

de Almeida, Edilberto Tiago; Rocha, Roberta Moraes

## Article

# Labor pooling as an agglomeration factor: Evidence from the Brazilian Northeast in the 2002-2014 period

EconomiA

### Provided in Cooperation with:

The Brazilian Association of Postgraduate Programs in Economics (ANPEC), Rio de Janeiro

*Suggested Citation:* de Almeida, Edilberto Tiago; Rocha, Roberta Moraes (2018) : Labor pooling as an agglomeration factor: Evidence from the Brazilian Northeast in the 2002-2014 period, EconomiA, ISSN 1517-7580, Elsevier, Amsterdam, Vol. 19, Iss. 2, pp. 236-250, <https://doi.org/10.1016/j.econ.2018.02.002>

This Version is available at:

<https://hdl.handle.net/10419/266921>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

# Labor pooling as an agglomeration factor: Evidence from the Brazilian Northeast in the 2002–2014 period<sup>☆,☆☆</sup>

Edilberto Tiago de Almeida<sup>a,\*</sup>, Roberta de Moraes Rocha<sup>b</sup>

<sup>a</sup> Universidade Federal de Pernambuco, Centro de Ciências Sociais Aplicadas, Av. dos Economistas, s/n Cidade Universitária 50670-420, Recife, PE, Brazil

<sup>b</sup> Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, Rodovia BR 104, Km 62 – Nova Caruaru – 55002-970, Caruaru, PE, Brazil

Received 1 April 2017; received in revised form 9 January 2018; accepted 28 February 2018

Available online 7 March 2018

## Abstract

This is a pioneering study of Brazil of the importance of labor pooling to explain industrial agglomeration in the Northeast of Brazil, employing firm-level microdata. We applied the theoretical model proposed by Krugman (1991) (labor market pooling model) with the adaptations by Overman and Puga (2010) to examine how firms react to shocks in the labor market that influence their productivity. For this purpose, we applied regression models in which we regressed the Ellison and Glaeser (1997) index as a function of a proxy for labor pooling, to capture exogenous shocks in the labor market while controlling for observed sector characteristics that vary in time and sector fixed effects. The results are consistent with a reduction in the level of industrial concentration in the period from 2002 to 2014. For labor pooling, as predicted, the role of the labor pooling variable is positive and significant. Thus, industries where, on average, plants face more idiosyncratic shocks relative to their industry are more spatially concentrated.

*JEL classifications:* R11; R12

*Keywords:* Labor pooling; Agglomeration; Industries

© 2018 The Authors. Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The agglomeration of industrial activity can be observed in many countries, including Brazil (Sobrinho and Azzoni, 2016). Given that the distribution of industries can be a determining factor of the growth and development of regions, many studies have been carried out to shed light of the effects of industrial concentration. Among these are works to construct concentration indicators (e.g., Ellison and Glaeser, 1997; Maurel and Sédillot, 1999; Duranton and Overman,

<sup>☆</sup> Peer review under responsibility of National Association of Postgraduate Centers in Economics, ANPEC.

<sup>☆☆</sup> A preliminary version of this article is the third place of the CNI Economics Award 2016. The authors thank anonymous referees for suggestions that have greatly improved the paper. The authors thank Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE) for the research funding.

\* Corresponding author.

E-mail addresses: [edilberto.almeida@ufpe.br](mailto:edilberto.almeida@ufpe.br) (E.T. de Almeida), [robertamrocha@ufpe.br](mailto:robertamrocha@ufpe.br) (R. de Moraes Rocha).

2005) and empirical studies of the aspects underpinning the concentration of productive agents (e.g., Rosenthal and Strange, 2001; Resende and Wyllie, 2004; Ellison et al., 2010; Overman and Puga, 2010; Burki and Khan, 2013).

On the other hand, the literature has been less effective in distinguishing the sources of productivity gains derived from agglomeration (Overman and Puga, 2010). Although formal micro-founded theoretical models exist (see Krugman, 1991; Fujita and Thisse, 2013; Duranton and Puga, 2004), from an empirical standpoint the distinction among these different mechanisms is complex. This complexity, in turn, mainly results from all these theories predictive productivity gains with agglomeration but through mechanisms that are hard to trace (Duranton and Puga, 2004; Overman and Puga, 2010).

Among these mechanisms, labor pooling, initially highlighted by Marshall (2009), can act as a force of attraction, leading industries to concentrate. This centripetal force results from the advantages that can be obtained by firms that are concentrated in places where labor is abundant (Marshall, 2009). A formalization of this argument was proposed by Krugman (1991), called the “labor market pooling model”, in which an establishment adjusts its level of production and employment in response to idiosyncratic productivity shocks. This model suggests that establishments tend to concentrate spatially in places with a large number of workers having the skills necessary for their production. The firms can obtain advantages from this labor supply and adapt better to exogenous productivity shocks without significant declines in expected profits in function of adjustments in local wages. The model predicts that the effect of labor pooling will be as stronger as greater for the heterogeneity of the establishment-specific shocks in the sector.

To empirically test the hypothesis of the theoretical model in the United Kingdom, Overman and Puga (2010) regressed a measure of spatial concentration, the Ellison and Glaeser (1997) Index (EGI), as a function of a variable meant to capture the potential for labor pooling by sectors, with sectorial controls. The index used by the authors has the advantage of being weighted by the firm size, so that the problem of firm dimensionality does not influence the magnitude of the index. Another advantage is that the index is formally developed and supported by a macro-founded model. In turn, the labor pooling variable was proposed by Overman and Puga (2010) to capture the effect suggested by the theoretical model that the sectors in which idiosyncratic shocks are more heterogeneous across firms tend to be spatially more concentrated. In support of this, the authors argued that if wages are high when a firm intends to increase its workforce in response to a positive shock, compatible with the situation that most firms in a specific sector are also increasing their workforce, for example, this reduces its ability to adjust the wages paid. On the other hand, when changes in the firm’s headcount are not accompanied by significant wage increases, the advantage is greater to be located in regions with a large concentration of workers with the skills necessary for that firm’s production. Starting from this idea, according to Overman and Puga (2010), to isolate the direct<sup>1</sup> effect of heterogeneity, it is necessary to consider the difference between the percentage change of the establishment’s employment and the percentage change of employment in the industry (sector) in absolute value. This variable will take higher values when an establishment is expanding employment while the rest of the industry is contracting, and vice versa, capturing the heterogeneity of the shocks.

In Brazil, despite the availability of data disaggregated by geographic region and sector, few studies have analyzed the industrial concentration in the country with careful theoretical and empirical support. Besides this, the majority<sup>2</sup> of studies applying the EGI have done so for the entire country. The exceptions are the studies of Maciente (2008), who compared the indicators obtained for the country’s macro-regions, and Araujo (2016), who obtained evidence from this index for transformation industry clusters in the Northeast. Araujo (2016) did an exploratory analysis of the EGI for the Northeast and shows that there is a spatial concentration of the industry in the region, but he not explain the local factors that may explain its results.

