The geographical distribution of Portuguese business incubators: An empirical approach

Gonçalo Rodrigues Brás¹,* • Miguel Torres Preto²

¹Faculty of Economics, University of Coimbra, Coimbra, Portugal
²Instituto Superior Técnico, IN+, Center for Innovation, Technology and Policy Research, Universidade de Lisboa, Lisbon, Portugal

Abstract
This study attempts to develop an ordered logit regression model to identify the determinants of the geographical distribution of business incubators (BIs) in Portuguese regions in 2015. The results show that Portuguese BIs are more likely to be geographically concentrated in regions where the number of patent applications is higher, usually in the larger regions (particularly in metropolitan areas), and mostly in regions with a university. The study clearly demonstrates the importance of Portuguese universities to the location of BIs.

Keywords: business incubators; universities; entrepreneurial ecosystems; Portuguese regions; ordered logit model
JEL Classification Codes: L26, L24, M13, O31

1. Introduction
BIs fall into what Mason and Brown (2014) term entrepreneurial ecosystems, which formally and informally coalesce to connect, mediate and govern the performance within the local entrepreneurial environment. According to the academic, policy, and business literature, entrepreneurial ecosystems are a critical tool for creating resilient economies based on entrepreneurial innovation (Spigel, 2017). Thus, the role of BIs and science parks should be highlighted due to their increasing relevance in recent years (Barbero, Casillas, Wright, & Garcia, 2014); moreover, science parks may integrate BIs within their campuses (Squicciarini, 2008).

BIs are organizations that create value through the provision of spaces and/or utilitarian services for start-ups and companies to assure their sustainable development (Tötterman & Sten, 2005); they can make a significant contribution to regional development, as demonstrated in south-eastern Spain (Mas-Verdú, Ribeiro-Soriano, & Roig-Tierno, 2015).

* Corresponding author. E-mail: goncalo.nr.bras@gmail.com.

DOI: 10.17811/ebl.8.3.2019.147-155
In the framework of entrepreneurial ecosystems, the importance of the role played by BIs makes it relevant to explore geographically bounded factors that may influence their location or concentration in some Portuguese regions. This research adds to the scarce literature on the topic due to its novelty in Portugal; moreover, it responds to the need for studies that shed light on the dispersion of BIs across Portugal. Despite the growing number of studies on business incubators as a formal mechanism for business creation (Schwartz & Hornych, 2010), these rarely address the factors influencing their location. Exceptions to this are the studies by Qian, Haynes, and Riggle (2011) and by Yu, Middleton, and Jackson (2010) which focus on BI clusters in USA and explore significant variables to explain the location of BIs; they are therefore of relevance to our specific topic. However, the research on BIs typically examines topics such as their success factors (Grimaldi & Grandi, 2005; Lee & Osteryoung, 2004; Mian, 1996; O'Neal, 2005), the various impacts of their screening processes (Aerts, Matthyssens, & Vandenbempt, 2007; Lumpkin & Ireland, 1988), the consequences of BI specialization (Schwartz & Hornych, 2010; Vanderstraeten, van Witteloostuijn, Matthyssens, & Andreassi, 2016), the link between BIs and innovation (Thierstein & Willhelm, 2001; Xiao & North, 2017) and the cooperation between universities and BIs (McAdam & McAdam, 2008; Mian, 1996).

This gap in the literature led to exploratory research aimed at exploring some variables related to Portuguese regions, namely in relation to their area, population, innovation and the presence of universities that might influence BIs' location in Portuguese districts. Due to the lack of official information from the business incubation activity in Portugal, this essay serves merely as the starting point for further research. Rather than studying the hypothetical contribution of BIs to the development of Portuguese regions based on available information, we address their spatial distribution in Portugal, notably with the aim of understanding the “asymmetrical distribution of BIs across Portuguese districts and the mismatch between the number of BIs and the number of new firms” (Brás & Preto, 2018, p. 275).

