Location Modelling and the Localization of Portuguese Manufacturing

Paulo Guimarães
Universidade do Minho and CEMPRE

Octávio Figueiredo
Universidade do Porto and CEMPRE

Douglas Woodward[1]
University of South Carolina

May 17, 2004

[1]Corresponding author. Address: The Moore School of Business, University of South Carolina, Columbia, South Carolina. E-mail: woodward@moore.sc.edu.

Abstract

The recent index proposed in Ellison & Glaeser (1997) is now well established as the preferred method for measuring localization of economic activity. We critically review this index and build on the McFadden's Random Utility (Profit) Maximization framework to develop an alternative measure that is more consistent with the theoretical construct underlying the original work of Ellison and Glaeser. Given that our method is regression based it goes beyond the descriptive nature of the EG index and allows us to evaluate how the localization measure behaves with changes in the systematic forces that drive firms' location decisions.

JEL classification: C25, R12, R39.

1 Introduction

Agglomeration is widely recognized as a source of increasing returns for individual firms in particular industries. For more than a century economists have examined why and to what extent these localization economies–internal to the local industry, but external to the firm–explain the spatial concentration of economic activity. Casual empiricism suggests that there is a marked tendency for industries to localize, i.e. to concentrate over and above overall economic activity. Alfred Marshall's classic examples included cutlery (Sheffield) and jewelry (Birmingham) in 19th century England. Contemporary examples abound, from the automotive industry in Michigan and semiconductors in California, to the often-cited footwear cluster of northern Italy and telecommunications in Finland. Yet how general and how strong is the tendency of industry to agglomerate in local areas?

The debate reignited by the advent of the "new economic geography", with its emphasis on the importance of external economies, has again brought these questions to the forefront of many scientists' research agendas. However, clear answers have been marred by the lack of an adequate approach to the measurement of an industry degree of localization. More recently, Ellison & Glaeser (1997) tackled this problem. Based on a Random Utility (Profit) Maximization model of industrial location (henceforth RUM), they proposed an index that captures the effect of those non-systematic forces (spillovers and natural advantages of the regions) that lead to spatial concentration of firms. In a short period of time, their work spawned a significant number of studies and rapidly emerged as the standard approach for measuring localization of economic activity. Yet, and despite its significant contribution, the index of Ellison & Glaeser (1997) (henceforth EG index) treats the systematic forces that lead to spatial concentration (e.g. wages, land costs, market accessibility and transportation costs) as a black-box. In their view, in the absence of natural advantages or spillovers, all regions exert the same pull on firms, regardless of sector of activity.

In this paper we critically review the EG index and contend that the link between the RUM framework and this index is fragile and should be strengthened. Thus, we build on McFadden's RUM framework to propose an alternative measure that is more faithful to the theoretical construct presented in Ellison and Glaeser's original work. Because we explicitly model the location decision of firms (our index is directly derived from a discrete choice model) we are able to go beyond the descriptive nature of the EG index and evaluate how the localization measure behaves with changes in those systematic forces that affect the firms' profit function.

The rest of the paper consists of four sections. The following section reviews traditional measures of spatial concentration and attendant problems. In section 3, we take a more in-depth look at the EG index and develop our alternative method for measuring localization. Section 4 provides an illustration using data on Portuguese industries and section 5 concludes.

2 The Measurement of Spatial Concentration

Past economists had no shortage of tools for measuring the geographical concentration of economic activity. Most prominent are Hoover's (1937) location quotient and a form of Gini coefficient, as applied by Krugman (1991). These measures quantify the discrepancy between the distribution of regional employment in a particular industry against the regional distribution of overall employment. But are these measures able to capture the concept of localization? A first obvious problem is that they are sensitive to the levels of concentration within the industry. Take as an example two industries which have identical measures for the Gini index. The first industry is composed of many independent firms, all equally sized and located in a single region, while the second industry is composed of just one firm operating a large establishment. The first case agrees more with the notion of spatial external economies, which may explain the clustering of all firms in that industry. But for the second industry, it is obvious that external economies are not a valid explanation to justify spatial concentration. In this second case, geographic concentration is entirely explained by industrial concentration and then by returns to scale.

Another problem is that these measures do not account for the inherent randomness of the underlying location decisions. Firms may exhibit some level of spatial concentration by chance. This idea can be explained by appealing to the balls and urns example often used in statistics. If one has, say, 10 urns (regions) and 10 balls (firms) and drops the balls at random into these urns then, even though all urns are equally probable, it is very unlikely that we will observe exactly one ball in each urn. Some clustering will necessarily occur and that is perfectly compatible with the idea that the balls were thrown at random (the firms' decisions were random). The above indexes are not able to control for this type of clustering.

It should be obvious, then, that these indices do not accurately measure an industry's degree of localization. The recent index proposed in Ellison & Glaeser (1997) overcomes some of the limitations. Like the Gini coefficient, the EG index attempts to measure the tendency of one industry to agglomerate in relation to the general tendency of all industries to agglomerate. Unlike its predecessors, however, it accounts for the inherent discreteness (lumpiness) that will be observed if location decisions are driven by chance alone and it expurgates the effect of industrial concentration, offering a standardized measure that can be readily used for temporal or inter-sectorial comparisons. Most notably, the EG index is rooted in the location choice model of Carlton (1983), which in turn is based on McFadden's RUM framework and has been the workhorse for empirical research on industrial location [e.g. Bartik (1985), Luger & Shetty (1985), Hansen (1987), Schmenner, Huber & Cook (1987), Coughlin, Terza & Arromdee (1991), Woodward (1992), Friedman, Gerlowski & Silberman (1992), Head, Ries & Swenson (1995), Guimarães, Figueiredo & Woodward (2000), and Figueiredo, Guimarães & Woodward (2002)].

Other authors have proposed measures closely related to the EG index. Based on a different theoretical argument, Maurel & Sedillot (1999) constructed an index which is similar to the EG index. By comparing the two formulas, they show that the difference between the indices has an expected value of zero. Also noteworthy is the work of Devereux, Griffith & Simpson (2004). They showed that the index of Ellison and Glaeser can be conveniently approximated by the difference between an index that measures geographic concentration and another that measures industrial concentration. In turn, Duranton & Overman (2002) have proposed a different approach to the measurement of spatial concentration. Their approach draws directly from methods well-known to spatial statisticians to measure concentration of spatial phenomena. They treat space as continuous and compute their measurements based on the Cartesian distances between each pair of plants. Treating space as continuous has an inherent appeal but their approach lacks a theoretical underpinning. Moreover, it is an essentially descriptive procedure that requires precise information (often unavailable) on the exact location of each business unit.

The new wave of literature initiated by Ellison & Glaeser (1997) has already generated a substantial amount of applied work. Beyond the ongoing research in the United States [Ellison & Glaeser (1997), Dumais, Ellison & Glaeser (2002) and Holmes & Stevens (2002)], recent studies characterizing industry localization can be found for France [Maurel & Sedillot (1999) and Houdebine (1999)], Belgium [Bertinelli & Decrop (2002)], UK [Devereux et al. (2004)] and Spain [Callejón (1997)]. Common to all studies is the finding that the majority of industries are localized.

In the following, we offer an approach to the measurement of localization of economic activity that builds on the conceptual approach of Ellison & Glaeser (1997), yet is grounded more solidly on the RUM framework. As will become clear in the next section, the link between the RUM location literature and the EG index is feeble. We show how the two can be better integrated. Also, contrary to the trend in the literature, we argue that using employment figures confounds the measurement of localization [as proposed by Ellison & Glaeser (1997)] and advocate the use of counts of plants.

