIA assumption is violated, biased coefficient estimates occur.

In order to accommodate the IIA assumption in the location choice analysis, some procedures have been developed. One possibility is to recur to the nested logit model, which implies a partition of the spatial choice set into subgroups (e.g., Head and Mayer 2004; Barrios et al. 2006; Basile et al. 2009). However, this assumes that the IIA assumption holds within each subgroup, which might be implausible.

A recent strand of the literature, introduced by Guimarães et al. (2003), modelled the location choice by means of a Poisson model, as the coefficients of the conditional logit model can be equivalently estimated by using a Poisson regression.<sup>3</sup> Under this formulation, the number of new firms that choose a specific location is a count variable and relates to a vector of local and sectorial characteristics.

<sup>3</sup> According to Schmidheiny and Brülhart (2009), the identical estimates from the Poisson and the conditional logit model have different economic implications. Accordingly, in the conditional logit model an additional agent attracted to a specific location means one less agent among the other locations in the relevant set, while in the Poisson model, that

It should be said that the Poisson regression model assumes that the conditional mean  $\lambda_j$  equals the conditional variance. However, the variance is frequently larger than the one assumed by the Poisson model, a result called overdispersion, which causes a type of heteroscedasticity as it yields downward biased estimates of the standard errors.

In order to account for overdispersion, the negative binomial model is usually adopted, which can be motivated as a mixture of Poisson distribution, where the mean  $\lambda_j$  follows the gamma distribution (Greene 2000). Therefore, the probability that the number of firms choose a specific location  $j$  ( $n_j$ ) is given by:

$$P(n_j) = \frac{\Gamma(\theta + n_j)}{\Gamma(1 + n_j)\Gamma(\theta)} \left( \frac{\lambda_j}{\lambda_j + \theta} \right)^{n_j} \left( \frac{\theta}{\lambda_j + \theta} \right)^\theta$$

Additionally, the negative binomial model may be estimated by maximum likelihood (e.g., Coughlin and Segev 2000; Holl 2004a).

Since Guimarães et al. (2003) most research on firms' location follows this approach and adopts the Poisson regression. This model finds its roots on the random utility maximization framework but takes advantage of the equivalence between the conditional logit model and the Poisson regression, which allows us to overcome a potential IIA violation (e.g., Arauzo-Carod and Manjón Antolín 2004; Barbosa et al. 2004; Basile 2004; Guimarães et al. 2004; Holl 2004a; Arauzo-Carod 2005). In cases of overdispersion, the negative binomial model is frequently employed (e.g., Barbosa et al. 2004; Holl 2004b; Autant-Bernard et al. 2006), being other possibilities the ZI Poisson or negative binomial in case of zero inflation data.<sup>4</sup> In this paper, we follow this approach and apply the Poisson regression. However, preliminary results evidenced overdispersion and, therefore, we recurred to the negative binomial model.

### 3 Data

#### 3.1 Plant births

Plant births from 1992 and 2007 are calculated from *Quadros do Pessoal* (DEEP–MTE, 1991–2007). This statistical database is built on a compulsory survey collected annually by the Portuguese Ministry of Employment for all business firms operating in Portugal.<sup>5</sup> The inquiry has collected information at the firm, plant and worker level since 1982, including data on firms' location, economic activity, capital structure or number of plants and gives a particular emphasis to the characteristics of the workforce.

Besides its reliability, the database has a longitudinal dimension that makes it particularly suitable for studying the topic of plants' birth. By using a unique identifying number attached to each firm and their establishments and employees, it was possible to merge data about firms, plants and labour force. Given that this identifying number was modified in 1991, we have limited the study to the period 1992 to 2007. As in previous research on location choice, only the manufacturing plants were eligible and we recurred to the Portuguese Classification of Economic Activities at two-digit level (CAE – Rev. 2–15 to 37) (INE 1994).<sup>6</sup> Additionally, the

raises the aggregate number of agents by one. Thus, the conditional logit model and the Poisson model can be viewed as polar cases, with the former representing zero-sum reallocations of firms or households across locations and the latter implying a positive-sum world.

<sup>4</sup> For a detailed description of the methods and results presented in literature of industrial location choice, please consider Arauzo-Carod et al. (2010).

<sup>5</sup> *Quadros do Pessoal* does not cover public administration or domestic service and excludes firms without wage-paid employees or start-ups with a very short life of up to 12 months.

<sup>6</sup> We considered the change of the Code of Economic Activity from Revision 1 to Revision 2 in 1994 and to Revision 2.1 in 2003.

*concelho* (or municipality)<sup>7</sup> was adopted as the geographic unit, with 275 municipalities being identified on mainland Portugal.<sup>8</sup> The reason for choosing this territorial unit relies on its relevance in Portugal: the municipalities are the layer of government just below the central government (in mainland Portugal) and, since 1974, the transference of competences to the municipalities registered a significant increase.

A plant was classified as new if it was the first time it appeared in the merged data set.<sup>9</sup> Therefore, we identify 61,177 new starting plants between 1992 and 2007, of which 89.11 per cent were new single plants and 10.89 per cent were new multi-plants. The geographical location by municipalities and NUTS 2 of these newly created establishments is presented in Figure 1, as well as its main descriptive statistics in Table 1.<sup>10</sup>

As we can observe, the most dynamic region is Norte, which accounts for over 55.7 per cent of total manufacturing plant births between 1992 and 2007. Additionally, there is strong evidence that single plants and multi-plants locate differently: Norte is responsible for the highest birth rate of single plants, while the highest rates for new multi-plants are observed in Lisbon (capital) followed by Porto. As well, the Wilcoxon signed-rank test for the equality of matched pairs of observations allows us to conclude that both distributions are not the same.<sup>11</sup> Therefore, as the birth of a single plant or a multi-plant, results from two different spatial decision-making processes, they should be treated separately.

We also consider the distribution of new single plants and new multi-plants across sectors (Table 2).

We therefore conclude for their uneven behaviour: sectors 18, 28 and 36 (CAE – Rev. 2) are the most relevant ones for new single plants, while for the new multi-plants the most relevant sectors are 15, 18 and 28.

In order to evaluate the spatial concentration of each sector by NUTS 3, we compute the location coefficient<sup>12</sup> for both new single plants and multi-plants, as well as for total industry.<sup>13</sup> We then observe the uneven concentration of each sector when considering different samples. In fact, and in the case of new single plants, the location coefficient reveals that sectors 35 and 19 are the most concentrated in space; while for new multi-plants, the geographical concentration

<sup>7</sup> The *concelho* is a small administrative region in Portugal, with a population that ranges between 1,924 and 564,657 inhabitants (INE 2001) and an average area of 323.5 km<sup>2</sup> (INE 2003b).

<sup>8</sup> We considered the change of the Code of Administrative Division in 1998 that introduced three new municipalities (Vizela, Trofa and Odivelas) and included them into the original ones, as our study is largely previous to 1998. Additionally, we excluded the islands of Açores and Madeira, as the number of new plants of the manufacturing sector born during 1992–2007 was quite small.

<sup>9</sup> We must note that plants are removed from the sample if they do not respond to three consecutive spells of the survey, meaning that shutdowns are accommodated although they cannot be distinguished from nonresponses.

<sup>10</sup> The Nomenclature of Territorial Units for Statistics (NUTS) (1989) (INE 1989) defines five regions at NUTS 2 and 28 homogeneous regions at NUTS 3 for mainland Portugal.