The overall aim of this exploratory research is to understand which variables are related to or can influence the location and concentration of BIs in Portuguese districts. Rubin and Babbie (2010: 166) note that “exploratory and descriptive studies, for example, do not test hypotheses”, and this research seeks to accomplish the above-mentioned objective rather than to establish and confirm specific hypotheses. Spatial economics has been studied extensively since von Thünen (1826), Krugman (1991) and other more recent works; however, this exploratory rather than confirmatory research is justified by the lack of a solid theoretical framework on the location of BIs. Nevertheless, this does not stop us from addressing some theoretical topics to support this empirical and exploratory research.

Stokan, Thompson, and Mahu (2015) show that BIs lead to firm growth partly through access to services within and external to the incubator. Besides providing infrastructures and business services to tenant firms, BIs stimulate their network opportunities (Bergek & Norman, 2008; Bollingtoft, 2012), particularly with suppliers, customers, competitors, financial institutions, consultants, investors, among other stakeholders. As such, this might be critical for the selection of a specific geographical location and could explain the creation of business clusters (Peña, 2004); therefore, Henderson, Kuncoro, and Turner (1995) also point out that start-ups might have benefits from both location and urbanization economies.

On the one hand, it is known that major cities act as innovation and creativity ecosystems (Krätke, 2012) and data from USA show that new patents are granted in larger urban centres (Bettencourt, Lobo, & Strumsky, 2007). Thus, patent data is often used as a proxy due to the difficulty of measure regional innovation (Ejermo, 2009). On the other hand, urbanization is usually measured by population density (Gössling & Rutten, 2007).

Tamásy (2007, p. 470) argue that “regions need to have a sufficient population to reach agglomeration economies, and possess a strong university (or other research-oriented organisation) to serve as a good host for an incubator”. However, it should be underlined that
BIs are only one part of the knowledge transfer outcomes in a university, and although BIs play an important role in this domain (Rothaermel & Thursby, 2005), this is not always the case (Berbegal-Mirabent, Lafuente, & Solé, 2013). Furthermore, taking into account the knowledge transfer in universities, the density and size of a region are linked to the innovative activity (Varga, 1998) and possibly to BIs.

To sum up, it is expected that BIs might be concentrated in regions with a university, a high population density, in bigger region areas and with an effective innovation environment (measured by number of patents).

2. Data and methods

The Brás and Preto (2018) database was used to identify the location of Portuguese BIs at the district level. A total of 171 BIs with known addresses were identified in the 20 Portuguese districts. This study uses quartiles to determine the maximum and minimum concentration of BIs in groups of four districts by each quartile – Table 1 and Table 2.

The ordered (by quartiles) and discrete dependent variable ranges from one (1st category – low concentration of BIs) to four (4th category – high concentration of BIs). Moreover, demographic, social, and economic variables drawn from publicly available data were tested at the Portuguese district level to explain BI concentration in these specific regions. This exploratory research seeks to estimate the likelihood of BIs being concentrated in some districts due to the presence of a university through a dummy variable (taking the value 0 or 1 to indicate the absence or presence of a university), based on the number of patent applications or even the population density and area of the districts.

Table 1. Geographical distribution of Portuguese BIs in 2015 and presence/absence of a public university.

<table>
<thead>
<tr>
<th>Portuguese districts</th>
<th>Number of BIs</th>
<th>Presence of at least one public university¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bragança</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Guarda</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Madeira</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>Vila Real</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>Santarém</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Viana do Castelo</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Viseu</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Azores</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Évora</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>Faro</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>Leiria</td>
<td>5</td>
<td>No</td>
</tr>
<tr>
<td>Portalegre</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td>Beja</td>
<td>7</td>
<td>No</td>
</tr>
<tr>
<td>Castelo Branco</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>Coimbra</td>
<td>9</td>
<td>Yes</td>
</tr>
<tr>
<td>Setúbal</td>
<td>9</td>
<td>No</td>
</tr>
<tr>
<td>Braga</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>Aveiro</td>
<td>19</td>
<td>Yes</td>
</tr>
<tr>
<td>Porto</td>
<td>26</td>
<td>Yes</td>
</tr>
<tr>
<td>Lisbon</td>
<td>42</td>
<td>Yes</td>
</tr>
</tbody>
</table>