3 Methodological Issues

Industrial location models based on the RUM framework provide an explanation for the spatial distribution of an industry. Idiosyncratic factors aside, firms choose locations that yield the highest profits. If we abstract from the dynamic questions, we can use the RUM theoretical framework to justify the geographic concentration of industries. Ellison & Glaeser (1997) used this approach. Yet, as argued here, the integration between the RUM and the derivation of the EG index can be strengthened to provide a more theoretically sound way to measure the degree of localization of an industry.

3.1 The EG Index

To motivate our approach, we now take a closer look at the derivation of the EG index. Let us assume at the outset that the economy is divided into J geographical units (regions). Also, we take as our reference a given industry which has exactly n_j plants located in each region j. Thus, $n = \sum_{j=1}^{J} n_j$ represents the total number of existing plants in our reference industry. Next, we briefly sketch how the EG index is obtained taking as a reference their model of "natural advantages". If firm i chooses to locate in region j then its profits will consist of

$$\ln \pi_{ij} = \ln \overline{\pi}_j + \varepsilon_{ij} \tag{1}$$

where $\overline{\pi}_j$ is a non-negative random variable reflecting the profitability of locating in area j for a typical firm in the industry. In this formulation of the model, Nature introduces the randomness in $\overline{\pi}_j$ by selecting for each region the characteristics that make it unique (their natural advantages). ε_{ij} is a random disturbance. If we assume that ε_{ij} is an identically and independently distributed random term with an Extreme Value Type I distribution[1] then, conditional on a realization of $\overline{\pi}_j$, we can apply McFadden's (1974) result to

[1] In the past this distribution has been referred to by other names such as Weibull, Gumbel and double-exponential [Louviere, Hensher & Swait (2000)].

obtain,

$$p_{j/\pi} = \frac{\exp(\ln \overline{\pi}_j)}{\sum_{j=1}^{J} \exp(\ln \overline{\pi}_j)} = \frac{\overline{\pi}_j}{\sum_{j=1}^{J} \overline{\pi}_j} \ , \tag{2}$$

which denotes the probability of a firm locating in region j. Thus, p_j is obtained from the Random (Profit) Utility Maximization framework of Carlton (1983) which, as mentioned earlier, gives support to the most recent studies of industrial location. To derive their index, Ellison & Glaeser (1997) introduced two parametric restrictions regarding the expected value and variance of p_j. Thus, they assume that the distribution of $\overline{\pi}_j$ is such that:

$$E(p_j) = x_j \ , \tag{3}$$

and that,

$$V(p_j) = \gamma x_j (1 - x_j) \ , \tag{4}$$

where x_j may be thought of as the probability of a firm locating in region j in the absence of any region specific advantages for that industry. Thus, the larger the discrepancy between x_j and p_j, the larger the influence that these region specific effects (say, natural advantages) play in the location decisions of firms in that industry. That difference is captured by the parameter γ (which we will refer to as the EG parameter) which belongs to the unit interval. It is easy to see that if $\gamma = 0$ then the industry will tend to replicate the pattern observed for the x_j (what Ellison and Glaeser call the dartboard model) and we can conclude that there is no spatial concentration in excess of what we would expect to occur. If, however, $\gamma > 0$, then the actual location probabilities of the industry will differ from x_j and in the limit, when $\gamma = 1$, each p_j has the largest variance and becomes a Bernoulli random variable. Thus, in the limit, all the investments for that industry would be located in a single region.

Ellison & Glaeser (1997) also show that the γ parameter may be derived from an alternative model that emphasizes industrial spillovers as the force leading to "excessive concentration". In any case, the theoretical motivation one uses is irrelevant because the two models are observationally equivalent and lead to the same functional form for the index, the practical implication being that we can not readily distinguish the two sources of geographic concentration (natural advantages and industrial spillovers).

To estimate γ for a particular industry they let x_j denote area j's share of total manufacturing employment. Here, the idea is that the model should on average reproduce the overall distribution of manufacturing activity. In a next step they considered the following "raw concentration index" of employment:

$$G_E = \sum_{j=1}^{J} (s_j - x_j)^2 \tag{5}$$

where, s_j denotes area j's share of employment in that industry and the x_js are as described above. Now, taking the expected value of G_E they obtain a function of γ and the authors use that relation to propose an estimator for γ. Their proposed estimator for γ (the EG index) is then

$$\widehat{\gamma}_{EG} = \frac{G_E - \left(1 - \sum_{j=1}^{J} x_j^2\right) H_E}{(1 - \sum_{j=1}^{J} x_j^2)(1 - H_E)} \ , \tag{6}$$

where H_E is the employment Herfindhal index for the industry and the expected value of G_E is replaced by its actual value. Note that the computation of the Ellison and Glaeser measure of localization requires employment and plant size information. But, as we will see next, one could obtain a more efficient estimator for γ relying on counts of plants.

3.2 A EG Index Based on Counts of Plants

Defining the "raw concentration index" as

$$G_F = \sum_{j=1}^{J} \left(\frac{n_j}{n} - x_j \right)^2 \tag{7}$$

and proceeding in a fashion similar to Ellison and Glaeser (see Appendix A) we derive the following alternative estimator for γ:

$$\widehat{\gamma}_A = \frac{nG_F - \left(1 - \sum_{j=1}^{J} x_j^2\right)}{(n-1)\left(1 - \sum_{j=1}^{J} x_j^2\right)} . \tag{8}$$

The above expression is very similar to that of the Ellison and Glaeser index. In the "raw concentration index", s_j is expressed in terms of plants (instead of employment) and the Herfindhal index is replaced by $1/n$. Like the estimator proposed by Ellison and Glaeser this estimator for γ is also, by construction, unbiased. Most notably, it has a much smaller variance. To see this note that:

$$\frac{V\left(\widehat{\gamma}_{EG}\right)}{V\left(\widehat{\gamma}_A\right)} = \left(\frac{n-1}{n(1-H_E)}\right)^2 \frac{V\left(G_E\right)}{V\left(G_F\right)}. \tag{9}$$

An heuristic argument suffices to justify the better efficiency of this estimator. If all plants had the same dimension, the indexes would be identical (H_E would be $1/n$). As the Herfindhal index increases, the first term of the product in the RHS of (9) increases. One would also expect the second term (the ratio of the variances) to be larger with increases in the Herfindhal index. Thus, we argue that a more precise estimate for γ is obtained if we ignore the confounding influence of plant size (employment) and work directly with counts of plants. From another perspective, Holmes & Stevens (2002) provide additional evidence against the use of an index based on employment plant size. These authors found evidence that plants located in areas where an industry concentrates (as measured by the EG index) are larger, on average, than plants in the same industry outside the same area, thus suggesting that the EG index will tend to overstate the degree of localization of an industry.

A clear disadvantage of the EG index is that it does not provide an indication of statistical significance.[2] In Appendix B we show how one can

[2] However, in latter work, Maurel & Sedillot (1999) provide an approximate test for the null hypothesis that $\gamma = 0$.

construct and implement an exact (non-parametric) test for the null hypothesis that $\gamma = 0$ for the $\widehat{\gamma}_A$ statistic.