<sup>11</sup> We may also refer that every municipality register the birth of single plants during the period under study, while 27 (9.82%) municipalities has zero births of multi-plants.

<sup>12</sup> The location coefficient (LC) is a measure of relative concentration of a sector. It compares the spatial distribution of a sector with the spatial distribution of total sectors:

$$LC = 0.5 \sum_{i=1}^I |x_{ik} / x_k - x_i / x|,$$

where  $i$  = region = 1, . . . ,  $I$ ;  $k$  = sector = 1, . . . ,  $K$ ;  $x_k = \sum_{i=1}^I x_{ik}$ ;  $x_i = \sum_{k=1}^K x_{ik}$ ;  $x = \sum_{k=1}^K \sum_{i=1}^I x_{ik}$ .

<sup>13</sup> Literature shows several measures of industrial concentration. According to Duranton and Overman (2005), the location coefficient, as the Herfindahl or the Gini indexes, while being extremely popular, do not take space into account and do not control for the level of industry concentration. In spite of their shortcomings, any of these measures would be suitable to show the uneven concentration of new single plants and new multi-plants. Pioneered by Ellison and Glaeser (1997), a second type of measures emerged that take under consideration space and that control for the overall agglomeration of manufacturing and for industrial concentration, while treating space as discrete. The ‘third generation’ of industrial concentration measures, introduced by Duranton and Overman (2005) with the ‘K-density’ measure, treats space as continuous. In addition, and according to Albert et al. (2011), the use of spatial statistics measures, such as Ripley’s K function, offers important advantages over traditional concentration indices, as it allows evaluating whether concentration exists, its intensity and at what distance its highest level is obtained.

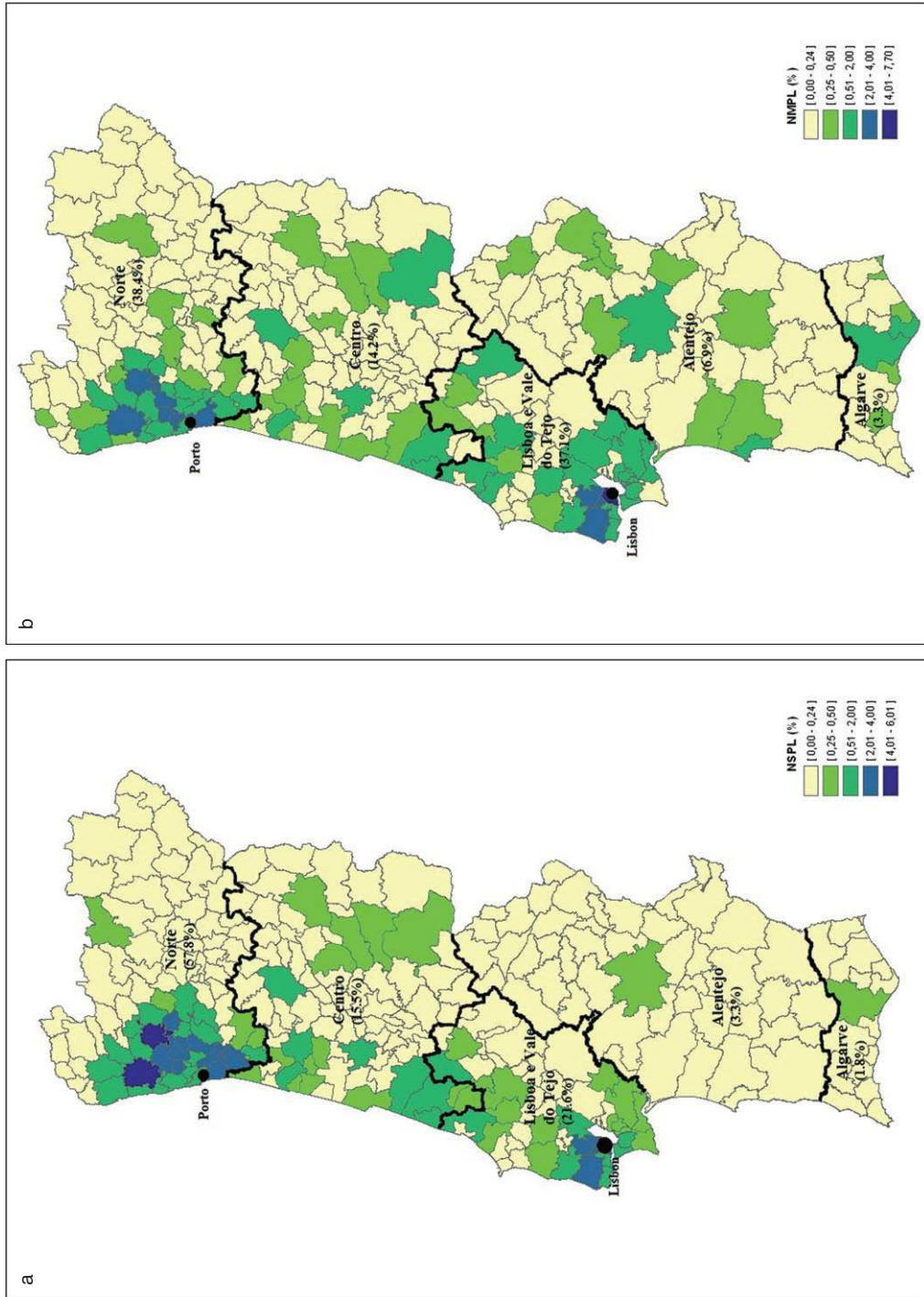


Fig. 1. New single plants and new multi-plants (1992–2007), by municipality and NUTS 2

**Table 1.** New single-plants and new multi-plants (1992–2007): Descriptive statistics

	New single-plants	New multi-plants
Number	54,512 (89.11%)	6,665 (10.89%)
Mean	11.099	1.134
Standard deviation	24.236	4.680
Min.	3	0
Máx.	141	18
Zeros	0	27

is very high for sectors such as 30, 23, 19 and 32.<sup>14</sup> If we compare with the concentration of total industry, we partially justify above evidence: sector 23 is highly concentrated for both new multi-plants and total industry; sector 19 is concentrated for both new single plants and multi-plants, as well as for total industry. However, the relative concentration of new single plants and new multi-plants cannot be explained by the spatial distribution of total industry: in fact, and for instance, the high concentration of sector 35 for new single plants or sectors 30 and 32 for new multi-plants is not related with the distribution of total industry. Therefore, we might conclude that the spatial pattern of new single plants and new multi-plants is dissimilar, and does not follow the spatial distribution of total industry.

### 3.2 Explanatory variables

Table 3 summarizes the main information about the explanatory variables that we use to study the location decision of single plants and multi-plants. As in previous research on location choice (e.g., Coughlin and Segev 2000; Guimarães et al. 2004; Holl 2004a), the variables influencing the choice of a particular location can be broadly classified into three categories: (i) cost variables, labour market features and technological characteristics; (ii) market size and its accessibility; and (iii) agglomeration economies.

#### 3.2.1 Supply variables

According to the least cost theory, inspired by the contributions of Launhardt (1993), Von Thünen (1966) and Weber (1929), land, labour and capital costs have a negative impact on firms' expected profit. Therefore, the optimal location of a firm corresponds to the least cost site.