¹ Only the public university’s (or universities’) main location of was considered, and not decentralized faculties or departments.
Moreover, logit models have been applied to geographical and spatial analysis (Bhat & Sener, 2009) as well as to regional contexts (Yung, Ho, & Chan, 2017; Zhou & Yu, 2017). The Maximum Likelihood method can be used to estimate parameters in ordinal logit regressions. This method can be employed if the observations are assumed to be independent (Hosmer, Lemeshow, & Sturdivant, 2013). The robustness of results was tested by estimating an ordered logit model. This model assumes that there is a natural ranking in the possible values of the dependent variable - BIs concentration (BlSc): a low concentration of BIs (value = 1), average/low concentration of BIs (value = 2), average/high concentration of BIs (value = 3) and high concentration of BIs (value = 4). This means that BlSc can be represented simply as:

\[ \text{BlSc}_i = \sum_{k=1}^{K} \beta_k X_{ik} + \varepsilon_i, \]

where BlSc\(_i\) is a linear function of \(K\) explanatory variables for each BI \((i)\), whose values are \(X_{ik}\) and where \(\beta_k\) represents the coefficients associated to those explanatory variables, namely: (i) number of patent applications, (ii) regional area, (iii) presence/absence of a university (dummy variable) and (iv) population density; \(\varepsilon_i\) is the error term\(^2\).

This ordered logit model for an ordinal response BlSc\(_i\) with \(C\) categories can be represented as an underlying continuous response with a set of \(C - 1\) cut points \(\alpha^*_c\) such that \(\text{BlSc}_i = \text{BlSc}_c\) if and only if \(\alpha^*_{c-1} < \text{BlSc}_i \leq \alpha^*_c\) through the cumulative logistic function\(^3\).

The endogeneity hypothesis of the explanatory variables was tested for the use of instrumental variables (IV), such as Population level, R&D firms, Higher Institutions related to science and technology, Number of higher education courses and Regional area under logarithm form. Hence, the Durbin-Wu-Hausman Test (Durbin, 1954; Hausman, 1978; Wu, 1973) was performed:

\[ H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS}) [\text{Var}(\hat{\beta}_{IV}) - \text{Var}(\hat{\beta}_{OLS})]^{-1} [\hat{\beta}_{IV} - \hat{\beta}_{OLS}] \chi^2_k \]

Using Durbin-Wu-Hausman Test, with a 5% significance level, it is concluded that all explanatory variables are exogenous since they have a \(p\)-value > 0.05\(^4\).

\(^2\) The error term has a variance \(\sigma^2 \sigma_i^2\), where \(\sigma_i^2 = h(h(z_i + \alpha^*_\sigma); h}\) is a given function, \(z_i\) is a vector of explanatory variables, \(\alpha^*_\sigma\) is normalized such that \(h(\alpha^*_\sigma) = 1\), and \(\sigma^2\) is determined by \(F\) in its logistic cumulative distribution function (ordered logit model).

\(^3\) \(F(\alpha^*_\sigma + X_i' \beta) = 1/(1 + \exp [-\alpha^*_\sigma - X_i' \beta])\)

\(^4\) Number of patent applications (\(p\)-value = 0.110123); Regional area (\(p\)-value = 0.948803); Presence/Absence of university (\(p\)-value = 0.522046); Population density (\(p\)-value = 0.445131).
3. Results

Considering the BIs concentration as the outcome variable, Table 3 shows the estimation results and the robustness indicators.