3.3 An Alternative Method for Measuring Localization

An implicit assumption in the work of Ellison and Glaeser is that in the absence of natural advantages (or spillover effects) all individual industries would be faced with the same location probabilities, $p_j(= x_j)$. If these p_js are obtained from the RUM framework, as is claimed, then this amounts to the underlying assumption that all industries would have identical profit functions. But, the systematic forces that drive the location of a chemical plant may be very different from those driving the location of an apparel plant. In other words, we claim that if natural advantages (or spillovers) were inexistent then one would still expect to find different patterns of location across industries, simply because industries value regional characteristics differently. For example, wages may be an important component of the profit function for the apparel industry but may not be a determinant factor in the locational decisions of chemical plants. To incorporate this dimension into the framework laid out by Ellison and Glaeser, we take a different route - we explicitly model the location decision process of firms and measure concentration in excess of that which would result if all industries were influenced by the same set of (observed) locational factors. That is, instead of approximating the "attractiveness" of a region by its share of manufacturing employment[3], we let each industry have a different valuation for the "attractiveness" of a region based on the particular combination of factors that are relevant for that industry.

Hence, we admit that the profit function faced by firm i in our reference

[3]At this point it should be noted that Ellison and Glaeser report the use of other alternatives to manufacturing employment such as the area and the population.

industry, if it decides to locate in region j, may be written as,

$$\log \pi_{ij} = \boldsymbol{\theta}'\mathbf{y}_j + \eta_j + \varepsilon_{ij} \ , \tag{10}$$

where, the $\mathbf{y}_j$ are regional characteristics that affect the location decisions of firms in all industries (systematic forces such as wages, land costs, market accessibility and transportation costs), $\boldsymbol{\theta}$ is a vector of parameters, and η_j is a (regional) random effect that picks the unobservable (non-systematic) locational advantages of that region for a particular industry. The other random term, ε_{ij}, is as defined earlier. Now, conditional on the $\eta_j s$ and again drawing on McFadden's (1974) result we can write,

$$p_{j/\boldsymbol{\eta}} = \frac{\exp(\boldsymbol{\theta}'\mathbf{y}_j + \eta_j)}{\sum_{j=1}^{J}\exp(\boldsymbol{\theta}'\mathbf{y}_j + \eta_j)} = \frac{\exp(\eta_j)\lambda_j}{\sum_{j=1}^{J}\exp(\eta_j)\lambda_j}. \tag{11}$$

The likelihood function (conditional on the $\eta_j s$) implied by the above expression is that of a conditional logit model:

$$L(n_1, n_2, ..., n_J/n, \boldsymbol{\eta}) = \prod_{j=1}^{J} p_{j/\boldsymbol{\eta}}^{n_j}. \tag{12}$$

which in turn is the kernel of a multinomial distribution with parameters $(p_{1/\boldsymbol{\eta}}, p_{2/\boldsymbol{\eta}}, ..., p_{J/\boldsymbol{\eta}}, n)$,

$$L(n_1, n_2, ..., n_J/n, \boldsymbol{\eta}) \propto n! \prod_{j=1}^{J} \frac{p_{j/\boldsymbol{\eta}}^{n_j}}{n_j!}. \tag{13}$$

Now, if we assume that the $\exp(\eta_j)s$ are i.i.d. gamma distributed with parameters $(\delta^{-1}, \delta^{-1})$- and thus with variance equal to δ-, then $\exp(\eta_j)\lambda_j$ also follows a gamma distribution with parameters $(\delta^{-1}\lambda_j, \delta^{-1})$. We know from Mosimann (1962) that in this case the $(p_1, p_2, ..., p_J)$ are Dirichlet distributed with parameters $(\delta^{-1}\lambda_1, \delta^{-1}\lambda_2, ..., \delta^{-1}\lambda_J)$. Therefore the unconditional likelihood function may be written as,

$$L(n_1, n_2, ..., n_J/n) = n! \int \prod_{j=1}^{J} \frac{p_j^{n_j}}{n_j!} g(p_1, p_2, ..., p_J) dp_1 dp_2, ..., dp_J \ . \tag{14}$$

The above integral has a closed form, whose solution is known as the Dirichlet-Multinomial distribution (Mosimann 1962):

$$L(n_1, n_2, ..., n_J/n) = \frac{n!\Gamma(\delta^{-1}\lambda_\bullet)}{\Gamma(\delta^{-1}\lambda_\bullet + n)} \prod_{j=1}^{J} \frac{\Gamma(\delta^{-1}\lambda_j + n_j)}{\Gamma(\delta^{-1}\lambda_j)n_j!} \tag{15}$$

where $\lambda_\bullet = \sum_{j=1}^{J} \lambda_j$. The resulting likelihood function offers no particular challenge and can be easily implemented. But the interesting feature of this approach is that now,

$$E(p_j) = \frac{\lambda_j}{\lambda_\bullet} \tag{16}$$

and

$$V(p_j) = \frac{1}{\delta^{-1}\lambda_\bullet + 1} . \frac{\lambda_j}{\lambda_\bullet} . \left(1 - \frac{\lambda_j}{\lambda_\bullet}\right) \tag{17}$$

and by analogy with (4) and the approach of Ellison and Glaeser we can interpret

$$\widetilde{\gamma} = \frac{1}{\delta^{-1}\lambda_\bullet + 1} = \frac{\delta}{(\lambda_\bullet + \delta)} \tag{18}$$

as an index of excessive spatial concentration for that industry, that is, an alternative estimator for the EG parameter. As δ (the variance of the region specific random error) increases, so does $\widetilde{\gamma}$ and in the limit, when δ tends to infinity, $\widetilde{\gamma}$ will tend to 1. On the other hand, $\widetilde{\gamma}$ will approach zero as δ tends to zero.[4] Because in this latter situation the Dirichlet-Multinomial distribution collapses to a standard multinomial distribution we can use a likelihood ratio test to test the hypothesis that the industry is more concentrated than what we would expect ($\delta = 0$).[5] To implement our model, we wrote the likelihood function in Stata (version 7) using that package's standard numerical

[4]Unlike the EG index, which often produces negative estimates, our estimator will always generate estimates that belong to the unit interval.

[5]Because we are testing a value which is in the boundary of the set of admissible values for δ, we follow the suggestion in Cameron & Trivedi (1998) and adjust the level of significance of the chi-square statistic accordingly. Also, we should note that to apply the likelihood ratio test, we need to rescale the likelihood function of the Conditional Logit model as in (13).

maximization routine (a modified Newthon-Raphson algorithm). To obtain starting values, we first estimated a Poisson regression- which in this context produces the same estimates for the variable coefficients as the conditional logit model [Guimarães, Figueiredo & Woodward (2003)]. Convergence was fast with a very small number of iterations (less than 10 for most cases).[6]

4 An Empirical Application: Localization of Portuguese Manufacturing Industries

4.1 Data and Variables

The availability of detailed plant establishment information by industry allowed us to apply our model to Portugal. Our main source of data was the "Quadros do Pessoal" database for 1999, the most recent available year. The "Quadros do Pessoal" is a yearly survey collected by the Ministry of Employment for all the existing companies operating in Portugal (except family businesses without wage earning employees) and covers 45,350 plants for the year of 1999.[7] Using this source, we tallied the number of plants as well as employment for each "concelho" in continental Portugal.[8] We rely on the 3-digit (103 industries) classification of the Portuguese Standard Industrial Classification system (CAE).[9] Using the 275 Portuguese "concelhos" as the spatial choice set, we estimated a location regression for each industry

[6] However, for eight industries the model did not converge. We took it as evidence that the data were not overdispersed enough. For these cases we let $\widetilde{\gamma} = 0$.