This study advance to obtain evidences of the importance of labor pooling to explain the industrial concentration in the meso-regions Northeast,<sup>3</sup> using data from the Annual List of Social Information of the Ministry of Labor (RAIS-MTE) at the firm level for the period from 2002 to 2014. To try to explain these distorting effects, we use regression models considering the EGI – calculated for the Northeast, broken down by sectors of the manufacturing industry

<sup>1</sup> The authors argue that studies such as Rosenthal and Strange (2001) incorporated the pooling effect indirectly through labor skills and that this type of proxy might not be suitable to isolate the effect, capturing effects other than labor pooling.

<sup>2</sup> Biderman (2004), Resende and Wyllie (2005), Laurent and Araújo (2007).

<sup>3</sup> By considering the spatial distribution of industry within a macro-region, the heterogeneity of geographic sub-units diminishes in relation to the analysis of distinct macro-regions. In other words, more homogeneous sub-units are compared. The Brazil is officially divided in to five macro-regions: North, Northeast, Midwest, Southeast and South.

(divisions of the National Classification of Economic Activities – CNAE 1.0)<sup>4</sup> – as the dependent variable, and labor pooling as explanatory variable, controlling for others factors that can influence the industrial spatial concentration.

In the next section, we review the empirical evidence obtained for Brazil regarding industrial concentration patterns, mainly from studies that have applied the [Ellison and Glaeser \(1997\)](#) Index. In the third section we focus on the model of [Krugman \(1991\)](#) and the considerations of [Overman and Puga \(2010\)](#) about labor pooling. The fourth section describes the empirical strategy, while the last two sections present the results and final remarks, respectively.

## 2. Empirical evidence of geographic concentration of industry in Brazil

Empirical studies for Brazil can be divided into two groups according to the research objective. One group contains studies with the aim of measuring industrial agglomeration ([Biderman, 2004](#); [Resende and Wyllie, 2005](#); [Lautert and Araújo, 2007](#); [Maciente, 2008](#); [Araujo, 2016](#)) while in the other group the authors studies try to explain the concentration in light of local factors ([Resende and Wyllie, 2004](#); [Silveira, 2005](#); [Resende, 2015](#)).

[Biderman \(2004\)](#) presented evidence on concentration for all the divisions of the CNAE in micro-regions in Brazil for 1991 and 2000 based on the EGI, earlier studies were descriptive in nature.<sup>5</sup> Among the results, the average degree of concentration did not vary significantly between the two years studied.

[Resende and Wyllie \(2005\)](#) applied three concentration measures, those formulated by [Ellison and Glaeser \(1997\)](#), [Maurel and Sédillot \(1999\)](#) and [Devereux et al. \(2004\)](#), to the Brazilian case using data on the manufacturing industry for 1995 and 2001 obtained from the RAIS.<sup>6</sup> To calculate these measures, they considered the four-digit CNAE classification for sectorial disaggregation and the country's micro-regions for geographic disaggregation. The main results were the high heterogeneity across the sectors studied, with the EGI being the concentration metric that varied the least between 1995 and 2001. Besides this, from the 20 top classifications ranked according to agglomeration degree, 14 belonged to technology-intensive divisions of the CNAE (high and medium-high) according to the classification proposed by [Cavalcante \(2014\)](#),<sup>7</sup> suggesting greater concentration levels of this type of productive activity.

In the same context, [Lautert and Araújo \(2007\)](#) calculated the EGI for the manufacturing industry in Brazil in the period from 1996 to 2001 based on the CNAE (divisions/2 digits and groups/3 digits) with data from the RAIS at state level of geographic disaggregation. As in the previous study, the evidence suggested that technology-intensive sectors were more concentrated. In the ranking constructed by the authors, for both years, 6 of the 10 most concentrated sectors in Brazilian states were technology-intensive.

With focus on the dynamics of the concentration of employment by regions and sectors in Brazil, [Maciente \(2008\)](#) calculated the EGI values for the period from 1994 to 2005. He found evidence that sectors employing fewer workers tend to be more concentrated. These results are in agreement with those of other studies in the sense that technology-intensive sectors do not employ large numbers of workers, in particular in the Northeast. Thus, his study indirectly suggests the tendency for concentration of these sectors.

[Araujo \(2016\)](#) provides evidence about the concentration of the manufacturing industry in micro-regions and states in the Northeast based on the EGI. The author calculated the index for 2005 and 2010 with data from the RAIS. Also found that technology-intensive sectors are more concentrated than others in the country. Earlier studies for Brazil, with a more descriptive nature, such as [Pacheco \(1999\)](#), had already indicated this pattern. Thus, the Northeast follows the national trend in terms of concentration of technology-intensive firms.

In the second group of studies, the work of [Resende and Wyllie \(2004\)](#) added to the literature by using a measure of agglomeration formally developed by [Maurel and Sédillot \(1999\)](#), similar to the EGI, to analyze the industrial

<sup>4</sup> The sectors are identified in [Table A2](#) in [Appendix A](#).

<sup>5</sup> We can highlight, for example, the studies of [Diniz and Crocco \(1996\)](#), [Azzoni \(1997\)](#), [Pacheco \(1999\)](#), [Saboia \(2000\)](#) and [Suzigan et al. \(2001\)](#).

<sup>6</sup> The reason for using this database given by the authors was the limitations of the data available from the Brazilian Institute of Geography and Statistics (IBGE) through the Monthly Industrial Survey – Physical Production (PIM-PF), Monthly Industrial Survey – General Data Version (PIM-DG) and Annual Industrial Survey (PIA), in particular the advantage of the RAIS by supplying data disaggregated to the firm level, necessary to construct the indices used.

<sup>7</sup> [Cavalcante \(2014\)](#) proposes an association between the technological classification of the Organization for Economic Cooperation and Development (OECD) and the National Classification of Economic Activities (CNAE 1.0), which allows each CNAE division to be studied according to its technological level.

concentration in the state of Rio de Janeiro, using various databases<sup>8</sup> for different years. For this purpose they regressed the agglomeration metric on a series of sectorial and local characteristics. They adopted the strategy of [Ellison and Glaeser \(1997\)](#) to obtain the weight of each municipality in the sectorial indicator, so as to harmonize the variables used in the empirical model. The main results were that local infrastructure availability had a relatively important role in explaining agglomeration while incentive policies were less important.

[Silveira \(2005\)](#) studied the specialization and concentration of industrial activity in Brazil for two periods, 1950–1985 and 1985–2000. His empirical exercise consisted of two steps. In the first, using data from the Brazilian Institute of Geography and Statistics (IBGE – Industrial Census for the first period and Annual Industrial Survey for the second) for Brazilian macro-regions. The results showed a trajectory of industrial desconcentration from the 1970s, occurring more intensively until 1985 and, in the later period (1985–2000), in a more attenuated. He also presented an analysis disaggregated by industrial sectors where the capital-intensive sectors showed higher concentration based on the location coefficient of Hoover<sup>9</sup>. In second step of his study to explain industrial concentration. The empirical strategy relied on regression models with panel data for the 1950–1985 period and cross-sectional regression models for 1996 and 2000, where the dependent variable was the Hoover index and the regressors were proxies for scale economies and sectorial linkages. The results suggested the importance of these factors, but the effects of public policies cannot be disregarded.