Table 3. Estimation results.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Ordered Logit</th>
<th>Robust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>Number of patent applications</td>
<td>.1063</td>
<td>.0144</td>
</tr>
<tr>
<td>Regional area</td>
<td>.0006</td>
<td>.0001</td>
</tr>
<tr>
<td>Presence/absence of university (dummy)</td>
<td>2.0228</td>
<td>.6534</td>
</tr>
<tr>
<td>Population density</td>
<td>.0056</td>
<td>.0042</td>
</tr>
<tr>
<td>_cut1</td>
<td>5.6054</td>
<td>1.2926</td>
</tr>
<tr>
<td>_cut2</td>
<td>7.5068</td>
<td>1.4648</td>
</tr>
<tr>
<td>_cut3</td>
<td>11.9427</td>
<td>1.7156</td>
</tr>
</tbody>
</table>

Wald test Wald chi2(4) = 76.61; Prob > chi2 = 0.0000
Pseudo R2 0.6516

There is no standard way of interpreting the coefficients of a logit regression: However, the positive signs of all coefficients and the absence of statistical significance of population density should not be ignored when explaining BIs' concentration. The Wald tests confirm the global consistency of the specified models, given that all the coefficients in the models are different than zero. In relation to pseudo R-squared value, the model has an excellent fit based on the thresholds defined by McFadden (1973).

However, following Hamilton (2012), the scores (S) can be stated a linear function of the explanatory variables:

\[ S_{\logit} = 0.1063PA + 0.0006A + 2.0228D \]  

(3)

According to Greene (2003), given the cut points, it is possible to predict probabilities for the ordered logit model. These refer to the cut points used to differentiate the adjacent levels of the BIs concentration and indicates the probability of a BI belong from a low to a high concentration area of BIs when values of the independent variables are evaluated at zero.

It is important not only to predict probabilities, but also to interpret the marginal effects, that is, the extent to which the (conditional) probability of the outcome variable changes when there is a change in the value of a regressor, ceteris paribus. In this case, the marginal effects show the change in probability for each BI concentration category (outcome variable) when the predictor or independent variable increases by one unit – Table 4.

Table 4 shows that the university dummy variable has a positive marginal effect for the last two categories of BI concentration, but a negative effect on the two first categories. This indicates there is more likely to be a higher BI concentration in Portuguese districts with universities.

\[ P(y_{\text{ordinal}} = \text{"low concentration of BIs"}) = P(S_{\logit} + u \leq \_cut1) = P(S_{\logit} + u \leq 5.6054) \]
\[ P(y_{\text{ordinal}} = \text{"average/low concentration of BIs"}) = P(\_cut1 < S_{\logit} + u \leq \_cut2) = P(5.6054 < S_{\logit} + u \leq 7.5068) \]
\[ P(y_{\text{ordinal}} = \text{"average/high concentration of BIs"}) = P(\_cut2 < S_{\logit} + u \leq \_cut3) = P(7.5068 < S_{\logit} + u \leq 11.9427) \]
\[ P(y_{\text{ordinal}} = \text{"high concentration of BIs"}) = P(\_cut4 < S_{\logit} + u ) = P(11.9427 < S_{\logit} + u) \]
Table 4. Average marginal effects.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Ordered logit (categories)</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patent applications</td>
<td>-.00502</td>
<td>-.00169</td>
<td>.00452</td>
<td>.00220</td>
<td></td>
</tr>
<tr>
<td>Regional area</td>
<td>-.00003</td>
<td>-.00001</td>
<td>.00003</td>
<td>.00001</td>
<td></td>
</tr>
<tr>
<td>Presence/absence of university (dummy)</td>
<td>-.09560</td>
<td>-.03213</td>
<td>.08589</td>
<td>.04185</td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>-.00026</td>
<td>-.00009</td>
<td>.00023</td>
<td>.00012</td>
<td></td>
</tr>
</tbody>
</table>

Note. All coefficients except “population density” are statistically significant at the 1% level of significance.