Previous research on location choice used land area (e.g., Coughlin et al. 1991; and Woodward 1992), property taxes (e.g., Carlton 1983), or the population density (e.g., Bartik 1985; Schmenner et al. 1987; Guimarães et al. 2000; Woodward et al. 2006) as a proxy for land costs, but frequently failed to establish its relevance. Following Bartik's (1985) approach, which assumes that industrial and residential users compete for the same space, the population density in each municipality is adopted in this research as a proxy for land costs, with a negative influence on the location choice expected to be observed.

In their location decisions, firms are also motivated by labour market conditions, particularly, labour costs and the qualifications of the workforce. The labour costs, measured by the

<sup>14</sup> Literature on industrial concentration also confirms these results: according to Duranton and Overman (2005), sectors 17, 18, 19, 22, 30, 31 32 and 33 are the most concentrated in UK, while in Spain, and according to Albert et al. (2011), the most concentrated sectors are 17, 19, 22, 24, 30, 31 32 and 33.

**Table 2.** New single plants and new multi-plants (1992–2007), by sectorial activity

CAE – Rev. 2	Manufacturing industry	New single plants*		New multi-plants*		Total industry**	
		%	LC	%	LC	CAE	LC
15	Manufacture of food products and beverages	10.07	0.32	17.31	0.20	15+16	0.22
16	Manufacture of tobacco	0.00	0.00	0.00	0.00		
17	Manufacture of textile	6.68	0.37	5.68	0.37	17	0.40
18	Manufacture of wearing apparel; dressing and dyeing of fur	20.60	0.37	12.87	0.35	18+19	0.58
19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	5.96	0.49	2.80	0.56		
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of straw and plaiting materials	9.66	0.28	5.22	0.28	20	0.24
21	Manufacture of paper and paper products	0.68	0.23	0.98	0.33	21+22	0.14
22	Publishing, printing and reproduction of recorded media	5.39	0.39	9.36	0.32		
23	Manufacture of coke, refined petroleum products and nuclear fuel	0.00	0.00	0.10	0.64	23	0.74
24	Manufacture of chemicals and chemical products	0.91	0.28	2.57	0.28	24	0.14
25	Manufacture of rubber and plastics products	1.30	0.29	2.14	0.24	25	0.29
26	Manufacture of other non-metallic mineral products	6.37	0.28	9.69	0.21	26	0.22
27	Manufacture of basic metals	0.48	0.27	0.35	0.37	27+28	0.14
28	Manufacture of fabricated metal products, except machinery and equipment	15.12	0.20	12.11	0.20		
29	Manufacture of machinery and equipment n.e.c.	3.47	0.30	5.55	0.21	29+30	0.13
30	Manufacture of office, accounting and computing machinery	0.00	0.00	0.05	0.81		
31	Manufacture of electrical machinery and apparatus n.e.c.	0.94	0.35	1.77	0.29	31+32+33	0.16
32	Manufacture of radio, television and communication equipment and apparatus	0.32	0.45	0.50	0.50		
33	Manufacture of medical, precision and optical instruments, watches and clocks	0.73	0.38	0.91	0.29		
34	Manufacture of motor vehicles, trailers and semi-trailers	0.41	0.38	0.76	0.40	34+35	0.18
35	Manufacture of other transport equipment	0.48	0.61	0.98	0.48		
36	Manufacture of furniture; manufacturing n.e.c.	10.16	0.35	7.77	0.24	36+37	0.25
37	Recycling	0.27	0.29	0.53	0.47		
15–37	Total manufacturing	100		100			

Note: LC = location coefficient; CAE = code of economic activity.

Sources: \*DEEP-MTE (1991–2007); \*\*INE (2007).



Table 3. Explanatory variables

Variable	Proxy	Expected sign	Mean	Standard deviation	Data source
Land costs	Population density, by municipality, 1992–2007	Negative	450.71	1,117.25	INE (1991–2001); INE (2002–2008); INE (2003b)
Labour costs	Real base-wage over regular working hours, by municipality and CAE, 1992–2007	Negative	297.22	184.94	DEEP–MTE (1991–2007); INE (1991–2008)
Human capital	Average years of schooling of the adult population, by municipality, 1992–2007	Positive	5.57	1.18	INE (1991); INE (2001)
Capital costs	<i>Derrama</i> plus other taxes over firms divided by total societies, by municipality, 1992–2007	Negative	538.79	779.82	DGAL (1992–2007); INE (1992–2007)
R&D	R&D expenditures per capita, by municipality, 1994–2009	Positive	0.03	0.21	OCES (1995, 1997, 1999, 2001, 2003, 2005, 2007, 2009); INE (1991–2001); INE (2002–2008)
Regional market	Per capita gross domestic product, by NUTS3, 1992–2007	Positive	5,498.46	1,763.95	INE (1991–2007)
Local market	Purchasing power index (IPPC), by municipality, 1993–2008	Positive	72.65	36.43	INE (1993–2009)
Market accessibility	Lowest geographical distance to Porto/Lisbon, by municipality	Negative	86.17	52.63	INE (2003a)
Localization economies	Share of manufacturing employment for each CAE – 2 digit, by municipality, 1991–2006	Positive	0.29	0.06	DEEP–MTE (1991–2007)
Agglomeration economies	Density of manufacturing and service plants (CAE D, G, H, I, J, K) per square kilometre, by municipality, 1992–2007	Positive	8.80	29.93	DEEP–MTE (1991–2007); INE (2003b)

average wage rate, are included in almost every study on location choice, while its coefficient is usually negative (e.g., Coughlin et al. 1991; Barbosa et al. 2004; Figueiredo et al. 2002; Holl 2004a, 2004b; Woodward et al. 2006; Basile et al. 2009).<sup>15</sup> Therefore, in this research, the proxy for labour costs is the real wage per working hour, for each municipality and code of economic activity (CAE) at 2-digit level.

The influence of population skills and abilities on the productivity of the workforce and, therefore, on firms' location choice, has also been included in several studies. Some authors considered the influence of specific skills on location choice (e.g., the number of engineers, Carlton 1983), or, more often, the education-level of population (e.g., Bartik 1985; Woodward 1992; Basile et al. 2009), but results were mostly inconclusive. In this study, the average years of schooling of the adult population is used as a proxy for human capital stock in each municipality, with a positive influence on location choice expected to be observed.

Several studies use the interest rate, corporate taxes or other local taxes to evaluate the influence of capital costs on location choice (e.g., Carlton 1983; Bartik 1985; Friedman et al. 1992; Woodward 1992; Head et al. 1999; Head and Mayer 2004; Basile et al. 2009), while its coefficient is often not significant. Other studies consider different local taxes: corporate vs. personal income taxes (Feld and Kirchgässner 2002); business vs. property taxes (Jofre-Monseny and Solé-Ollé 2010), and concluded that local taxes deter new manufacturing plants. Additionally, variables capturing the effects of bureaucracy and market regulation on location choice are sometimes considered (e.g., Friedman et al. 1992; Basile et al. 2009). In this study, capital costs are measured by the taxes over companies collected by municipalities, which include both *derrama*<sup>16</sup> and other taxes over firms.

In addition to traditional location determinants, the influence of *technological characteristics* on location choice has been considered. Therefore, the Research and Development (R&D) *per capita* expenditures at the municipality level is introduced. Taking into consideration the composition of R&D expenditures and its geographical distribution, a small but positive impact of the R&D expenditures on location choice is expected to be observed.<sup>17</sup>

### 3.2.2 Demand variables

Another strand of the literature, motivated by Fetter (1924), Hotelling (1929) and Chamberlin (1950), focuses on the influence of demand variables, such as the market size and its accessibility, on the location choice.