Also in the second group of studies, [Resende \(2015\)](#) used an industrial co-agglomeration index and related it with labor pooling to seek evidence of the importance of this local factor to explain co-agglomeration. Based on [Ellison et al. \(2010\)](#), the author calculated the co-agglomeration index that relates pairs of industrial sectors using data on the manufacturing industry in the state of Rio de Janeiro in 2010. He quantified the industrial co-agglomerations and analyzed its relation with a proxy for labor pooling and other controls. Among the results, the author stressed the importance of labor pooling and sectorial characteristics (inputs) to explain co-agglomeration.

As shown above, most of the studies in Brazil using the micro-founded index of EGI have analyzed the country as a whole. Here we focus on the Northeast, to provide indicators of the concentration of the manufacturing industry, with emphasis on technology-intensive sectors, in its meso-regions. The results presented in the cited literature indicate that this type of industrial activity tends to be the most concentrated in Brazil. In this sense, our interest was to ascertain if this pattern holds for the Northeast based on the EGI at the regional level. Differently from the [Araujo study \(2016\)](#), which makes an exploratory analysis of the EGI, this research advances in the sense of explaining the concentration from the labor pooling.

In addition, we examine evidence about the importance of labor pooling to explain industrial concentration based on the model of [Krugman \(1991\)](#), following the empirical strategy used by [Overman and Puga \(2010\)](#), which allows us to test the theoretical hypothesis of that model by directly measuring the heterogeneity effect of sector-specific productivity shocks. This is our main contribution to the debate on the theme.

### 3. The model

According to [Duranton and Puga \(2004\)](#), the micro-fundamentals of urban agglomeration economies can be divided into three types, based on market linkages (sharing), labor market efficiency (matching) and learning mechanisms (learning). In this article, we investigate the attractive force generated by the concentration of workers in spaces, through the idea of labor pooling introduced by [Marshall \(2009\)](#): the concentration of specialized labor generates gains for firms and workers, such as by reducing the time of hiring and increasing the ability of firms to adjust to productivity shocks.

[Krugman \(1991\)](#) proposed a theoretical model to explain how this force acts on the distribution of productive activity. In the model,<sup>10</sup> industry is composed of a discrete number of firms that produce using a labor factor considered to be

<sup>8</sup> RAIS to construct the agglomeration metric and other information about industry; FIRJAN, CIRJ, SESI, SENAI and IEL for information about tax incentive policies; Fundação CIDE for information from the input–output matrix for the state of Rio de Janeiro and social and economic information for the state’s cities; BNDES for credit data of industrial sectors.

<sup>9</sup> The Hoover coefficient, concentration index analogous to the Gini Coefficient, is obtained from the construction of a Location Curve (analogous to the Lorenz Curve) (see [Hoover \(1936\)](#)).

<sup>10</sup> Based on [Krugman \(1991\)](#) and [Overman and Puga \(2010\)](#).

homogeneous for the activity in which the firm is specialized. The production function of each firm  $I$  can be described by the following equation:

$$y_i = (\beta + \varepsilon_i)l_i - \frac{1}{2}\psi l_i^2 \quad (1)$$

where  $\beta$  captures the additional change in output due to a marginal increment in the labor force ( $l_i$ );  $\varepsilon_i$  denotes the effect of the specific productivity shocks of the firm, with  $\varepsilon_i \sim (0, \sigma^2)$ , and  $\psi$  measures the decreasing returns of production.

Each firm pays the local wage, given by:

$$w = \beta + \varepsilon_i - \psi l_i \quad (2)$$

After a productivity shock, each firm hires more or fewer workers until equilibrium is established again. Assuming that the average wage in the locale depends on the number of workers, number of firms and productivity shock of each firm, the expected wage is given by:

$$E(w) = \beta - \psi \frac{L}{n} \quad (3)$$

considering that  $\varepsilon_i \sim (0, \sigma^2)$ .

From Eq. (3), note that the expected wage is positively related with the number of firms: as the number of firms in the location increases, the labor force/firm ratio declines, thus increasing the marginal productivity of this factor and consequently wages. On the other hand, it is inversely related to the number of workers, due to the reduction of productivity, and with the intensity of the decreasing returns ( $\psi$ ).

Considering the labor costs, the firm's profit can be denoted by:

$$\pi_i = y_i - w l_i = (\beta + \varepsilon_i)l_i - \frac{1}{2}\psi l_i^2 - (\beta + \varepsilon_i - \psi l_i)l_i = \frac{(\beta - w + \varepsilon_i)^2}{2\psi} \quad (4)$$

From the previous equation, profit is given by a convex function of the idiosyncratic shocks and wages. The expected profit is given by:

$$E(\pi_i) = \frac{[\beta - E(w)]^2 + \text{var}(\varepsilon_i - w)}{2\psi} \quad (5)$$

Substituting Eq. (3) and developing the term  $\text{var}(\varepsilon_i - w)$  in Eq. (5) yields:

$$E(\pi_i) = \frac{\psi}{2} \left( \frac{L}{n} \right)^2 + \frac{\text{var}(\varepsilon_i) + \text{var}(w) - 2\text{cov}(w, \varepsilon_i)}{2\psi} \quad (6)$$

with  $\text{var}(w) = \frac{\sigma^2}{n}$  and  $\text{cov}(w, \varepsilon_i) = \frac{\sigma^2}{n}$ .

From the preceding equation, in the absence of shocks the profit would be given by the first term on the right side of the equation. In this case, there would be a positive correlation between the labor/firm ratio and the expected profit, in function of the lower expected wage. The second term on the right side of Eq. (6) captures the labor pooling effect. "Each firm benefits from sharing its labor market with more firms in face of idiosyncratic shocks" (Duranton and Puga, 2004, p. 17).

The expected profit is positively correlated with  $\text{var}(\varepsilon_i)$  and  $\text{var}(w)$  because it is convex in relation to  $\varepsilon_i$  and  $w$ . On the other hand, it is negatively related with  $\text{cov}(\varepsilon_i, w)$ . In this last case, as pointed out by Overman and Puga (2010), if wages are higher when the firm intends to expand its production in response to a positive shock and lower when the firm intends to cut its production in response to a negative shock, the firm's ability to adjust will be limited, in the sense that the expected profit will decline due to the wage increases. This is the basic idea of the micro-foundation model of labor pooling as a source of agglomeration (Duranton and Puga, 2004; Overman and Puga, 2010). In other words, firms prefer places where the shocks are not strongly reflected in wages.