[7] For a thorough description of this database see, for example, Mata, Portugal & Guimarães (1995) and Cabral & Mata (2003). Unless otherwise noted the "Quadros do Pessoal" was the source for all the information used in this paper.

[8] The concelho is an administrative region in Portugal. In recent years some new concelhos have been created by the incorporation of parts of existing "concelhos". To maintain data compatibility, we used the spatial breakdown of 275 "concelhos" that was still valid in 1997. These have an average area of 322.5 squared kilometers.

[9] Revision 2 of the CAE.

(the Dirichlet-Multinomial model), as well as the corresponding measure of excessive concentration (localization) given by (18).

The choice of regressors for our location model was dictated by location theory. Location theory distinguishes three different sets of factors driving the firm's location decision problem: external economies, costs of production factors, and accessibility (transportation costs) to input and final demand markets. External economies can arise from two different sources. Localization economies are those external economies that result from the spatial concentration of firms of a particular industry in a given region and that are internalized by firms of that particular industry. In our model, this effect on firm's location decisions is captured through $\widetilde{\gamma}$ (along with natural advantages of the regions). The other externality, urbanization economies, accrues from the clustering of general economic activity in a given area and benefits all plants locating in that particular area. Urbanization economies are proxied in our model by the "concelho" density of service and manufacturing establishments per square kilometer in 1999.

To control for the impact of factor prices, we obtained information on the cost of labor and land. Labor costs are measured by an index of the average manufacturing base wage rate in 1999.[10] Since industrial and residential users compete for land, one may argue that when modeling with small areas and controlling for urbanization, as in our case, land costs can be proxied by population density. Consequently, following the suggestion of Bartik (1985), we use population density to approximate land costs.[11] We did not consider the cost of capital because it is practically invariant across alternatives. Interest rates do not differ regionally, and despite some minor differences in municipal taxes, the overall tax burden on manufacturing activity comes

[10] Because we are not using real wages, a higher average manufacturing base wage rate in a given "concelho" can also indicate the presence of a highly-skilled workforce. If investors are willing to pay higher wages for more qualified workers, the coefficient of this variable is expected to be positive.

[11] We used population for the year of 1996, taken from the National Institute of Statistics (INE).

mostly from taxes set at the national level.

To account for market accessibility at a given location (and transportation costs) we enter two variables in the model. The drive time distance from each "concelho" to the Porto-Lisbon corridor (the more urbanized coastal side of the country) measures large-scale accessibility, i.e. access to the largest markets. Small-scale accessibility, i.e., access to regional markets, is proxied by the distance in time by road from each "concelho" to the administrative center (the capital) of the related "distrito".[12]

4.2 Results

4.2.1 Localization of Portuguese Industries

We computed the localization index $\widetilde{\gamma}$ for each of the 3-digit SIC industries at the "concelho" level.[13] As indicated before, for a small number of industries (8) the model did not converge and we assumed that γ was zero. For 17 industries, the $\widetilde{\gamma}$ index was not statistically different from zero at the 95 per cent level of confidence. For the remaining 75 industries (75 per cent of the 100 industries analyzed) we find evidence of "excess of concentration" ($\gamma > 0$). Therefore, a high percentage of Portuguese manufacturing industries appear to be localized. This result corroborates similar evidence for others countries.[14]

[12] The "distrito" is an higher administrative region level composed of several adjacent "concelhos". Continental Portugal is divided in 18 "distritos". The time distance variables report to the year of 1996. They were constructed using an algorithm that selected the shortest time route between locations, using as parameters the average traveling speed for the particular type of road as well as a road network compiled from road maps (ACP 1998/9; Michelin 1999) and detailed information from the Portuguese Road Institute (Instituto Português de Estradas). We thank Adelheid Holl for making this unpublished data available for the present study.

[13] Our dataset contains information for 100 3-digit SIC industries. For SICs 231, 233, and 300, the "Quadros do Pessoal" dataset did not report any plant in 1999.

[14] Ellison & Glaeser (1997) found that 446 out of 459 4-digit SIC industries in the United States were localized ($\widehat{\gamma}_{EG} > 0$). Based on a test of statistical significance Maurel

As previously observed for others countries as well, the localization index displays a very skewed distribution, the majority of industries showing slight levels of localization. This pattern is displayed in Figure 1, where we show a histogram of $\widetilde{\gamma}$ at the "concelho" level for the 100 3-digit SIC industries.

(Figure 1, Page 27)

Tables 1 and 2 provide information for individual industries. In Table 1 we list the 22 sectors that have a $\widetilde{\gamma}$ index that is significantly different from zero and above the industry average. Among them, we find a large number of traditional sectors for which localization is associated with the historical specialization of Portuguese particular regions (e.g. tannery, jewelry, textiles, footwear, and cork industries). Again, this pattern coincides with evidence for other countries that suggests that typically traditional industries are highly localized.[15] Table 1 also shows that several more technologically advanced industries (such as fabrication of radio and television apparatus, artificial and synthetic fibers, automobiles, and measuring and controlling devices) exhibit higher than average levels of localization. As could be expected, shipbuilding and industries that process sea products are also among the most localized industries.

Table 2 displays the group of non-localized sectors (i.e. those for which we do not reject the null hypothesis that $\gamma = 0$).[16] For this last group it is important to distinguish our measure of localization from a simple measure of geographic concentration. While some of these sectors (such as tobacco, petroleum refining or aircraft and space vehicles fabrication) are highly concentrated in space, this concentration is almost entirely explained by industrial concentration, and thus by returns to scale rather than natural advantages or external economies associated with firms' clustering.

& Sedillot (1999) found that 77% of the 273 4-digit French industries display "excess of concentration". Similar results were found for the UK by Devereux et al. (2004).

[15] See Ellison & Glaeser (1997), Table 4, Maurel & Sedillot (1999), Tables 1 and 2, Devereux et al. (2004), Tables 4 and 5, and Krugman (1991), Appendix D.

[16] This Table also includes the eight sectors for which we did not find evidence of overdispersion.

(Tables 1 and 2, Pages 28 and 30)

4.2.2 Comparison with the EG Index

We now compare our estimates of localization of Portuguese manufacturing industries with those provided by the EG index ($\widehat{\gamma}_{EG}$) and the alternative EG index based on counts of plants ($\widehat{\gamma}_A$). If we first look at the extent of localization across the 100 3-digits sectors, we find very similar results for the three measures. 60, 68 and 75 per cent of the industries exhibit "excess of concentration", according to the EG index, the alternative EG index based on counts, and our index, respectively[17].

In Figure 2, we display the box-whisker plots for the three measures. To increase readability the graph omits a few extreme (high) values for each one of the distributions. Clearly, all distributions show the same pattern of skewness with increasing interquartile ranges. Nevertheless, as we anticipated, our proposed measure of localization (labeled as DM index in the figure) produces much smaller estimates for γ when compared with the EG index ($\widehat{\gamma}_{EG}$) and the alternative EG index ($\widehat{\gamma}_A$). We take these results as confirmatory evidence that the original EG index tends to overstate the degree of localization of industries.

(Figure 2, Page 29)

If we now look at the hierarchy of individual industries, we find a significant degree of concordance between the three indexes. The Spearman rank correlation coefficient between $\widetilde{\gamma}$ and $\widehat{\gamma}_{EG}$ is 0.41 and if we consider the rank correlation between the indexes based on counts of plants ($\widetilde{\gamma}$ and $\widehat{\gamma}_A$) this coefficient increases to 0.61.[18] Furthermore, as a quick inspection of Table 1 will reveal, among the top 22 most localized industries according to $\widetilde{\gamma}$ we

[17] These figures are based on the statistical tests of significance indicated before. For the EG index, the test was implemented as in Maurel & Sedillot (1999).