The market size, measured by per capita income, is frequently used in location studies (e.g., Coughlin et al. 1991; Coughlin and Segev 2000; Guimarães et al. 2004), while others use the gravitation personal income as a proxy for the potential market, as it intends to capture both market size and its accessibility (e.g., Friedman et al. 1992; Woodward 1992; Head and Mayer 2004; Holl 2004b; Basile et al. 2009). In most studies, a positive influence of market size on location choice was evidenced.

Following previous research, and in order to capture the influence of market size on location choice, two possibilities are considered in this research: first, we use the Purchasing Power

<sup>15</sup> Except for the location decisions of foreign firms, as a positive coefficient frequently emerges, revealing the positive effect of high wages, correlated with high qualifications of the workforce, on the location choice (e.g., Friedman et al. 1992; Woodward 1992; Guimarães et al. (2000).

<sup>16</sup> *Derrama* (or municipal surcharge) is a local municipal tax that can be charged annually by municipal authorities up to a maximum of 10 per cent of the amount paid in corporate tax (IRC). Other local but non-corporate taxes include the Municipal Property Transfer Tax (IMT) and the Municipal Property Tax (IMI).

<sup>17</sup> In Portugal, and for the year 1999, about 66.5 per cent of the R&D expenditure was made by the Government and Higher Education sector (OCES 1999).

Index<sup>18</sup> for each municipality, which intends to capture the influence of local market on location choice. Alternatively, and taking into consideration the small size of most Portuguese municipalities, we use the per capita gross domestic product at regional level (NUTS 3) as a proxy for the regional market.

Additionally, some authors consider the influence of market accessibility on firms' location decisions by introducing the geographical or time distance to an important market (city centre or capital), airport or port (Hansen 1987; Guimarães et al. 2000; Holl 2004a, 2004b). Therefore, in this study the shortest physical distance between each municipality to Porto or Lisbon is employed, which are the most important cities in Portugal, both equipped with international airports, ports and railway stations. It is expected to observe that high distances to Porto or Lisbon negatively influence the decision to locate new investments.

### 3.2.3 Agglomeration economies

Both location theory and empirical studies claim for the relevance of agglomeration economies. The literature usually distinguishes between urbanization economies, which are external to firms and industries but internal to a city (e.g., access to large population centres and large and diversified service and manufacturing sectors) and localization economies, which are external to firms but internal to an industry (e.g., access to specialized labour force and communication economies).

Nearly all studies conclude for the positive influence of localization economies on location choice, usually measured by the number of firms or employment in each manufacturing sector (e.g., Hansen 1987; Guimarães et al. 2004; Head and Mayer 2004), or by an industrial diversity index (e.g., Holl 2004a, 2004b; Arauzo-Carod 2005). Regarding urbanization economies, some researchers use a measure of global industrial activity in the region (e.g., Carlton 1983; Bartik 1985; Coughlin et al. 1991), while others use a proxy for the volume of the economic activity, such as per capita income, population size or density of industrial and services activities (e.g., Coughlin and Segev 2000; Guimarães et al. 2004). In this study, to account for the localization economies, the share of manufacturing employment for each CAE – 2 digit in each municipality is considered, while the urbanization economies are measured by the density of manufacturing and service plants per square kilometre in each municipality.

Finally, and in order to control for unobservable region characteristics that might affect firms' location choice, a dummy variable for each region, NUTS 3 level has been included.

## 4 Empirical results

We model the location choices of new single plants and new multi-plants between 1992 and 2007 through a discrete choice analysis. Table 4 summarizes the main estimation results of a standard conditional logit model by means of its equivalence with the Poisson model, which is guaranteed by using a set of dummy variables for each combination of year and 2-digit CAE sector, as demonstrated by Guimarães et al. (2003). (Regressions (1) to (4) refer to single plant location choices, while regressions (5) to (8) refer to new multi-plant location decisions. In both samples, we consider either the influence of the regional market (regressions (1), (3), (5) and (7)) or the local market (regressions (2), (4), (6) and (8)). Additionally,

---

<sup>18</sup> The Purchasing Power Index (IPPC) intends to capture the purchasing power in each municipality. It is an index built by means of a model of factorial analysis and recurring to a set of 20 variables that were selected according to an expenditure criterion upon a larger group of 70 variables (INE 2004).

Table 4. Poisson estimation results: New single plants and new multi-plants (1992–2007)

Variables	Single plants				Multi-plants			
	with dummy by year × CAE		with dummies by year × CAE and NUTS3		with dummy by year × CAE		with dummies by year × CAE and NUTS3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land costs	-0.620* (0.000)	-0.621* (0.000)	-1.194* (0.000)	-1.190* (0.000)	-0.666* (0.000)	-0.536* (0.000)	-1.097* (0.000)	-0.988* (0.000)
Labour costs	-1.002* (0.000)	-0.867* (0.000)	-0.899* (0.000)	-0.900* (0.000)	-0.428* (0.000)	-0.338* (0.000)	-0.560* (0.000)	-0.583* (0.000)
Human capital	-2.964* (0.000)	-2.173* (0.000)	-1.219* (0.000)	-1.257* (0.000)	0.408* (0.010)	0.332 (0.082)	0.803* (0.000)	0.158 (0.524)
Capital costs	0.065* (0.000)	0.067* (0.000)	0.072* (0.000)	0.072* (0.000)	0.077* (0.004)	0.077* (0.000)	0.077* (0.000)	0.077* (0.000)
R&D	0.066* (0.000)	0.067* (0.000)	0.057* (0.000)	0.056* (0.000)	0.069* (0.000)	0.064* (0.000)	0.080* (0.000)	0.073* (0.000)
Regional market	0.478* (0.000)	—	0.135 (0.431)	—	0.446* (0.000)	—	0.096 (0.833)	—
Local market	—	-0.203* (0.000)	—	-0.027 (0.342)	—	0.295* (0.000)	—	0.534* (0.000)
Market accessibility	-0.109* (0.000)	-0.133* (0.000)	-0.198* (0.000)	-0.199* (0.000)	-0.128* (0.001)	-0.161* (0.000)	-0.049 (0.354)	-0.048 (0.362)
Localization economies	0.754* (0.000)	0.745* (0.000)	0.745* (0.000)	0.750* (0.000)	0.509* (0.000)	0.509* (0.000)	0.540* (0.000)	0.544* (0.000)
Urbanization economies	1.137* (0.000)	1.143* (0.000)	1.330* (0.000)	1.325* (0.000)	0.916* (0.000)	0.779* (0.000)	1.181* (0.000)	1.048* (0.000)
Constant	12.191* (0.000)	15.143* (0.000)	14.433* (0.000)	15.480* (0.000)	1.714 (0.028)	3.373* (0.000)	6.671 (0.080)	5.908* (0.000)
Log likelihood	-39,920.9	-40,053.6	-35,823.8	-35,823.6	-11,021.8	-11,033.6	-10,534.4	-10,520.0
Pseudo R <sup>2</sup>	0.5311	0.5296	0.5793	0.5796	0.2288	0.2279	0.2629	0.2639
LR test	90,449.94 (0.000)	90,184.42 (0.000)	98,644.09 (0.000)	98,644.38 (0.000)	6,538.72 (0.000)	6,515.03 (0.000)	7,513.44 (0.000)	7,542.13 (0.000)
LR test regional dummies = 0			6,754.23 (0.000)	6,781.14 (0.000)			789.52 (0.000)	791.22 (0.000)
Pearson statistic	83,097.48 (0.000)	83,309.44 (0.000)	71,905.00 (0.000)	71,912.66 (0.000)	33,426.1 (0.000)	33,590.01 (0.000)	32,960.91 (0.000)	32,915.36 (0.000)

Notes: *p*-values are in parentheses.  
\* significant at 0.05 level of significance.

we introduce a dummy variable for each NUTS 3 region, in order to capture other non-observable regional characteristics (regressions (3), (4) and (7), (8)). All explanatory variables are included in their logarithmic form.