Using  $\text{var}(w) = \sigma^2/n$  and  $\text{cov}(\varepsilon_i, w) = \sigma^2/n$  and considering, besides these terms, the fact that  $\text{var}(\varepsilon_i) = \sigma^2$  in Eq. (6), we obtain:

$$E(\pi) = \frac{\psi}{2} \left( \frac{L}{n} \right)^2 + \left( 1 - \frac{1}{n} \right) \frac{\sigma^2}{2\psi} \quad (7)$$

The expected profits of the firms in the same place and sector are equal. Additionally, the effect of labor pooling measured by the second term on the right side of Eq. (7) is stronger when  $\sigma^2$  is larger. This implies that the benefits of pooling are greater when the heterogeneity of the shock in the sector is higher. Thus, sectors with more heterogeneous shocks should be more prone to concentrate.

Krugman's (1991) model suggests that establishments tend to concentrate spatially in places with a large number of workers possessing the skills needed for their production. Firms can obtain advantages from this supply of labor and better adapt to exogenous productivity shocks without significant reduction of their expected profit. As shown by the model, the effect of labor pooling will be greater the higher the heterogeneity of the shock in the specific sector is. Based on this theoretical hypothesis, according to Overman and Puga (2010), to isolate the effect of labor pooling, one takes the difference between the percentage change in the establishment's employment and the percentage change in the sector's employment in absolute value. This variable will assume higher values when an establishment is expanding its employment and the rest of the sector is reducing its employment, or vice versa, capturing the heterogeneity of the shocks. This is the empirical strategy followed here to build the model.

The potential labor pooling and the concentration measures used in this study, as described in the next section, were calculated from microdata from the RAIS, which contains data on individual firms, identified by taxpayer number. This research advances in relation to the supracite studies for Brazil when approaching the labor pooling from a theoretical model.

#### 4. Empirical model and database

The empirical model used here follows that developed by Overman and Puga (2010). Based on the theoretical arguments that labor pooling is important to explain the concentration of industrial activity, the expression to be estimated is represented by the following equation:

$$\gamma_{st} = \alpha + \psi P_{st} + \lambda L_{st} + \nu E_{st} + \delta_s + \varepsilon_{st} \quad (8)$$

The dependent variable,  $\gamma_{st}$ , is the EGI;  $s$  denotes the sectors of the manufacturing industries (15, ..., 37);  $t$  indexes the years ( $t=2003, \dots, 2014$ );  $P_{st}$  is a proxy for labor pooling, constructed as proposed by Overman and Puga (2010), consisting of the difference between the percentage variation of employment in a particular establishment and the percentage change in the sector's employment in absolute value;  $L_{st}$  and  $E_{st}$  are control variables (proxies for intersectorial dependencies and scale economies, respectively), which may be important in explaining the concentration incorporated in the model as described by Silveira (2005), who applied the method used by Kim (1995) to the Brazilian case;  $\delta_s$  is the sector fixed effect; and  $\varepsilon_{st}$  is the stochastic error term *i.i.d.* ( $0, \sigma^2$ ).

The EGI is represented by Eq. (9):

$$\gamma_s = \left\{ \left[ \sum_z (\delta_{zs} - x_z)^2 / \left( 1 - \sum_z x_z^2 \right) \right] - \sum_{k=1}^{N_s} \tau_{sk}^2 \right\} / \left( 1 - \sum_{k=1}^{N_s} \tau_{sk}^2 \right) = \frac{G_s - H_s}{1 - H_s} \quad (9)$$

The EGI is composed of two other indexes. One is the raw concentration index,  $G_s$ , which is defined by Eq. (10):

$$G_s = \sum_z (\delta_{zs} - x_z)^2 / \left( 1 - \sum_z x_z^2 \right) \quad (10)$$

where  $z$  stands for the geographic subunits;  $\delta_{zs}$  is the participation of  $z$  in the employment level of national (or regional, as here) sector  $s$ ; and  $x_z$  is the participation of  $z$  in the total industrial employment.

A limitation of  $G_s$ , as indicated by Ellison and Glaeser (1997), is its sensitivity with production scale of industries. Small number of firms with large production scale would cause overestimation of the indicator, indicating sectors with naturally higher production scales to be more concentrated.

Another measure that can serve as the base for constructing more refined concentration indexes is the Herfindahl–Hirshman index. It is a scale measure that attributes weights according to the market shares of each firm and is necessarily contained in an interval with lower bound given by  $1/n$  and upper bound of 1 (Resende and

Wyllie, 2005). Assume that industrial sector  $s$  has  $N_s$  firms, each with shares of employment in the sector given by  $\tau_{s1}, \dots, \tau_{sN_s}$ . The Herfindahl–Hirshman index is defined according to Eq. (11).

$$H_s = \sum_{k=1}^{N_s} \tau_{sk}^2 \quad (11)$$

Therefore, the EGI is a weighting of Eq. (10) by Eq. (11), which overcomes the limitation of another advantage of the EGI is that it is derived from a micro-founded location choice model developed by Ellison and Glaeser (1997).

Based in Overman and Puga (2010), the proxy for labor pooling was obtained according to Eq. (12):

$$P_t = |\Delta\%emp_{estab} - \Delta\%emp_s| \quad (12)$$

where  $\Delta\%emp_{estab}$  = percentage change of employment in the establishment; and  $\Delta\%emp_s$  = percentage change of employment in the sector (industry). The heterogeneity of the productivity shock is measured by the absolute value of the difference between the two variation metrics. In the form described, this indicator varies for each establishment, so to aggregate the variable by sector, we used the average for each industry.

The control for inter-sectorial dependencies was formulated according to expression (13).

$$L_{st} = \sum_{p=1}^9 C_{pst} / \sum_{p=1}^9 VTI_{pst}, \quad (13)$$

where  $C_{pst}$  = costs of raw materials, intermediate materials and components, indexed by place  $p$  of sector  $s$  in period  $t$ ; and  $VTI_{pst}$  = value of industrial transformation, with the same indexes.

Finally, the control for scale economies was constructed according to the equation below:

$$E_{st} = emp_{st} / nest_{st} \quad (14)$$

where  $emp_{st}$  = number of jobs in sector  $s$  in period  $t$ ; and  $nest_{st}$  = number of establishments in sector  $s$  in period  $t$ .

The main source of data was the RAIS-MTE, which contains data broken down by divisions of the CNAE and by meso-regions, and also micro-data, enabling identification of each firm according to its taxpayer enrollment number (National Register of Legal Entities - CNPJ), allowing following the yearly behavior. The use of this database enabled implementing the strategy proposed by Overman and Puga (2010) in the Brazilian case, more specifically the Northeast region, through the longitudinal data available for all the manufacturing firms in the region between 2002 and 2014 (defined in function of data availability). It should be mentioned that these data refer to the formal labor market.

We also considered the data from the Annual Industrial Survey (PIA) published by the IBGE, reflecting the characteristics of the sectors of the manufacturing industry. In this base, the data on the industrial transformation value (VTI) and costs of raw materials, intermediate materials and components up to 2007 are broken down according to the CNAE 1.0 and for subsequent years the CNAE 2.0. Therefore, in constructing the time series for these variables we relied on the table of correspondence between the CNAE 1.0 and CNAE 2.0 disclosed by the IBGE.

## 5. Analysis of the results

The results are presented in two steps. In the first we analyze the concentration patterns based on the EGI calculated for the Northeast and in the second we present the results of the empirical model.