[18] The rank correlation between $\widehat{\gamma}_{EG}$ and $\widehat{\gamma}_A$ is 0.59. All correlation coefficients are statistically different from zero.

find 11 and 13 industries for a similar ranking based on $\widehat{\gamma}_{EG}$ and $\widehat{\gamma}_A$, respectively. Similarly, Table 2 shows that among the 25 non-localized industries according to $\widetilde{\gamma}$ we find 17 and 19 industries that are also classified as non-localized based on $\widehat{\gamma}_{EG}$ and $\widehat{\gamma}_A$, respectively. Thus, despite a substantially different methodological approach, our index produces agreeable results with the other two indexes.

4.2.3 Impact of Changes on Location Factors

Given that we explicitly model the location decision process of firms, we are able to perform exercises of comparative statics to determine how our localization index changes under an alternative scenario for the allocation of regional resources. From expression (18), it is obvious that anything that will increase industry profits will reduce the weight that localization economies (natural advantages of the regions) have on driving the firms' location decisions. If we compute the elasticity of $\widetilde{\gamma}$ with respect to one of the variables entering the profit equation, say variable k, we obtain $-\theta_k(1-\widetilde{\gamma})$ (we are taking into account that all explanatory variables are already entered in logarithmic form). Thus, those variables that are more capable of affecting profits (with higher profit coefficients) are precisely the ones that offer the highest potential to counterbalance the effects of local spillovers and natural advantages. This means that if wages have the highest profit coefficient (assumed negative) then a decrease of 1 per cent in the average cost of the workforce across regions will increase profits everywhere and will diminish the relative importance of localization economies and natural advantages, leading to a smaller level of "excessive concentration", more than an equivalent percentage change in any of the other factors affecting profits. But, on the other hand, we can see in the above expression for the elasticity of $\widetilde{\gamma}$, that the impact of any change is smaller for those industries that are more localized.

To gain some insight into the factors affecting localization we computed for each of the 3-digit SIC industries the elasticity of $\widetilde{\gamma}$ with respect to the explanatory variables introduced in our model. In Figure 3 we summarize the results of our calculations.[19] We find that wages (with a negative coefficient and thus capturing the cost of the workforce) have the highest elasticity for 15 industries (out of 92) while wages (with a positive coefficient and thus more likely to proxy the quality of the workforce) have the largest elasticity for 17 industries. Land costs and urbanization economies are the variables with the highest impact for 3 and 9 industries, respectively. On the other hand, large-scale accessibility has the largest (positive) impact for 38 industries in contrast with small-scale accessibility which is more relevant for only 3 industries. Thus, it seems fair to conclude that accessibility to the larger markets of the Porto-Lisbon coastal corridor (and transportation costs) is the factor with the highest potential to offset the influence that localization economies (and natural advantages of the regions) exert on firms' location decisions.

To reinforce this conclusion, we computed the impact on the average of $\widetilde{\gamma}$ across 3-digit SIC industries resulting from a 10 per cent decrease (across regions) for each one of the variables entering in the profit equation. Results are shown in Figure 4. Again, large-scale accessibility is the variable with the highest average impact. A 10 per cent decrease in this variable across regions (and thus a 10 per cent decrease in transportation costs from "concelhos" to the more urbanized coastal side of the country) results in a 13,73 per cent decrease on the average value of $\widetilde{\gamma}$ across 3-digit SIC industries, while the same elasticities for wages, land costs, urbanization economies and small-scale accessibility are 6,64, 2,05, 4,78, and 0,04 per cent, respectively.

(Figures 3 and 4, Pages 31 and 32)

[19] We restrict our analysis to those 92 sectors for which the regressions converged. Whenever the regression coefficients were not statistically significant we set the corresponding elasticities to zero.

5 Conclusion

Because it overcomes several significant pitfalls found in past measures, the index proposed in Ellison & Glaeser (1997) is now well established as the preferred method to measure localization of economic activity. In this paper we critically review the EG index, contending that the link between the Random Utility (Profit) Maximization framework and the Ellison and Glaeser measure is fragile and should be strengthened. We argue that the EG index treats the systematic forces that lead to spatial concentration as a black-box. In Ellison and Glaeser's view, in the absence of spillovers and natural advantages, all regions would exert the same pull on firms, regardless of sector of activity. Nevertheless, even in the absence of these non-systematic forces, one should still expect to find different patterns of location across industries, simply because firms from different sectors value regional factors differently.

Building on the McFadden's Random Utility (Profit) Maximization framework, we develop an alternative measure that is more consistent with the theoretical construct underlying the original work of Ellison & Glaeser (1997). With our approach, we are able to simultaneously compute the locational probabilities and the localization index. Hence, our method goes beyond the descriptive nature of the EG index and allows us to evaluate how the localization measure behaves with changes in the systematic forces that drive firms' location decisions.

References

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**(1), 14–22.

Bertinelli, L. & Decrop, J. (2002), Geographical agglomeration: The case of belgian manufacturing industry. Federal Planning Bureau of Belgium, Working Paper n.14.

Cabral, L. & Mata, J. (2003), 'On the evolution of the firm size distribution: Facts and theory', *American Economic Review* **93**(4), 1075–1090.

Callejón, M. (1997), 'Concentratión geográphica de la industria y economias de aglomeratión', *Economia Industrial* (5), 61–68.

Cameron, A. C. & Trivedi, P. K. (1998), *Regression Analysis of Count Data*, Econometric Society Monographs, Cambridge University Press, Cambridge, UK.

Carlton, D. (1983), 'The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables', *The Review of Economic and Statistics* **65**(3), 440–449.

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

Devereux, M. P., Griffith, R. & Simpson, H. (2004), 'The geographic distribution of production activity in the UK', *Regional Science and Urban Economics* . forthcoming.

Dumais, G., Ellison, G. & Glaeser, E. (2002), 'Geographic concentration as dynamic process', *The Review of Economics and Statistics* **84**(2), 193–204.

Duranton, G. & Overman, H. (2002), Testing for localisation using micro-geographic data. Centre for Economic Policy Research, Discussion Paper n. 3379.

Ellison, G. & Glaeser, E. (1997), 'Geographic concentration in u.s. manufacturing industries: A dartboard approach', *Journal of Political Economy* **105**, 889–927.

Figueiredo, O., Guimarães, P. & Woodward, D. (2002), 'Home field advantage: Location decisions of portuguese entrepreneurs', *Journal of Urban Economics* **52**(2), 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**(4), 403–418.

Guimarães, P., Figueiredo, O. & Woodward, D. (2000), 'Agglomeration and the location of foreign direct investment in Portugal', *Journal of Urban Economics* **47**(1), 115–135.

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

Hansen, E. (1987), 'Industrial location choice in Sao Paulo,Brazil: A nested logit model', *Regional Science and Urban Economics* **17**(1), 89–108.

Head, K., Ries, J. & Swenson, D. (1995), 'Agglomeration benefits and location choice: Evidence from japanese manufacturing investments in the United States', *Journal of International Economics* **38**(3-4), 223–247.

Holmes, T. & Stevens, J. (2002), 'Geographic concentration and establishment scale', *Review of Economics and Statistics* **84**, 682–690.

Hoover, E. (1937), *Location Theory and the Shoe and Leather Industries*, Harvard University Press, Cambridge, MA.