At a first glance, we conclude that some location determinants have a similar performance across both samples: land and labour costs, agglomeration economies and R&D are always significant and with the expected sign, while capital costs never have the expected sign. Nevertheless, there are some specificities: the accessibility to main markets has a better performance in the single plant sample while the local market is only relevant in the multi-plant case. In both samples, we observe that the likelihood-ratio test sustains the hypothesis of inclusion of dummies by NUTS 3.

However, the Pearson statistics for the goodness of fit lead us to reject the hypothesis that the variance equals the mean at 1 per cent of significance, evidencing overdispersion. This reveals that the Poisson regression is not adequate to our data, suggesting that we should try the negative binomial model.

Tables 5 and 6 resume the main estimation results of a negative binomial model (Negbin II) for single plants and multi-plants respectively. As before, we run a standard negative binomial model with a dummy for each combination of year and 2-digit CAE sector in order to ensure equivalence with the conditional logit model and introduce either the regional market (regressions (9) and (21)) or the local market influence (regressions (10) and (22)). Similarly, we consider dummies by NUTS 3 regions (regressions (11), (12) and (23), (24)), so as to capture other regional characteristics that might affect firms' location choice.

As we can observe from both likelihood-ratio and Wald tests, all regressions perform quite well. Additionally, the likelihood-ratio test of  $\alpha = 0$  indicates that the Poisson dispersion is greater than 1 (overdispersion), which confirms that the negative binomial model is more suitable to our data.<sup>19</sup> Also, the likelihood ratio tests between new single plants and new total plants (Table 5), as well as between new multi-plants and new total plants (Table 6), confirms that the coefficient estimates between the two samples versus alternative total plants are not similar.

Let us consider the negative binomial model without 'specific-effects' by municipalities (regressions (9) to (12) and (21) to (24)). A simple likelihood-ratio test for the inclusion of dummy variables by region or the increase of the 'pseudo-R<sup>2</sup>' sustains the hypothesis that there are other regional characteristics that affect plants' location.

As we can observe from regressions (11), (12) and (23), (24), some variables have a regular behaviour across both samples. In fact, we conclude that the most relevant location determinants for both single plants and multi-plants are the urbanization and localization economies and both land and labour costs. Compared with previous research, most studies corroborate the importance of agglomeration economies and labour costs for location choice (e.g., Coughlin et al. 1991; Figueiredo et al. 2002; Barbosa et al. 2004). At the same time, few authors confirm the influence of land costs (e.g., Guimarães et al. 2004), which may be attributed to the use of the proxy population density, which is sensitive to the size of the spatial unit.<sup>20</sup> Our results also reveal that capital costs are almost never significant or with the correct sign, which is confirmed in several studies (e.g., Carlton 1983; Woodward 1992; Head and Mayer 2004), and can be justified by the absence of noteworthy differences in the cost of capital across Portuguese municipalities. As well, our estimates show that the human capital variable has an irregular behaviour, which is also evidenced in the literature (e.g., Guimarães et al. 2000; Arauzo-Carod

<sup>19</sup> Additionally, the Vuong statistics of the ZI Negative Binomial versus the standard negative binomial for the new single plants and new multi-plants was  $-0.42$  and  $-0.03$ , respectively, supporting the hypothesis that the standard negative binomial is more suitable to our data.

<sup>20</sup> In this study, and similarly to Figueiredo et al. (2002), this limitation is partially solved by the small size of each spatial unit.

Table 5. Negative binomial estimation results: New single plants (1992–2007)

Variables	Negative binomial model – new single plants														
	without specific effects					with random effects by municipality					with fixed effects by municipality				
	with dummy by year × CAE	with dummies by year × CAE and NUTS3	with dummy by year × CAE	with dummies by year × CAE and NUTS3	with dummy by year × CAE	with dummies by year × CAE and NUTS3	with dummy by year × CAE	with dummies by year × CAE and NUTS3	with dummy by year × CAE	with dummies by year × CAE and NUTS3	with dummy by year × CAE	with dummies by year × CAE and NUTS3	with dummy by year × CAE	with dummies by year × CAE and NUTS3	
(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)				
Land costs	-0.641* (0.000)	-0.560* (0.000)	-1.01* (0.000)	-0.986* (0.000)	-0.513* (0.000)	-0.542* (0.000)	-0.481* (0.001)	-0.486* (0.000)	-0.404* (0.000)	-0.397* (0.000)	-0.312* (0.005)	-0.324* (0.008)			
Labour costs	-0.745* (0.000)	-0.702* (0.000)	-0.706* (0.000)	-0.728* (0.000)	-0.922* (0.000)	-0.945* (0.000)	-0.934* (0.000)	-0.930* (0.000)	-0.834* (0.000)	-0.821* (0.000)	-0.852* (0.000)	-0.936* (0.000)			
Human capital	-1.376* (0.000)	-1.446* (0.000)	-0.195 (0.129)	-0.418* (0.014)	2.652* (0.000)	2.983* (0.000)	0.579* (0.042)	0.667* (0.021)	2.768* (0.000)	3.165* (0.000)	0.275 (0.418)	0.250 (0.505)			
Capital costs	0.009 (0.374)	0.010 (0.370)	0.025* (0.027)	0.026 (0.092)	-0.031 (0.057)	0.124 (0.124)	-0.037* (0.025)	-0.032 (0.052)	-0.039* (0.020)	-0.029 (0.076)	-0.042* (0.035)	-0.039 (0.074)			
R&D	0.037* (0.000)	0.035* (0.000)	0.052* (0.000)	0.048* (0.000)	0.013 (0.084)	0.012 (0.113)	0.026* (0.001)	0.027* (0.006)	0.015 (0.059)	0.014 (0.070)	0.021* (0.023)	0.023* (0.022)			
Regional market	0.282* (0.000)	-	0.211 (0.432)	-	0.630* (0.000)	-	0.468* (0.028)	-	0.739* (0.000)	-	0.388 (0.087)	-			
Local market	-	0.208 (0.052)	-	0.211* (0.008)	-	0.169* (0.023)	-	-0.089 (0.192)	-	0.185* (0.011)	-	-0.076 (0.392)			
Market accessibility	-0.148* (0.000)	-0.169* (0.000)	-0.198* (0.000)	-0.189* (0.000)	-0.177* (0.007)	-0.217* (0.001)	-0.421* (0.001)	-0.423* (0.005)	-	-	-	-			
Localization economies	0.552* (0.000)	0.552* (0.000)	0.552* (0.000)	0.551* (0.000)	0.594* (0.000)	0.594* (0.000)	0.571* (0.000)	0.570* (0.000)	0.597* (0.000)	0.597* (0.000)	0.576* (0.000)	0.557* (0.000)			
Urbanization economies	1.082* (0.000)	1.030* (0.000)	1.124* (0.000)	1.088* (0.000)	0.431* (0.000)	0.449* (0.000)	0.664* (0.000)	0.674* (0.000)	0.333* (0.000)	0.342* (0.000)	0.614* (0.000)	0.669* (0.000)			
Constant	9.833* (0.000)	11.044* (0.000)	10.201* (0.000)	10.967* (0.000)	2.155 (0.057)	6.445* (0.000)	7.715* (0.000)	11.76* (0.000)	1.710 (0.559)	4.114* (0.000)	8.221* (0.000)	12.99* (0.000)			
Log likelihood	-31.236	-31.250	-30.399	-30.387	-29.304.2	-29.318.7	-29.138.6	-29.140.1	-29.397.5	-29.416.4	-28.239.1	-28.242.1			
Pseudo R <sup>2</sup>	0.2012	0.2009	0.2227	0.2229											
LR test	15.740.4	15.711.9	17.414.4	17.421.2											
LR test of $\alpha = 0$	17.000 (0.000)	18.000 (0.000)	11.000 (0.000)	12.000 (0.000)	14.049.9 (0.000)	14.010.1 (0.000)	14.351.8 (0.000)	14.339.7 (0.000)	14.828.9 (0.000)	14.779.9 (0.000)	14.187.6 (0.000)	8.558.37 (0.000)			
Wald test			1.335.15 (0.000)	1.372.41 (0.000)			233.42 (0.000)	252.23 (0.000)			337.68 (0.000)	363.11 (0.000)			
LR test regional dummies = 0			931.0 (0.000)	938.9 (0.000)	752.7 (0.000)	759.2 (0.000)	791.1 (0.000)	804.8 (0.000)	2.512.7 (0.000)	2.506.8 (0.000)	3.370.8 (0.000)	3.369.3 (0.000)			