### 5.1. Evidence of industrial concentration in the Northeast

We calculated the yearly EGI, as described in the preceding section, for the Northeast of Brazil for the period from 2002 to 2014, broken down by sectors of the manufacturing industry (CNAE 1.0 divisions). The geographic sub-units considered in the analysis were the 42 meso-regions<sup>11</sup> that compose the Northeast macro-region. The objective was to learn whether the firms of any of these industrial sectors are concentrated in these meso-regions. When the EGI is positive, this can indicate the existence of distortionary factors acting on the location choice of productive plants, leading

<sup>11</sup> Meso-regions in Brazil are official geographical subdivisions created by Brazilian Institute of Geography and Statistics (IBGE).



Table 1  
Descriptive statistics of the EGI for the Northeast.

Mean	Median	Minimum	Maximum	Standard deviation	Coefficient of variation	Obs.
All manufacturing sectors						
0.0548	0.0368	−0.0542	0.2901	0.0559	1.0207	299
Not technology intensive						
0.0378	0.0321	−0.0062	0.1851	0.0289	0.7633	195
Technology intensive						
0.0865	0.055	−0.0542	0.2901	0.0769	0.8883	104

Source: Prepared by the authors.

Table 2  
Descriptive statistics of the EGI for 2002 and 2014.

	All sectors		Non technology intensive		Technology intensive	
	2002	2014	2002	2014	2002	2014
Mean	0.0656	0.0368	0.0531	0.0236	0.0893	0.0616
Std. dev.	0.0532	0.0375	0.0438	0.0199	0.0638	0.0504
Minimum	0.0121	0.0035	0.0134	0.0035	0.0121	0.0075
Maximum	0.1937	0.1222	0.1851	0.0848	0.1937	0.1222

Source: Prepared by the authors.

to concentration. The index, however, does not distinguish these effects, which can be associated with natural factors of the meso-region and/or spillovers generated by firms that chose the location in question previously (economies of agglomeration).

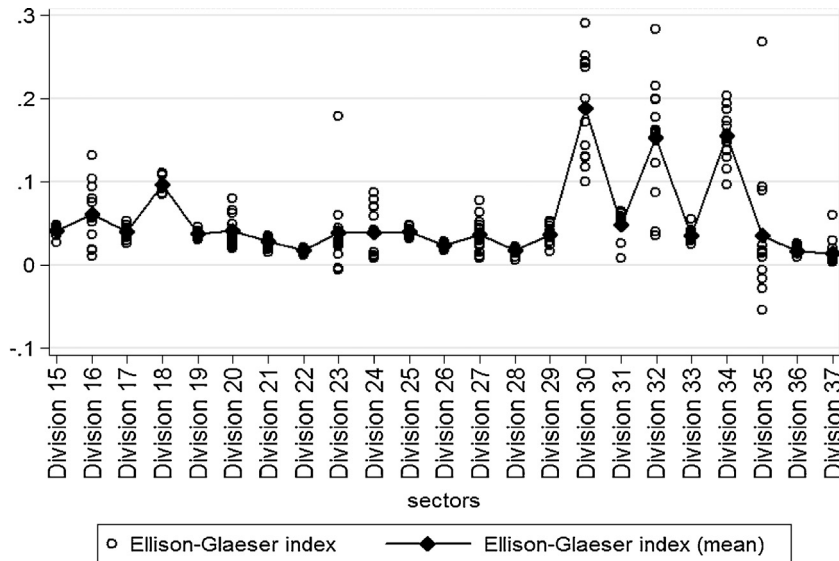
Besides this, given the existing empirical evidence on Brazil about the concentration of technology-intensive sectors, we sought to verify if these types of firms are more concentrated than others in the Northeast. In this respect, [Table 1](#) presents the descriptive statistics of the EGI considering the entire manufacturing industry and the sectors that are and are not intensive in technology using classification proposed by [Cavalcante \(2014\)](#).

For all years and sectors, the average EGI was 0.0548, indicating that location choice in the Northeast is not random. On average, the technology-intensive sectors had higher EGI than the other sectors. This can indicate the existence of factors that act more strongly on these sectors regarding geographic concentration.

We also observed a difference between the means for the first and last years considered in the sample, obtaining evidence of possible changes in concentration patterns with time. Considering all the sectors of the manufacturing industry, as shown in [Table 2](#), is it possible to identify a reduction of the average EGI between 2002 and 2014. The average value of this variable also declined for the technology-intensive sectors in the period, although this reduction was relatively lower compared to the average decrease observed in the other sectors. These results provide evidence that just as for the country as a whole, as found by [Resende and Wyllie \(2005\)](#) and [Lautert and Araújo, 2007](#), in the Northeast, technology-intensive sectors are more concentrated than the other manufacturing sectors. These results are also in accordance with those presented by [Araujo \(2016\)](#), who obtained empirical evidence for the Northeast and stated that these sectors tend to be more spatially concentrated in the region.

At the same time, these statistics for the concentration indicator suggest that the distortionary effects that influence where manufacturing plants are located declined during the period analyzed.<sup>12</sup> It is also important to mention that this analysis considered average values for all the sectors, but each sector has its own trend, which does not necessarily follow the general trend observed. These findings thus suggest the existence of a process of deconcentration of industry in the

<sup>12</sup> We also calculated the concentration measure for each sector for 2002 and 2014, shown in [Table A1](#) in [Appendix A](#). To check for any change in the concentration pattern in each sector, we ranked them according to the magnitude of the EGI in the two years. As can be seen in the column with the positions for 2002, five of the eight sectors considered to be sectors with high and medium-high technology intensity are among the ten first positions in 2002 and six are among the top ten positions in 2010.



Graph 1. Heterogeneity of the EGI among industrial sectors.

Source: Prepared by the authors.

Northeast, as pointed out earlier, possibly resulting from exogenous factors such as tax incentives and infrastructure improvements.

The data indicate the existence of significant heterogeneity of the sectorial concentration patterns in the meso-regions of the Northeast. To a certain extent this was expected, given that each sector has specific demands. This heterogeneity can be better observed in [Graph 1](#), which depicts the average variation of the concentration index of the sectors considering fixed temporal effects.