Similarly, the marginal effect of the other continuous variables is negative for the two lowest categories of BI concentration, but positive for the two highest categories.

When the concentration of BIs is low or average/low, the district area decreases the probability of having a concentration of BIs. In Portuguese districts with a high concentration of BIs, the district area increases the probability of having a concentration of BIs.

In the same vein, when the concentration of BIs is low or average/low, the number of patent applications reduces the probability of having a concentration of BIs. In Portuguese districts with a higher concentration of BIs, the number of patent applications raises the probability of having a concentration of BIs. Population density did not show statistical significance to explain concentration of BIs at the Portuguese district level.

4. Concluding remarks

The results indicate there is more likely to be a strong concentration of BIs in Portuguese districts with one or more universities in the vicinity, where the number of patent applications is higher and in the largest areas. Taking into account the value-added contributions of university technology business incubators to tenant firms, the geographical proximity between Portuguese BIs and universities is in line with the findings of Mian (1996). A Portuguese entrepreneurial university may positively influence its regional development (Ferreira, Leitão, & Raposo, 2006), but little is known about the role of the knowledge transfer among all Portuguese universities and BIs. Despite this, the research presented here seeks to raise policymakers’ awareness of the relevance of a desirable holistic approach to the distribution of Portuguese BIs, as advocated by Spigel (2016) in Scotland.

The confirmed prominence of a university near BIs implies a holistic reflexion from Portuguese policymakers about the entrepreneurial ecosystem. Therefore, additional global policies are required with the capacity to engage BIs, universities and regional clusters, as opposed to the current “go with the flow” policy that prevails in Portugal in this domain.

Although the empirical knowledge on how the business incubator and the university relate to each other is still very limited (Rothschild & Darr, 2005), Gstraunthaler (2010, p. 400) argued that “business incubators are often seen as part of a resource transfer to enable firms to develop the inventions made in the university environment”, this also helps explain the probability of a strong concentration of BIs where the number of patent applications is higher. In this vein, a strong concentration of BIs is more likely to be found in Portuguese districts that have a larger number of specialized BIs. Squicciarini (2008) notes that science park tenants have a higher number of patent applications, which indirectly confirms the link between the concentration of BIs and the number of patent applications (in Portuguese districts with specialized BIs). However, the influence of patents should be viewed with caution because knowledge transfer depends on regional factors related to technological development and entrepreneurial culture (Berbegal-Mirabent et al., 2013).

The district area decreases the probability of having a strong concentration of BIs because the larger Portuguese districts are in rural areas. However, after an inflection point, in the two last
categories (average/high and high number of BIs), the district area increases the probability of having a strong concentration of BIs, as they belong to metropolitan areas; this is similar to the findings of Qian et al. (2011).

Nevertheless, the relevance of population density to explain the location of BIs was not confirmed; this is in contrast to findings for the USA (Qian et al., 2011). This might be explained by the regional idiosyncrasies in Portugal; for instance, some Portuguese districts have a low concentration of BIs and high population density (Madeira), while others have average/high concentration of BIs and low population density (Beja, Portalegre and Castelo Branco).

Due to the lack of official information about BIs and their location in Portugal, this cross-sectional work is limited to 2015. Longitudinal data would consolidate our findings but would exist consistent, harmonized, systematic and comparable data provided by a Portuguese institution. This would allow us to specify a panel data model or even a dynamic panel data model that sheds light on the long and short run effects on BIs concentration.

Acknowledgments

Article financed by Faculty of Economics University of Coimbra (project FEUC/IE) and by FEDER—Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020—Operacional Programme for Competitiveness and Internationalisation (POCI), and by Portuguese funds through FCT—Fundação para a Ciência e a Tecnologia in the framework of the project PTDC/IVC PEC/5514/2014. Both authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the paper.

References