Notes: *p*-values are in parentheses.

\* significant at 0.05 level of significance.

Table 6. Negative binomial estimation results: New multi-plants (1992–2007)

Variables	Negative binomial model – new multi-plants											
	without specific effects						with fixed effects by municipality					
	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
Land costs	-0.781* (0.000)	-0.615* (0.000)	-1.109* (0.000)	-0.983* (0.000)	-1.108* (0.000)	-1.023* (0.000)	-1.137* (0.000)	-1.079* (0.000)	-0.879* (0.001)	-0.881* (0.042)	-0.680* (0.021)	-0.711* (0.018)
Labour costs	-0.347* (0.001)	-0.306* (0.003)	-0.460* (0.000)	-0.485* (0.000)	-0.639* (0.000)	-0.634* (0.000)	-0.647* (0.000)	-0.649* (0.000)	-0.484* (0.000)	-0.477* (0.000)	-0.468* (0.000)	-0.435* (0.000)
Human capital	0.728* (0.000)	0.273 (0.242)	1.052* (0.000)	0.338 (0.234)	1.169* (0.016)	0.710 (0.171)	0.832 (0.142)	0.416 (0.480)	-0.223 (0.744)	-0.218 (0.693)	-0.442 (0.322)	-0.483 (0.254)
Capital costs	0.079* (0.008)	0.081* (0.000)	0.076* (0.014)	0.077* (0.008)	0.071 (0.052)	0.075* (0.031)	0.082* (0.019)	0.086* (0.013)	0.057 (0.146)	0.059 (0.118)	0.059 (0.185)	0.061 (0.156)
R&D	0.056* (0.000)	0.047* (0.001)	0.074* (0.000)	0.067* (0.000)	0.013 (0.449)	-0.017 (0.349)	0.005 (0.765)	0.007 (0.682)	-0.035 (0.168)	-0.041 (0.149)	-0.044 (0.122)	-0.045 (0.149)
Regional market	0.402* (0.000)	- (0.000)	0.160 (0.753)	- (0.000)	0.398 (0.111)	- (0.575)	0.257 (0.596)	- (0.494)	0.039 (0.924)	- (0.552)	0.214 (0.512)	- (0.412)
Local market	- (0.000)	0.536* (0.000)	- (0.000)	0.576* (0.000)	- (0.005)	0.575* (0.005)	- (0.005)	0.494* (0.017)	- (0.089)	0.552 (0.089)	- (0.118)	- (0.118)
Market accessibility	-0.140* (0.002)	-0.166* (0.000)	-0.111 (0.068)	-0.104 (0.086)	-0.061 (0.642)	-0.088 (0.502)	-0.157 (0.476)	-0.172 (0.431)	- (0.537)	- (0.528)	- (0.522)	- (0.518)
Localization economies	0.467* (0.000)	0.472* (0.000)	0.506* (0.000)	0.510* (0.000)	0.531* (0.000)	0.530* (0.000)	0.529* (0.000)	0.529* (0.000)	0.537* (0.000)	0.528* (0.000)	0.522* (0.000)	0.518* (0.000)
Urbanization economies	1.012* (0.000)	0.841* (0.000)	1.166* (0.000)	1.019* (0.000)	1.406* (0.000)	1.305* (0.000)	1.461* (0.000)	1.354* (0.000)	1.328* (0.000)	1.301* (0.000)	1.514* (0.000)	1.485* (0.000)
Constant	1.457 (0.112)	2.411 (0.104)	5.450 (0.199)	5.051* (0.000)	6.184* (0.015)	7.472* (0.000)	8.146 (0.060)	8.684* (0.000)	8.602* (0.018)	5.114 (0.183)	8.452 (0.125)	10.212 (0.059)
Log likelihood	-10.715.0	-10.713.5	-10.383.5	-10.370.6	-10.050.6	-10.047.6	-10.020.9	-10.018.0	-10.119.7	-10.116.6	-10.088.6	-10.087.2
Pseudo R <sup>2</sup>	0.1681	0.1682	0.1937	0.1947								
LR test	4.325.8	4.328.9	4.988.9	5.014.74								
LR test of $\alpha = 0$	613.31 (0.000)	640.10 (0.000)	301.77 (0.000)	298.84 (0.000)								
Wald test					3.232.4 (0.000)	3.228.23 (0.000)	3.320.4 (0.000)	3.328.3 (0.000)	3.688.7 (0.000)	3.671.2 (0.000)	3.528.6 (0.000)	3.519.4 (0.000)
LR test regional dummies = 0			718.22 (0.000)	740.46 (0.000)			88.17 (0.000)	88.32 (0.000)			75.21 (0.004)	74.18 (0.008)
LR test multi vs total plants	41.970.3 (0.000)	42.017.1 (0.000)	40.962.4 (0.000)	40.971.9 (0.000)	39.259.9 (0.000)	39.301.4 (0.000)	39.026.5 (0.000)	39.049.0 (0.000)	41.068.3 (0.000)	41.106.4 (0.000)	39.671.8 (0.000)	39.679.1 (0.000)

Notes: p-values are in parentheses.  
\* significant at 0.05 level of significance.

2005; Basile et al. 2009), and may be justified by the aggregate nature of the indicator, which does not allow us to evaluate the importance of some specific skills (e.g., engineers) for the location choice.