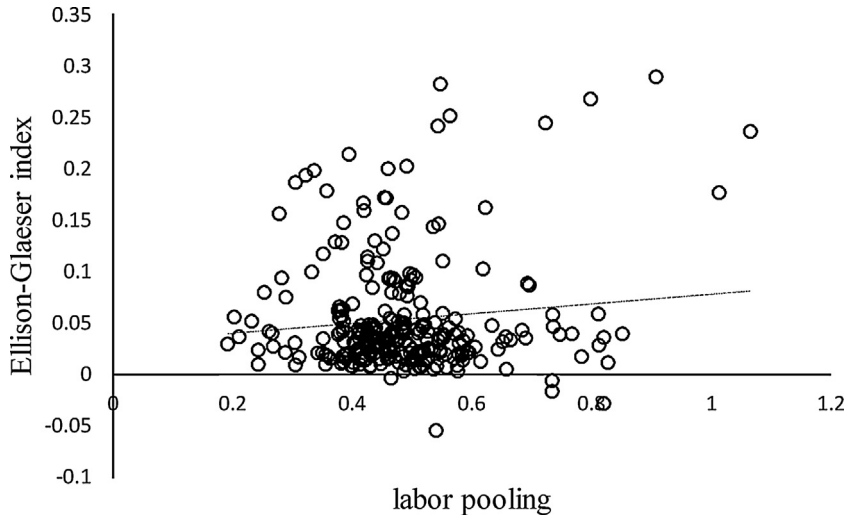
The analysis broken down by sectors again reveals higher values associated with technology-intensive sectors, such as the averages observed for Divisions 30 (0.1886), 32 (0.1532) and 34 (0.1519). For these, the variance around the means also is relatively higher. Division 36 was the industrial sector with the lowest mean (0.0169).

The empirical evidence found in this first step of the study suggests that in general, considering the entire manufacturing industry, the existence of a tendency for deconcentration among the meso-regions of the Northeast. In this respect, the results presented here are in accordance with those of [Silveira \(2005\)](#), [Resende and Wyllie \(2005\)](#) and [Lautert and Araújo, 2007](#) for Brazil as a whole, indicating that the concentration pattern observed nationally also holds regionally. The results also agree with the findings of [Araujo \(2016\)](#), who also analyzed the concentration in the Northeast. However, some of the technology-intensive sectors with growth potential, as pointed out by [Pacheco \(1999\)](#) for Brazil, presented a tendency for concentration, as is the case of Division 35, with EGI in 2002 equal to 0.012 and in 2014 equal to 0.094.

With respect to the heterogeneity among the sectors, [Resende and Wyllie \(2005\)](#) stated that since their study covered the entire country, thus reflecting its strong heterogeneity, the differences found were expected. In our study, with regional focus, we found much less pronounced heterogeneity. Nevertheless, it was still possible to observe divergent concentration patterns. These differences, in turn, might have been determined by the characteristics of each sector. This was the motivation to check for possible effects on concentration of factors intrinsic to the market. The next subsection contains a discussion and evidence from the empirical model described in the preceding section on this theme.

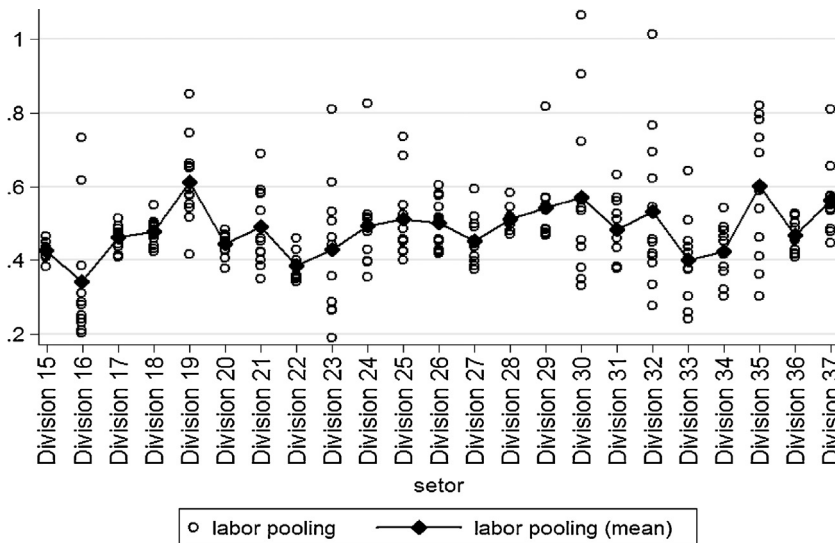
## 5.2. Results of the empirical model

To investigate the concentration of industry due to the distortionary effects generated by local factors, especially by the concentration of workers, we apply the empirical strategy suggested by [Overman and Puga \(2010\)](#) to construct the labor pooling variable. The idea that access to workers is an important factor influencing the concentration of industrial activity has existed since [Marshall \(2009\)](#). With the article by [Krugman \(1991\)](#), this argument gained formal



Graph 2. Ellison–Glaeser index versus labor pooling.

Source: Prepared by the authors.



Graph 3. Heterogeneity of labor pooling among the sectors.

Source: Prepared by the authors.

expression from labor pooling model, described in the Section 3. The theoretical hypothesis is that sectors with more heterogeneous productivity shocks will tend to concentrate spatially.

The dispersion of the EGI values in relation to labor pooling is shown in Graph 2.

The labor pooling variable was highly heterogeneous among sectors, as can be observed in Graph 3. Considering fixed time and the mean of the variable for each sector, it is possible to make inferences about the behavior of this variable among sectors.

In the way it was constructed, following the empirical strategy of Overman and Puga (2010), this variable can isolate the effect of the concentration of workers. It will be higher when a specific establishment is expanding employment (hiring workers) while the rest of the sector is reducing the number of workers, or vice versa.

The correlation coefficient between the EGI and labor pooling also suggests a positive relation between these variables. This provides empirical support for the theoretical hypothesis of Krugman’s model for the manufacturing industry in the Northeast, as can be seen in Table 3.

Table 3  
Correlation matrix of the variables of the empirical model.

	EGI	Labor pooling	Scale	Link
EGI	1			
Labor pooling	0.1536*	1		
Scale	−0.0424	−0.0492*	1	
Link	0.3726*	0.007	−0.2459*	1

\* Indicate the coefficients are significant at 5%.

Table 4  
Descriptive statistics of the variables of the empirical model.

Variable	Mean	Stand. dev.	Minimum	Maximum	Num. obs.
EGI	0.0539	0.0561	−0.0542	0.2901	276
Labor pooling	0.4841	0.1274	0.1909	1.0643	272 <sup>a</sup>
Scale	29.7344	48.2090	2.6817	377.7314	276
Link	1.1326	0.5897	0.1620	3.7365	235 <sup>b</sup>

<sup>a</sup> Outlier observations dropped from the sample.

<sup>b</sup> Information not available for the period after 2007.

Table 5  
Results of the empirical models.

Variable	Pooled OLS			Fixed effect	
	Model 1	Model 2	Model 3	Model 4	Model 5
Labor pooling	0.0481*** (0.0264)	0.0622** (0.0246)	0.0444*** (0.0241)	0.0537* (0.0191)	0.0535* (0.0201)
Scale		0.0001 (0.0001)	0.0001 (0.0001)	−0.0001 (0.0001)	−0.0001 (0.0001)
Link		0.0347* (0.0056)	0.0250* (0.0059)	0.0048 (0.0075)	−0.0008 (0.008)
Dum. technology			0.0310* (0.0074)		
Constant	0.0301** (0.0132)	−0.0192 (0.0145)	−0.0094 (0.0142)	0.0231 (0.0136)	0.0395** (0.0158)
Dum. time	No	No	No	No	Yes
R <sup>2</sup>	0.012	0.165	0.226	0.046	0.096
F	3.331	14.954	16.464	3.291	1.595
Num. obs.	272	231	231	231	231

Note: \*, \*\* and \*\*\* indicate the coefficients are significant at 1%, 5% and 10% respectively. Standard deviations of the estimates are between parentheses.