Houdebine, M. (1999), 'Concentration géographique des activités et spécialisation des départments français', *Economie et Statistique* **326-327**(6-7), 189–204.

Krugman, P. (1991), *Geography and Trade*, M.I.T. Press, Cambrige, MA.

Louviere, J., Hensher, D. & Swait, J. (2000), *Stated Choice Methods: Analysis and Application*, Cambridge University Press, Cambridge, UK.

Luger, M. & Shetty, S. (1985), 'Determinants of foreign plant start-ups in the United States: Lessons for policymakers in the Southeast', *Vanderbuilt Journal of Transnational Law* **18**, 223–245.

Mata, J., Portugal, P. & Guimarães, P. (1995), 'The survival of new-plants: Start-up conditions and post-entry evolution', *International Journal of Industrial Organization* **13**, 459–481.

Maurel, F. & Sedillot, B. (1999), 'A measure of the geographic concentration in french manufacturing industries', *Regional Science and Urban Economics* **29**, 575–604.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, *in* P. Zarembka, ed., 'Frontiers in Econometrics', Academic Press, NY, chapter 4, pp. 105–142.

Mosimann, J. (1962), 'On the compound multinomial distribution, the multivariate beta-distribution, and correlations among proportions', *Biometrika* **49**, 65–82.

Schmenner, R., Huber, J. & Cook, R. (1987), 'Geographic differences and the location of new manufacturing facilities', *Journal of Urban Economics* **21**(1), 83–104.

Woodward, D. (1992), 'Locational determinants of japanese manufacturing start-ups in the United States', *Southern Economic Journal* **58**(3), 690–708.

A Derivation of the EG index based on counts of plants

In the context of the EG model, the number of investments of a given industry in region j, conditional on the total number of investments in the industry, and on the vector of locational probabilities ($\mathbf{p} = p_1, p_2, ..., p_J$), follows a binomial law with parameters:

$$\begin{aligned} E(n_j/\mathbf{p}) &= np_j \\ V(n_j/\mathbf{p}) &= np_j(1-p_j) \end{aligned}$$

We now define a "raw index of concentration" as:

$$G_F = \sum_{j=1}^{J} \left(\frac{n_j}{n} - x_j\right)^2$$

Expanding terms we obtain,

$$G_F = \frac{1}{n^2}\sum_{j=1}^{J} n_j^2 + \sum_{j=1}^{J} {x_j}^2 - \frac{2}{n}\sum_{j=1}^{J} n_j x_j$$

and the expected value of the above equation gives:

$$\begin{aligned} E(G_F/\mathbf{p}) &= \frac{1}{n^2}E\left(\sum_{j=1}^{J} n_j^2\right) + \sum_{j=1}^{J} x_j^2 - \frac{2}{n}E\left(\sum_{j=1}^{J} n_j x_j\right) \\ &= \frac{1}{n^2}\left(\sum_{j=1}^{J} np_j - np_j^2 + n^2 p_j^2\right) + \sum_{j=1}^{J} {x_j}^2 - 2\sum_{j=1}^{J} p_j x_j \\ &= \frac{1}{n} + \frac{(n-1)}{n}\sum_{j=1}^{J} p_j^2 + \sum_{j=1}^{J} x_j^2 - 2\sum_{j=1}^{J} p_j x_j \end{aligned}$$

Applying the law of iterated expectations we get,

$$
\begin{aligned}
E(G_F) &= \frac{1}{n} + \frac{(n-1)}{n} E\left(\sum_{j=1}^{J} p_j^2\right) + \sum_{j=1}^{J} x_j^2 - 2E\left(\sum_{j=1}^{J} p_j x_j\right) \\
&= \frac{1}{n} + \frac{(n-1)}{n} \sum_{j=1}^{J} \left(\gamma x_j - \gamma x_j^2 + x_j^2\right) + \sum_{j=1}^{J} x_j^2 - 2\sum_{j=1}^{J} x_j^2 \\
&= \frac{1+\gamma(n-1)}{n} \left(1 - \sum_{j=1}^{J} x_j^2\right)
\end{aligned}
$$

and, as in Ellison & Glaeser (1997), the estimator for γ is obtained by replacing the $E(G_F)$ by the observed value of G_F and solving for γ. The proposed estimator is:

$$
\widehat{\gamma}_A = \frac{nG_F - \left(1 - \sum_{j=1}^{J} x_j^2\right)}{(n-1)\left(1 - \sum_{j=1}^{J} x_j^2\right)}.
$$

B A test of statistical significance for $\widehat{\gamma}_A$

Under the null hypothesis that $\gamma = 0$, the $p_j = x_j$ for all j and the observed spatial distribution of the investments for the particular industry follows a multinomial distribution,

$$
P(n_1, n_2, ..., n_J) = n! \prod_{j=1}^{J} \frac{x_j^{n_j}}{n_j!}
$$

Because we can associate a probability of occurrence to each possible distribution of the n investments we may also construct a distribution for the estimator of $\widehat{\gamma}_A$ under the null hypothesis that $\gamma = 0$. To do this, we may simply enumerate all possible values of the multinomial distribution. A simple example will help understand the argument. Suppose that we have 3 regions and 4 investments. Admit for the moment that $(x_1 = x_2 = x_3 = 1/3)$. The next table lists all possible spatial distributions of these investments, the associated probability, and the estimated concentration index ($\widehat{\gamma}_A$):

Table B.1 Distribution of Investments by Regions

n_1	n_2	n_3	$\widehat{\gamma}_A$	$P(n_1, n_2, n_3)$
4	0	0	1.00	1.23%
3	1	0	0.25	4.94%
3	0	1	0.25	4.94%
2	2	0	0.00	14.81%
2	1	1	-0.25	7.41%
2	0	2	0.00	14.81%
1	3	0	0.25	4.94%
1	2	1	-0.25	7.41%
1	1	2	-0.25	7.41%
1	0	3	0.25	4.94%
0	4	0	1.00	1.23%
0	3	1	0.25	4.94%
0	2	2	0.00	14.81%
0	1	3	0.25	4.94%
0	0	4	1.00	1.23%

This information can be used to construct the distribution for $\widehat{\gamma}_A$ which simply aggregates all common estimates and their probability. Thus, the distribution of $\widehat{\gamma}_A$ given $x_1 = x_2 = x_3 = 1/3$, $n = 4$, and $\gamma = 0$ is:

Table B.2: Statistical Distribution of the Estimator

$\widehat{\gamma}_A$	$f(\widehat{\gamma}_A)$	$F(\widehat{\gamma}_A)$
-0.25	44.44%	44.44%
0.00	22.22%	66.67%
0.25	29.63%	96.30%
1.00	3.70%	100.00%

From this simple example, we can see that if we had obtained an estimate of 1 for γ we could be fairly confident that $\gamma > 0$, given that the probability of that happening was only 3.7 per cent. But any other estimate would be a plausible outcome if the true value of γ were 0. Using this approach, we can test the probability that $\gamma = 0$ for any given number of investments and vector of locational probabilities.