Nevertheless, there are some noteworthy differences across both samples, namely on the demand side. In fact, we note that new single plants are particularly sensitive to the accessibility to main markets and infrastructures with elasticities around 0.20, while the market size is not very relevant. At the same time, new multi-plants are strongly influenced by the local market, with elasticities about 0.58, while the market accessibility loses relevance. These results accords with intuition, as multi-plants have access to more information about sites and are less risk averse than single plants and therefore they privilege access to multiple local markets, while new single plants are more sensitive to the risk and therefore, they privilege access to a wider market area. Also, the relevance of demand variables for location choice is confirmed in many studies (e.g., Coughlin et al. 1991; Friedman et al. 1992; Head et al. 1999; Guimarães et al. 2000; Head and Mayer 2004; Holl 2004b; Basile et al. 2009). Finally, the R&D variable has a significant and positive influence on location choice in both samples and supports the evidence of R&D spillovers.

Our results also find confirmation in previous studies on location choice for the Portuguese case, such as Figueiredo et al. (2002), Barbosa et al. (2004), and Holl (2004a, 2004b), who find relevance for labour costs, market size and its accessibility, while capital costs were not significant.

Following Guimarães et al. (2004),<sup>21</sup> and in order to more effectively control for the potential violation of the IIA assumption, we then consider ‘specific effects’ by municipality, either by considering random (regressions (13) to (16) and (25) to (28)) or fixed (regressions (17) to (20) and (29) to (32)) effects by municipality.

To begin with the single plant location choices, the differences between the log-likelihoods of the models with specific effects and the corresponding without specific effects models are statistically significant, providing evidence that the inclusion of specific effects by municipality is convincing. Likewise, the inclusion of dummy variables by region is also supported by a likelihood-ratio test, which gives reason for the existence of supra-municipality characteristics that are not captured by other variables. Finally, and in order to test for the inclusion of random or fixed effects by municipality, we perform a Hausman test that tests the null hypothesis that the coefficients estimated by the efficient random effects estimator are the same as the ones estimated by the consistent fixed effects estimator. We compare regressions (15) and (19) and the resulting statistics equals 0.04, which supports the hypothesis of random effects by municipality.

As before, and focusing on regressions (15) and (16), we observe that single plants’ location choices are strongly influenced by labour costs and agglomeration economies. In fact, we estimate that, everything else constant, a 1 per cent increase in labour costs leads to about a 0.93 per cent decrease in the number of new single plants, while the corresponding elasticity for urbanization and localization economies is about 0.57 and 0.67, respectively. Additionally, land costs and market accessibility are also relevant for the location of new investments, with elasticities of about 0.49 and 0.42, respectively. Furthermore, the R&D variable, while significant, has the lowest elasticity.

We then focus on the location choices of new multi-plants. As before, our estimation results support the inclusion of specific effects by municipality and dummies by region. Also, the statistics of the Hausman test for the inclusion of random versus fixed effects (regressions (27), (28) and (31), (32)) equal 0.01, which supports the hypothesis that the specific effects are not correlated with the explanatory variables.

<sup>21</sup> Guimarães et al. (2004) claimed that to more effectively control for the violation of the IIA assumption, one should include an additional specific effect to each alternative, which would collapse to the Poisson model with random effects if the additional term is a random variable or into the Poisson model with fixed effects if the additional term is a fixed effect.



From our results, and focusing on regressions (27) and (28), we observe that although the main location determinants are still the agglomeration economies and both land and labour costs, there are some important differences. The first notable difference is the importance of urbanization economies and land costs for the launching of new multi-plants, with elasticities of about 1.4 and 1.1, respectively. Another distinction is the relevance of the local market size, rather than market accessibility or regional market, which may be deduced from both overall and individual significance tests. In fact, we estimate that everything else constant, a 1 per cent increase in local market size leads to about a 0.5 per cent increase in the number of new multi-plants, while market accessibility and regional market are often non-significant.

We then conclude that new multi-plants are particularly sensitive to urbanization economies, land costs and local market size, while new single plants are more responsive to labour costs, both localization and urbanization economies and market accessibility. These results may be explained by attitudes to risk, phase in the life cycle of firms and access to information. New single plants are at the beginning of their life cycles and are more risk averse and therefore, they privilege the access to larger market areas and to regions with lower production costs. New multi-plants benefit from the experience of the mother-firm, namely in the labour market, and therefore, the importance of labour costs decreases, while being less risk averse which make them exploit diversified local markets. At the same time, new multi-plants have access to more information about sites and therefore, they are more able to identify places with lower production costs. Additionally, the joint optimization of the plant location of multi-plant firms may justify a different behaviour from single plant firms (Smith 1981). Finally, we note that new multi-plants are particularly attracted by areas of high density where they can realize increased opportunities and also by the proximity to the political decision centre (Lisbon) (Figure 1).

## 5 Concluding remarks

Using micro-level data, this paper explored the importance of location determinants for firms' location choice by assuming that start-ups behave differently according to the number of their plants.

Our results confirmed the relevance of agglomeration economies for plants' location choice, which is in line with previous studies. In fact, for both single plants and the multi-plants, the concentration of economic activity clearly attracts new investments. Additionally, urbanization economies seem to have a particular relevance for the location decision of new multi-plants, which might reveal that they privilege access to areas of high density.

This research also evidenced that start-ups are strongly influenced by land and labour costs. In fact, we observed that new multi-plants are particularly sensitive to land costs, while new single plants are more responsive to labour costs, which may be explained by the life cycle of firms. On the contrary, the hypothesis concerning the negative influence of capital costs on location choice is not confirmed in our study, which can be justified by the absence of significant differences in the cost of capital across Portuguese municipalities. Also, we observed that the human capital variable has an uneven behaviour that can be explained by its aggregate nature. Furthermore, our results confirmed the presence of spillovers, with the elasticity of plant births with respect to R&D expenditures being positive but quite small.

These results are particularly appealing and are in line with the evidence that new multi-plants benefit from multi-plant economies, matured entrepreneurship and easier access to information, which make them less risk averse. Therefore, new multi-plants privilege access to high density areas and are more able to exploit new local markets. On the contrary, single plants are at the beginning of their life cycles and so, they are more risk averse, which implies that the accessibility to main markets and infrastructures and lower production cost sites are the key

location determinants. From the standpoint of local politicians, these results are quite interesting and might suggest different policy options to attract single plants (e.g., construction of accessibilities) and multi-plants (e.g., soil policies).

The study of industrial location choice may proceed with some improvements in our research. An appealing research topic relates to the influence of local policies, particularly, fiscal options, on start-ups' location choices. Another area to focus on is the human capital variable in order to evaluate the influence of different qualifications of workforce on firms' location choice.