Additionally, we estimated the empirical model described in Section 4 considering data by sector and over time, to enable following the trajectory of a variable for a determined sector in the period from 2003 to 2014. As pointed out in the previous section, the EGI was calculated for the 2002–2014 period, but since the labor pooling variable was obtained from the variation of the level of employment for each establishment and each sector, the 2002 observation was lost. Table 4 presents the descriptive statistics of the variables included in the empirical model.

The results of the estimated regressions are reported in Table 5. Initially, disregarding the panel structure, the results of the pooled OLS estimation can be observed in columns 2, 3 and 4, which show the results of models 1, 2 and 3, respectively. In these three model specifications, the parameters were significant for the labor pooling variable.

In Model 2, which includes controls for scale economies and sectorial linkages, the labor pooling parameter was, albeit subtly, higher than that obtained in Model 1, in which the concentration indicator was regressed only on labor pooling. This suggests that the expected effect based on the theoretical model holds even when including other control variables for sectorial characteristics. In Model 3, we included a dummy variable for the technology-intensive sectors, which has value 1 for sectors considered intensive in technology (strongly or moderately) and 0 otherwise. The estimate of the parameter for this variable was significant at 1%, indicating that sectors that rely intensively on technology are more concentrated. This pattern was already indicated from analysis of the EGI, so the empirical model corroborates this argument. This analysis can be considered a complementary approach, one more formal from a statistical point of view.

In Model 4 and Model 5, who was adding time dummies, the parameters were obtained from fixed-effect estimators, meaning that control has been added to unobserved characteristics that are fixed in time. The parameter associated with labor pooling, on both models, was again statistically significant after these controls and as experienced with a magnitude infer to Model 2, when these controls are not considered. These results suggest that the labor pooling, originally highlighted by Marshall as an important factor of agglomeration economy location, is important to explain the concentration of the manufacturing industry in the Northeast's meso-regions, region characterized by labor-intensive industries. Regarding the magnitude of the parameter for labor pooling, in general, the estimates obtained in this research are smaller than those presented by [Overman and Puga \(2010\)](#) for the United Kingdom. This may suggest that in the Brazilian Northeast, although the concentration of workers influences the spatial distribution of firms, the importance of this local factor is apparently less. In order to determine in which of these two cases this type of externality could be stronger, a more detailed analysis would be necessary that considers a great part of the different characteristics between the two economies.

## 6. Final considerations

The aim of this study was to analyze the patterns of industrial concentration in Brazil's Northeast region and to investigate the importance of labor pooling to explain it. For that purpose, we quantified the concentration of the manufacturing industry through the EGI. This indicator provided evidence about the levels of industrial concentration as well as the dynamics of this process. The results obtained indicated there was a small reduction in the average of the index between 2002 and 2014, suggesting a trend for industrial deconcentration in the region. However when considering the results broken down by sectors (division of activities) of the manufacturing industry, some of these showed a stronger agglomeration trend, such as technology-intensive industries, which were on average more concentrated than the other sectors. These results are in accordance with previous studies cited in this study.

Based on the arguments of [Marshall \(2009\)](#) about labor pooling and the micro-founded formalization proposed by [Krugman \(1991\)](#) for this agglomeration factor, we conducted empirical tests to obtain evidence of the importance of this local factor to explain the industrial concentration in the Northeast. Following the empirical strategy of [Overman and Puga \(2010\)](#), it was possible to obtain evidence corroborating the theoretical model, according to which labor pooling is positively related to concentration, suggesting that industrial firms can benefit from the concentration of workers. This hypothesis was checked by applying two distinct estimators, and for both cases the results were statistically significant and in accordance with the expectation, even with the inclusion of controls for unobservable characteristics fixed in time and controls for economies of scale and sectorial linkages.

Suggestions for future research would be to expand the study to the entire country, and to consider the heterogeneity of labor qualification level, by incorporating in the model variables that capture the "quality" of workers' education and training, besides the "quantity" factor.

This study contributes to the empirical literature for Brazil by presenting evidence that geographic concentration of industrial firms provides productive advantages from the availability of specialized labor when compared to firms that are spatially isolated. This is an important result that can help guide regional policies aiming to make firms more productive, in particular policies to strengthen local productive arrangements. Given these findings, we recommend that the industrial policies seeking to increase productivity in the country should consider the potentials of regions as well as the productive vocation of local establishments and workers, bearing in mind that industrial concentration should not be considered harmful to the economy. On the contrary, market forces that generate productive externalities can spur the country's productivity.

## Appendix A.

Table A1  
EGI for the sectors of the manufacturing industry in the Northeast and concentration ranking.

Division	Description	2002		2014	
		γ <sub>s,2002</sub>	Position	γ <sub>s,2014</sub>	Position
<b>30</b>	Manufacture of office machines and informatics equipment	0.1937	1	0.0999	3
16	Manufacture of tobacco products	0.1851	2	0.0368	7
<b>32</b>	Manufacture of electronic goods and communication equipment	0.1541	3	0.1222	1
<b>34</b>	Manufacture and assembly of automotive vehicles and truck trailers	0.1194	4	0.1147	2
<b>24</b>	Manufacture of chemical products	0.0963	5	0.0098	18
20	Manufacture of wood products	0.0906	6	0.0203	14
18	Manufacture of clothing and accessories	0.0820	7	0.0848	5
37	Recycling	0.0758	8	0.0035	23
<b>31</b>	Manufacture of electrical machines, apparatuses and materials	0.0706	9	0.0075	21
27	Basic metallurgy	0.0624	10	0.0083	20
23	Manufacture of coke, refining of petroleum, preparation of nuclear fuels and production of alcohol	0.0462	11	0.0217	13
17	Manufacture of textile products	0.0450	12	0.0251	12
25	Manufacture of rubber articles and plastic material	0.0402	13	0.0337	9
15	Manufacture of food and beverage products	0.0372	14	0.0268	11
19	Preparation of leather and manufacture of leather goods, travel articles and footwear	0.0355	15	0.0342	8
<b>29</b>	Manufacture of machinery and equipment	0.0348	16	0.0394	6
<b>33</b>	Manufacture of medical-hospital instruments, precision and optical instruments, equipment for industrial automation, chronometers and clocks/watches	0.0331	17	0.0289	10
26	Manufacture of non-metallic mineral products	0.0245	18	0.0176	15
21	Manufacture of pulp, paper and paper products	0.0217	19	0.0149	16
36	Manufacture of furniture and sundry industries	0.0187	20	0.0089	19
22	Publishing, printing and reproduction of recordings	0.0178	21	0.0114	17
28	Manufacture of metal products, except machinery and equipment	0.0134	22	0.0060	22
<b>35</b>	Manufacture of other transport equipment	0.0121	23	0.0940	4

Source: Authors.

Note: The sectors (divisions) in boldface are intensive in technology according to [Cavalcante \(2014\)](#).

