

# The Location Quotient – Assembly and application of methodological enhancements

Strotebeck, Falk

5 July 2010

Online at https://mpra.ub.uni-muenchen.de/47988/ MPRA Paper No. 47988, posted 03 Jul 2013 11:00 UTC

### The Location Quotient -

# Assembly and application of methodological enhancements

Falk Strotebeck Faculty of Economics, Ruhr-University Bochum Universitätsstr. 150 44801 Bochum falk.strotebeck@rub.de

**Abstract:** The location quotient is an easy to use and often used indicator for identifying the clustering of industries, even though it struggles with some problems. This paper assembles different kinds of enhancements of the coefficient taken from the literature, which offer improvements regarding both the accurateness of the quotient and the interpretation of the results. This paper also combines two existing methodological enhancements in a new way and further applies the methods to analyze the localisation of the biotechnology industry in Germany, using data, taken from the BIOCOM Year- and Addressbook of 2005.

Keywords: location quotient, coefficient of localization, decomposition, cluster JEL classifications: R10; R12; C6

**Introduction:** By Porter's work (PORTER 1990) the analysis of the positive influence of regional agglomeration of industries has gained in importance again. <sup>1</sup> This not just concentrates on the output of the scientific community but also can be seen by the current innovation politics of the German government (e.g. cluster initiatives). Porter (1998) defined clusters as *"geographic concentrations of interconnected companies and institutions in a particular field."*<sup>2</sup> The geographic proximity between the relevant actors in this case increases the impact of the four factors Porter describes in his Diamond-Model. The four factors include the competition between actors, the specialisation on the local demand (and the pressure accompanied by that), the existence of suppliers and supporting actors and the availability of natural resources, specialised workers, capital and infrastructure.<sup>3</sup> Crucial points of this argumentation can even be traced back to VON THÜNEN (1826/1875) and MARSHALL (1890). E. g. VON THÜNEN explains the opportunity for big firms to generate economies of scale and considers the local demand (which must be high enough) as a crucial precondition. Secondly he mentions the positive effects of a

<sup>&</sup>lt;sup>1</sup> Vgl. Porter, M. E. (1990).

<sup>&</sup>lt;sup>2</sup> Porter, M. E. (1998), p. 78.

<sup>&</sup>lt;sup>3</sup> Vgl. Martin, R.; Sunley P. (2003), p. 8.

specialised labour market pool and the higher productivity of workers when they are able to do what they are best in without being distracted by additional tasks. Von THÜNEN also analyses the impact of local competitors regarding the setting of prices and he explains the positive impacts of local supporting firms when thinking about maintenance or need for spare parts and the transport of goods.<sup>4</sup> MARSHALL also describes central factors which happen to increase their impact when actors are located in geographical proximity. Mentioned are a specialised labour market pool, economies of scale either generated by big firms or by connected specialised and flexible small firms, the availability of input factors and infrastructure.<sup>5</sup>

The concept of clusters, as presented by Porter, is criticized by MARTIN AND SUNLEY (2003) because of the vagueness of the definition. Neither regarding the geographic proximity nor the relatedness of industries information is given to determine how to identify a cluster correctly.<sup>6</sup> Nevertheless a lot of measures are used throughout the literature to identify clusters.<sup>7</sup> None of these can overcome the *fuzziness* of the underlying concept but they indeed differ regarding pragmatism, explanatory power and theoretical foundation. Following FRATESI (2008) and DURANTON AND OVERMAN (2005) specific criterions regarding measures of agglomeration can be presented. First of all a measure which is meant to identify the agglomeration of an industry should correct for the distribution of the overall economic activity. Otherwise an agglomeration of an industry in a region might just be caused by the fact that there is a lot of economic activity in the region, but the industry does not really stick out in that case.<sup>8</sup> The effect of urbanization can be crossed out that way and the resulting effect shows the pure localization.<sup>9</sup> Secondly it does make sense to correct for the distribution of the firm size regarding the industry. <sup>10</sup> When (as it most often is the case) data on employment is used to measure localization, without correcting for firm size, one big firm can lead to a high value of the measure whereas one wants to identify the localization of a bunch of firms in one region. A very difficult problem to solve is the comparability across regions (FRATESI 2008). It is a problem, which is also known as the modifiable areal unit problem. The resulting measure should be independent from the size of

<sup>&</sup>lt;sup>4</sup> Vgl. Von Thünen, J. H. (1875), pp. 124-129, Fujita, M.; Thisse, J.-F. (2002), pp. 10-11, Reichelt, R. (2008), pp. 54-56.

<sup>&</sup>lt;sup>5</sup> Vgl. MARSHALL, A. (1895), pp. 348-353, FUJITA, M.; THISSE, J.-F. (2002), pp. 7-8, GORDON, I. R.; MCCANN, P. (2000), p. 516. Or as KRUGMAN (1991) puts it: Specialized labor market pool, industry specific inputs and technological spillovers. (KRUGMAN, P. 1991, pp. 36-38.