## References

- Albert JM, Casanova MR, Orts V (2011) Spatial location patterns of Spanish manufacturing firms. *Papers in Regional Science*, doi: 10.1111/j.1435-5957.2011.00375.x
- Arauzo-Carod JM (2005) Determinants of industrial location: An application for catalan municipalities. *Papers in Regional Science* 84: 105–119
- Arauzo-Carod JM, Liviano-Solis D, Manjón Antolín M (2010) Empirical studies in industrial location: An assessment of their methods and results. *Journal of Regional Science* 50: 685–711
- Arauzo-Carod JM, Manjón Antolín M (2004) Firm size and geographical aggregation: An empirical appraisal in industrial location. *Small Business Economics* 22: 299–312
- Autant-Bernard C, Mangematin V, Massard N (2006) Creation of biotech SMEs in France: The role of local environment. *Small Business Economic Journal* 26: 173–187
- Basile R (2004) Acquisition versus Greenfield investment: The location of foreign manufacturers in Italy. *Regional Science and Urban Economics* 34: 3–25
- Basile R, Castellani D, Zanfei A (2009) National boundaries and the location of multinational firms in Europe. *Papers in Regional Science* 88: 733–748
- Barbosa N, Guimarães P, Woodward D (2004) Foreign firm entry in an open economy: The case of Portugal. *Applied Economics* 36: 465–472
- Barrios S, Görg H, Strobl E (2006) Multinationals' location choice, agglomeration economies, and public incentives. *International Regional Science Review* 29: 81–107
- Bartik T (1985) Business location decisions in the United States: Estimates of the effects of unionization, taxes, and other characteristics of states. *Journal of Business and Economic Statistics* 3: 14–22
- 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* 65: 440–449
- Chamberlin EH (1950) *The theory of monopolistic competition*. Harvard University Press, Cambridge
- Coughlin C, Segev E (2000) Location determinants of new foreign-owned manufacturing plants. *Journal of Regional Science* 40: 323–351
- Coughlin C, Terza J, Arrondee V (1991) State characteristics and the location of foreign direct investment within the United States. *The Review of Economics and Statistics* 73: 675–683
- DEEP-MTE (1991–2007) *Quadros do pessoal*. DEEP-MTE, Lisboa
- DGAL (1992–2007) *Finanças municipais*. DGAL, Lisboa
- Duranton G, Overman H (2005) Testing for localization using micro-geographic data. *Review of Economic Studies* 72: 1077–1106
- Ellison G, Glaeser E (1997) Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy* 105: 889–927
- Feld LP, Kirchgässner G (2002) The impact of corporate and personal income taxes on the location of firms and on employment: Some panel evidence for the Swiss cantons. *Journal of Public Economics* 87: 129–155
- Fetter F (1924) The economic law of market areas. *Quarterly Journal of Economics* 38: 520–529
- Figueiredo O, Guimarães P, Woodward D (2002) Home-field advantage: Location decisions of Portuguese entrepreneurs. *Journal of Urban Economics* 52: 341–361
- Friedman J, Gerlowski D, Silberman J (1992) What attracts foreign multinational corporations? Evidence from branch plant location in the United States. *Journal of Regional Science* 32: 403–418
- Greene WH (2000) *Econometric analysis*. Prentice Hall International, Upper Saddle River, NJ
- Guimarães P, Figueiredo O, Woodward D (2000) Agglomeration and the location of foreign direct investment in Portugal. *Journal of Urban Economics* 47: 115–135
- Guimarães P, Figueiredo O, Woodward D (2003) A tractable approach to the firm location decision problem. *The Review of Economics and Statistics* 85: 201–204
- Guimarães P, Figueiredo O, Woodward D (2004) Industrial location modeling: Extending the random utility framework. *Journal of Regional Science* 44: 1–20

- Hansen E (1987) Industrial location choice in São Paulo, Brazil. *Regional Science and Urban Economics* 17: 89–108
- Head K, Mayer T (2004) Market potential and the location of Japanese investment in the European Union. *The Review of Economics and Statistics* 86: 959–972
- Head K, Ries J, Swenson D (1999) Attracting foreign manufacturing: Investment promotion and agglomeration. *Regional Science and Urban Economics* 29: 197–218
- Holl A (2004a) Start-ups and relocations: Manufacturing plant location in Portugal. *Papers in Regional Science* 83: 649–668
- Holl A (2004b) Transport infrastructure, agglomeration economies, and firm birth: Empirical evidence from Portugal. *Journal of Regional Science* 44: 693–712
- Hotelling H (1929) Stability in competition. *Economic Journal* 39: 41–57
- INE (1989) *Nomenclatura das unidades territoriais para fins estatísticos*. INE, Lisboa
- INE (1991) *Recenseamento geral da população e habitação (Resultados definitivos)*. INE, Lisboa
- INE (1991–2001) *Estimativas definitivas da população residente intercensitárias*
- INE (1991–2007) *Contas regionais*. INE, Lisboa
- INE (1991–2008) *Índice de preços no consumidor*. INE, Lisboa
- INE (1992–2007) *Ficheiro central de empresas e estabelecimentos (FUE)*. INE, Lisboa
- INE (1993–2009) *Estudo sobre o poder de compra concelhio*. INE, Lisboa
- INE (1994) *Classificação Portuguesa das actividades económicas – Revisão 2*. INE, Lisboa
- INE (2001) *Recenseamento geral da população e habitação (Resultados definitivos)*. INE, Lisboa
- INE (2002–2008) *Estimativas provisórias da população residente*. INE, Lisboa
- INE (2003a) *Base geográfica de referência da informação (BGRI)*. INE, Lisboa
- INE (2003b) *Referenciação territorial (REFTER)*. INE, Lisboa
- INE (2004) *Estudo sobre o poder de compra concelhio: Nota metodológica*. INE, Lisboa
- INE (2007) *Sistema de contas integradas das empresas*. INE, Lisboa
- Jofre-Monseny J, Solé-Ollé A (2010) Tax differentials in intraregional firm location: Evidence from new manufacturing establishments in Spanish municipalities. *Regional Studies* 44: 663–677
- Launhardt W (1993) *Mathematical principles of economics*, trans. by Schmidt H from *Mathematische begründung der volkswirtschaftslehre* (1885). Edward Elgar, Cheltenham
- Lee Y (2006) Relocation patterns in US manufacturing. Working Paper 06-24, Federal Reserve Bank of Cleveland
- Levinson A (1996) Environmental regulations and manufacturers' location choices: Evidence from the Census of Manufactures. *Journal of Public Economics* 62: 5–29
- McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) *Frontiers in econometrics*. Academic Press, New York
- OCES (1995, 1997, 1999, 2001, 2003, 2005, 2007, 2009) *Inquérito ao potencial científico e tecnológico*. Ministério da Ciência, Tecnologia e Ensino Superior, Lisboa
- Pellenberg PH, van Wissen LJG, van Dijk J (2002) Firm migration. In: McCann P (ed) *Industrial location economics*. Edward Elgar Publishing, Cheltenham
- Schmenner R, Huber J, Cook R (1987) Geographic differences and the location of new manufacturing facilities. *Journal of Urban Economics* 21: 83–104
- Schmidheiny K, Brühlhart M (2009) On the equivalence of location choice models: Conditional logit, nested logit and Poisson. CESifo Working Paper Series 2726, CESifo Group Munich
- Smith DM (1981) *Industrial location: An economic geographical analysis*. Wiley, New York
- Von Thünen JH (1966) *Von Thünen's Isolated State*, trans. by Hall P from *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie* (1875). Pergamon Press, Oxford
- Weber A (1929) *Alfred Weber's Theory of the location of industries*, trans. by Friedrich CJ from *Über den standort der industrien* (1909). University of Chicago Press, Chicago
- Woodward D (1992) Locational determinants of Japanese manufacturing start-ups in the United States. *Southern Economic Journal* 58: 690–708
- Woodward D, Figueiredo O, Guimarães P (2006) Beyond the Silicon Valley: University R&D and high-technology location. *Journal of Urban Economics* 60: 15–32