However, it is not always feasible to construct the distribution of $\widehat{\gamma}_A$ by numerically evaluating all possible distributions of investments by regions (as we did in Table B.1). The number of terms that will need to be computed amounts to $\binom{n+J-1}{J-1}$. If, for example, $n = 20$ and $J = 10$, then we get 10,015,005 different cases. If n is increased to 40 we will have 2,054,455,634 different cases. In this case, instead of computing the exact distribution, we will randomly sample from this known distribution and generate an empirical cumulative distribution function for $\widehat{\gamma}_A$. Thus, in an application, we should test our hypothesis for each sector by generating a large number of draws (say 10,000) from a multinomial distribution with parameters $(n; x_1, x_2, ..., x_J)$. For each one of these samples, we will compute an estimate of γ and the value reported for our test will be the value of the empirical cumulative distribution evaluated at the observed value for $\widehat{\gamma}_A$.

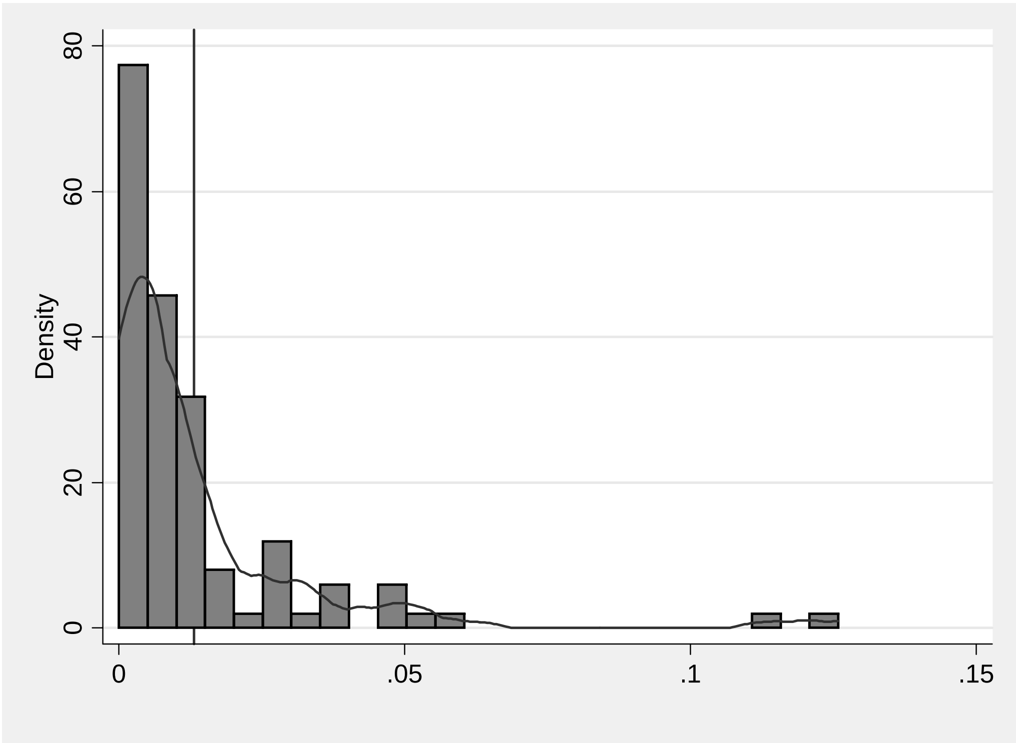

Figure 1: **Histogram of $\widetilde{\gamma}$ at the "concelho" level.**

Table 1: **Geographic Concentration, by Most Localized Industries According to $\widetilde{\gamma}$**

3-digit SIC Industry (Portuguese CAE-Rev2)	$\widetilde{\gamma}$	**Number of Plants**	**Rank** $\widetilde{\gamma}$	**Rank** $\widehat{\gamma}_A$	**Rank** $\widehat{\gamma}_{EG}$
354- Motorcycles and Bicycles	0.126	45	1	4	1
191- Leather Tanning and Finishing	0.115	110	2	1	5
362- Jewelry and Related Products	0.060	561	3	2	7
172- Broadwoven Fabric Mills	0.052	256	4	9	11
173- Dyeing and Finishing Textiles	0.048	275	5	16	21
193- Footwear	0.048	1932	6	7	15
171- Yarn Spinning Mills	0.048	226	7	17	22
351- Shipbuilding and Repairing	0.038	155	8	19	31
323- Radio and Television Apparatus (reception)	0.037	28	9	n.s.	3
176- Knit Fabric Mills	0.036	284	10	14	14
335- Watches, Clocks, and Clockwork Operated Devices	0.032	15	11	n.s.	n.s.
247- Artificial and Synthetic Fibers	0.029	12	12	n.s.	9
152- Sea Products Processing	0.029	106	13	22	23
192- Other Leather Products	0.029	244	14	23	28
341- Automobiles	0.027	15	15	n.s.	n.s.
177- Knit Article Mills	0.026	715	16	15	25
275- Ferrous and Nonferrous Foundries	0.025	154	17	28	48
313- Electric Cables and Related Products	0.021	31	18	n.s.	n.s.
205- Cork and Other Wood Products	0.020	1197	19	3	6
264- Brick, Roofing Clay Tile, and Related Products	0.019	197	20	31	38
262- Refractory and Non-refractory Ceramics	0.018	685	21	13	33
332- Measuring, Analyzing, and Controlling Instruments	0.018	29	22	24	n.s.

Note: *n.s.*- not significantly different from zero at 95% confidence.

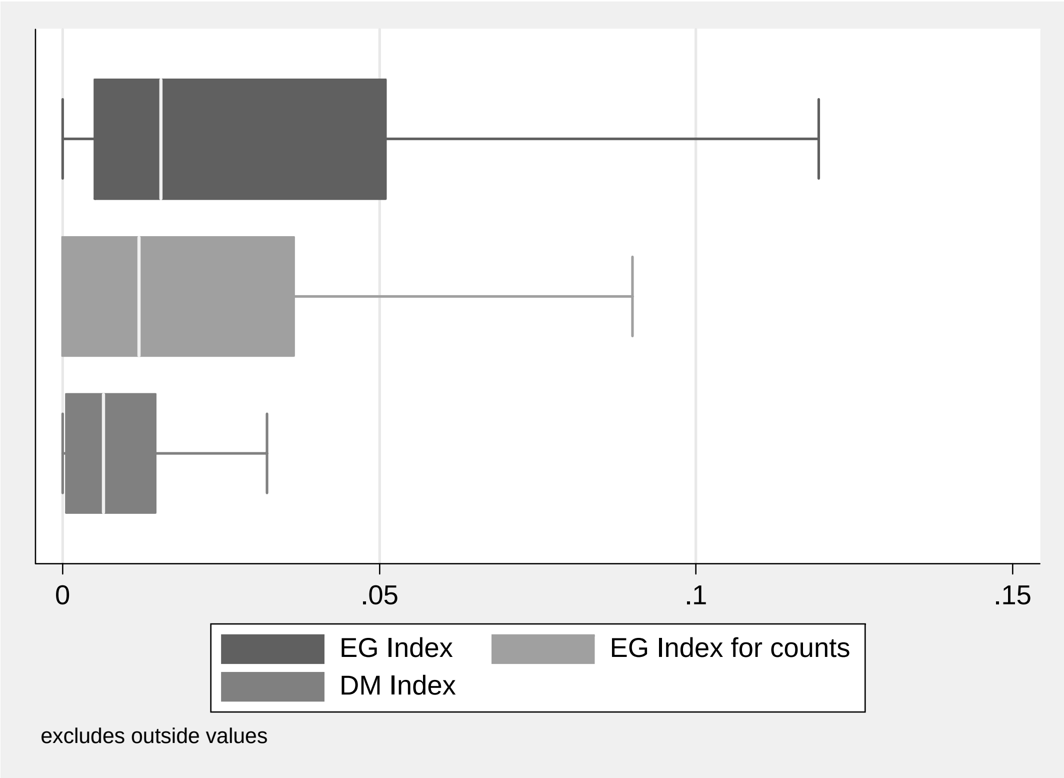

Figure 2: **Box-Whisker plots for the three localization indexes.**

Table 2: **Geographic Concentration, by Non Localized Industries According to $\widetilde{\gamma}$**