Table A2  
Technological intensity of the divisions (two-digit CNAE classification) that compose the manufacturing industry (CNAE 1.0).

Division	Description	OECD classification
15 ... 37	<b>Manufacturing industry</b>	
15	Manufacture of food and beverage products	Low
16	Manufacture of tobacco products	Low
17	Manufacture of textile products	Low
18	Manufacture of clothing and accessories	Low
19	Preparation of leather and manufacture of leather goods, travel articles and footwear	Low
20	Manufacture of wood products	Low
21	Manufacture of pulp, paper and paper products	Low

Table A2 (Continued)

Division	Description	OECD classification
22	Publishing, printing and reproduction of recordings	Low
23	Manufacture of coke, refining of petroleum, preparation of nuclear fuels and production of alcohol	Medium–low
24	<b>Manufacture of chemical products</b>	<b>Medium–high</b>
25	Manufacture of rubber articles and plastic material	Medium–low
26	Manufacture of non-metallic mineral products	Medium–low
27	Basic metallurgy	Medium–low
28	Manufacture of metal products, except machinery and equipment	Medium–low
29	<b>Manufacture of machinery and equipment</b>	<b>Medium–high</b>
30	<b>Manufacture of office machines and informatics equipment</b>	<b>High</b>
31	<b>Manufacture of electrical machines, apparatuses and materials</b>	<b>Medium–high</b>
32	<b>Manufacture of electronic goods and communication equipment</b>	<b>High</b>
33	<b>Manufacture of medical-hospital instruments, precision and optical instruments, equipment for industrial automation, chronometers and clocks/watches</b>	<b>High</b>
34	<b>Manufacture and assembly of automotive vehicles and truck trailers</b>	<b>Medium–high</b>
35	<b>Manufacture of other transport equipment</b>	<b>Medium–high</b>
36	Manufacture of furniture and sundry industries	Low
37	Recycling	Low

Source: Cavalcante (2014).

## References

- Araujo, J.E.S., 2016. *Novas evidências da distribuição espacial das indústrias do Nordeste brasileiro: 2005-2010*. Dissertação (Mestrado). Universidade Federal de Pernambuco, Programa de Pós-Graduação em Economia – PPGECON, Pernambuco.
- Azzoni, C.R., 1997. Concentração regional e dispersão das rendas per capita estaduais: análise a partir de séries históricas estaduais de PIB, 1939–1995. *Estudos econômicos* 27 (3), 341–393.
- Biderman, C., 2004. Concentração das atividades produtivas no Brasil: diferenças inter-setoriais, regionais e sua dinâmica. Instituto de Pesquisas Econômicas. Programa de Seminários Acadêmicos. Seminário n° 1/2004–FGV-EESP, São Paulo.
- Burki, A.A., Khan, M.A., 2013. Agglomeration economies and their effects on technical inefficiency of manufacturing firms: evidence from Pakistan. In: Unpublished Working Paper. International Growth Centre.
- Cavalcante, L.R., 2014. *Classificações tecnológicas: uma sistematização*. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Devereux, M.P., Griffith, R., Simpson, H., 2004. The geographic distribution of production activity in the UK. *Reg. Sci. Urban Econ. (Elsevier)* 34 (5), 533–564.
- Diniz, C.C., Crocco, M.A., et al., 1996. Reestruturação econômica e impacto regional: o novo mapa da indústria brasileira. *Nova Economia, Economics Department, Universidade Federal de Minas Gerais (Brazil)* 6 (1), 77–103.
- Duranton, G., Overman, H.G., 2005. Testing for localization using micro-geographic data. *Rev. Econ. Stud. (Oxford University Press)* 72 (4), 1077–1106.
- Duranton, G., Puga, D., 2004. Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, vol. 4. Elsevier, pp. 2063–2117.
- Ellison, G., Glaeser, E.L., 1997. Geographic concentration in us manufacturing industries: a dartboard approach. *J. Polit. Econ.* 105 (5).
- Ellison, G., Glaeser, E.L., Kerr, W.R., 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *The American economic review. Am. Econ. Assoc.* 100 (3), 1195–1213.
- Fujita, M., Thisse, J.-F., 2013. *Economics of Agglomeration: Cities, Industrial Location, and Globalization*. Cambridge University Press.
- Hoover, E.M., 1936. The measurement of industrial localization. *Rev. Econ. Stat.* XVIII, 162–171.
- Kim, S., 1995. Expansion of markets and the geographic distribution of economic activities: the trends in us regional manufacturing structure, 1860–1987. *Q. J. Econ. (JSTOR)*, 881–908.
- Krugman, P.R., 1991. *Geography and Trade*. MIT Press.
- Lautert, V., Araújo, N.C.M.D., 2007. Concentração industrial no Brasil no período 1996–2001: uma análise por meio do índice de Ellison e Glaeser. *Economia Aplicada (SciELO Brasil)* 11 (3), 347–368.
- Maciente, A.N., 2008. A concentração do emprego no Brasil entre 1994 e 2005: aspectos regionais e setoriais. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Marshall, A., 2009. *Principles of Economics: Unabridged Eighth Edition*. Cosimo, Inc.

- Maurel, F., Sédillot, B., 1999. A measure of the geographic concentration in French manufacturing industries. *Reg. Sci. Urban Econ. (Elsevier)* 29 (5), 575–604.
- Overman, H.G., Puga, D., 2010. Labor pooling as a source of agglomeration: an empirical investigation. *Agglomeration Economics*, vol. 2010. University of Chicago Press, pp. 133–150.
- Pacheco, C.A., 1999. Novos padrões de localização industrial? tendências recentes dos indicadores da produção e do investimento industrial. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Resende, M., 2015. Industrial coagglomeration: some state-level evidence for Brazil. *Nova Economia (SciELO Brasil)* 25 (1), 181–194.
- Resende, M., Wyllie, R., 2004. Ubicación industrial y políticas de incentivos locales en Brasil: una investigación empírica. *Perspectivas* 2, 9–24.
- Resende, M., Wyllie, R., 2005. Aglomeração industrial no Brasil: um estudo empírico. *Estudos Econômicos (São Paulo) (SciELO Brasil)* 35 (3), 433–460.
- Rosenthal, S.S., Strange, W.C., 2001. The determinants of agglomeration. *J. Urban Econ. (Elsevier)* 50 (2), 191–229.
- Saboia, J., 2000. Desconcentração industrial no Brasil nos anos 90: um enfoque regional. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Silveira, R., 2005. Concentração industrial regional, especialização geográfica e geografia econômica: evidências para o Brasil no período 1950–2000. *Revista Econômica do Nordeste (Fortaleza)* 36 (2), 189–208.
- Sobrinho, E.M.G., Azzoni, C.R., 2016. Aglomerações industriais relevantes do Brasil em 2010. *Revista Brasileira de Estudos Regionais e Urbanos* 9 (1), 1–18.
- Suzigan, W., et al., 2001. Aglomerações industriais no estado de São Paulo. *Economia Aplicada* 5 (4), 695–717.