3-digit SIC Industry (Portuguese CAE-Rev2)	$\widetilde{\gamma}$	**Number of Plants**	**Rank**		
			$\widetilde{\gamma}$	$\widehat{\gamma}_A$	$\widehat{\gamma}_{EG}$
160- Tobacco	0.000	2	76	n.s.	n.s.
355- Miscellaneous Transportation Equipment	0.000	4	n.s.	n.s.	n.s.
296- Arms and Ammunition	0.048	7	n.s.	n.s.	n.s.
363- Musical Instruments	0.108	8	n.s.	5	4
283- Steam Generators	0.000	8	n.s.	n.s.	13
242- Agricultural Chemicals	0.000	10	76	n.s.	18
314- Electric Batteries and Related Products	0.002	11	n.s.	n.s.	n.s.
271- Primary Iron Industries	0.018	13	n.s.	n.s.	2
232- Petroleum Refining	0.004	13	n.s.	8	16
364- Sporting Goods	0.000	13	76	n.s.	n.s.
353- Aircraft and Space Vehicles	0.000	13	76	n.s.	n.s.
272- Iron and Steel Pipes and Tubes	0.000	15	76	n.s.	n.s.
333- Controlling Devices for Manufacturing	0.000	15	76	n.s.	n.s.
223- Gravure Printing	0.045	16	n.s.	27	n.s.
352- Railroad Equipment	0.000	20	n.s.	n.s.	n.s.
322- Radio and Television Apparatus (emission)	0.000	23	76	n.s.	n.s.
183- Fur Articles	0.009	27	n.s.	n.s.	n.s.
334- Optical, Photographic, and Cinematographic Instruments	0.011	28	n.s.	21	n.s.
268- Miscellaneous Nonmetallic Mineral Products	0.006	29	n.s.	n.s.	17
365- Games and Toys	0.003	29	n.s.	n.s.	41
273- Other Iron and Steel Primary Industries	0.003	30	n.s.	n.s.	n.s.
371- Recycling of Metal Products	0.003	37	n.s.	n.s.	n.s.
265- Cement and Related Products	0.006	56	n.s.	38	n.s.
263- Ceramic Wall and Floor Tile	0.000	57	76	30	30
311- Electrical Motors, Generators, and Transformers	0.004	83	n.s.	n.s.	n.s.

Note: *n.s.*-not significantly different from zero at 95% confidence.

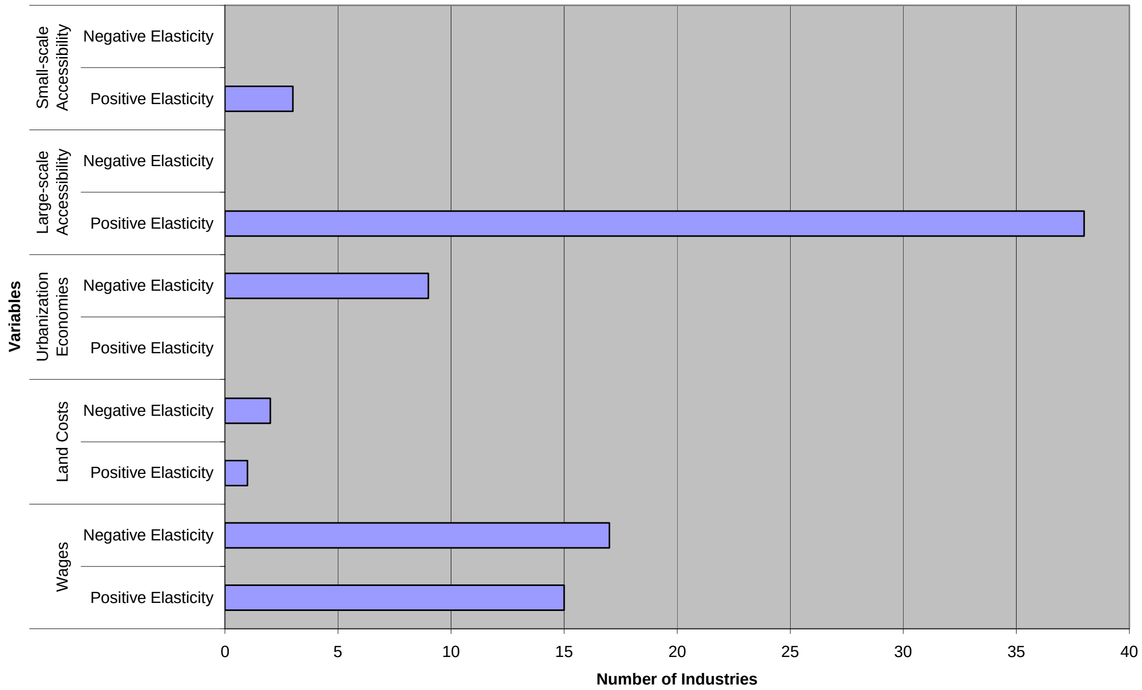

Figure 3: **Highest elasticities of $\widetilde{\gamma}$ by 3-digit SIC industries.**

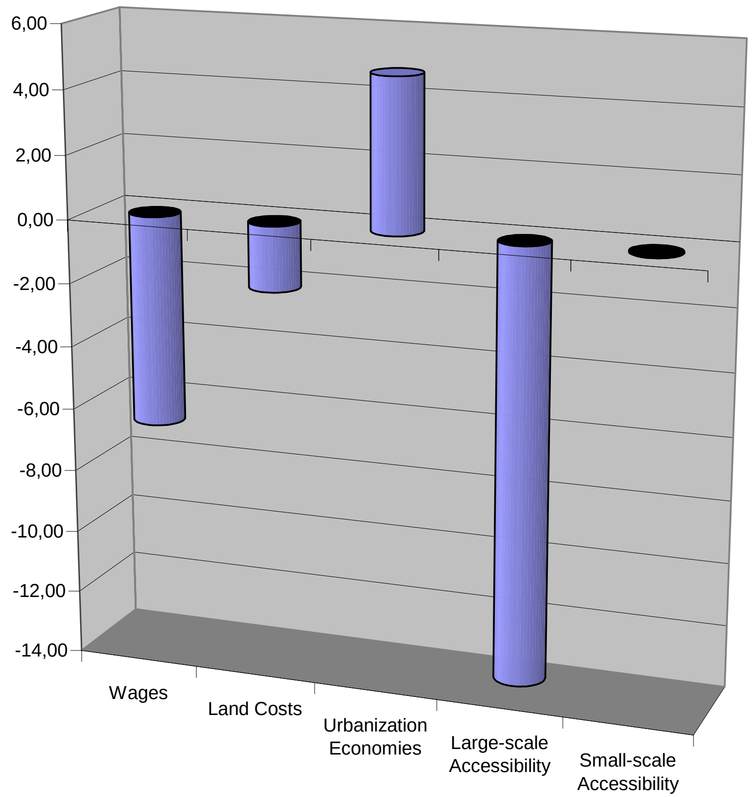

Figure 4: **Impact on the average of $\widetilde{\gamma}$ across 3-digit SIC industries.**