<sup>&</sup>lt;sup>6</sup> Vgl. Martin, R.; Sunley P. (2003), pp. 10-11.

<sup>&</sup>lt;sup>7</sup> As for example the Ellison/Glaeser Index (see Ellison, G.; GLAESER E. L. 1997; or also Alecke, B.; UNTIEDT, G. 2008 and FESER, E. J. 2000), Maurel/Sedillot Index (MAUREL, F.; SEDILLOT, B. 1999), the locational Gini Coefficient (KRUGMAN, P. 1991, p. 55; AUDRETSCH, D. B., FELDMAN, M. P. 1996, p. 633; SHELBURNE, R. C., BEDNARZIK, R. W. 1993, p. 11; BRÜLHART, M., TRAGER, R. 2005, p. 600; AMITI, M. 1997, pp. 24 – 25) or the Alpha-Index (DEVEREUX, M. P.; GRIFFITH, R.; SIMPSON, H. 1999)

<sup>&</sup>lt;sup>8</sup> Vgl. Fratesi, U. (2008), p. 735.

<sup>&</sup>lt;sup>9</sup> Vgl. Duranton, G.; Overman, H. G. (2005), pp. 1077-1078.

<sup>&</sup>lt;sup>10</sup> Vgl. Fratesi, U. (2008), p. 735.

regions observed and the way the areas are demarcated. <sup>11</sup> The result of a measure should at best also be comparable between industries, which postulates that it is independent from the size of the industry (for example measured by the employment) and the number of firms. An additional criterion can be the availability of a test of significance regarding the measurement results and last but not least the measure should be easy to use and (just one wish more) it should only rely on data which is easy to get one hands on. <sup>12</sup>

In this brief paper an easy to use and well known measure, the location quotient, is described, accompanied by some enhancements which help the quotient to overcome some of its shortcomings. The enhancements are then combined in a new way. Whereas the next section functions as an explaining part, the paper also contains a practical application of the presented methods for the German biotechnology industry.

# The Location Quotient and its methodological enhancements

The Location Quotient: The location quotient (LQ) can be used to see if the employment of an industry in a subregion is above or below the average. The average is given by the employment share of the industry regarding the overall employment in the nation. For example: The LQ can be calculated on federal state level to see how the employment share of the automotive industry, in relation to the overall employment of that federal state, differs to the employment share of the industry in relation to the overall employment on national level. In other words, the LQ therefore is the relation of the industry employment share regarding the overall employment in the analyzed subregion to the industry employment share regarding the overall employment in the nation. The LQ offers the opportunity to compare industries which differ in size because of the calculation using the employment shares on the national level as base magnitudes.<sup>13</sup> The LQ for a subregion *i* can be shown as follows:<sup>14</sup>

$$LQ_{i} = \frac{\frac{E_{i,r}}{E_{r}}}{\frac{E_{i,n}}{E_{n}}}$$

<sup>&</sup>lt;sup>11</sup> Vgl. FRATESI, U. (2008), p. 736. See regarding the MAUP OPENSHAW, S.; TAYLOR, P. J. (1979) and DARK, S. J., BRAM, D. (2007) and OPENSHAW, S. (1984).

<sup>&</sup>lt;sup>12</sup> Vgl. Fratesi, U. (2008), p. 736.

<sup>&</sup>lt;sup>13</sup> Vgl. Figueiredo, O.; Guimaraes, P; Woodward, D. (2007), pp. 4-5.

<sup>&</sup>lt;sup>14</sup> Vgl. Isserman, A. M. (1977), p. 34.

 $E_{i,r} = \text{Industry employment } i \text{ in } sub \text{region } r$   $E_r = \text{Overall employment in } sub \text{region } r$   $E_{i,n} = \text{Industry employment } i \text{ in region } n$   $E_n = \text{Overall employment in region } n$ 

A LQ above 1 (below 1) indicates an employment share of the observed industry above (below) the average.<sup>15</sup> Values above one are therefore often interpreted as indicators for a cluster because of the concentration of the employment.<sup>16</sup>

**Testing for significance:** One thing about the LQ which is often demanded is, that if the value of the LQ is above one, there is no indication regarding a significant deviation from zero. An enhancement to test for significance is given by MOINEDDIN, BEYENE AND BOYLE (2003).<sup>17</sup>

With

and

$$g_i = \frac{E_{i,r}}{E_r}$$

$$g = \frac{E_{i,n}}{E_n}$$

by using the approximated<sup>18</sup> variance  $V\left(\frac{g_i}{g}\right) = \frac{g_i(1-g_i)}{E_r g^2} + \frac{g_i^2(1-g)}{E_n g^3} - \frac{2g_i^2(1-g_i)}{E_n g^3}$  with

$$\frac{g_i}{g} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{g_i(1-g_i)}{E_r g^2} + \frac{g_i^2(1-g)}{E_n g^3} - \frac{2g_i^2(1-g_i)}{E_n g^3}}$$
 the  $(1-\alpha)$  confidence interval for the LQ of

each observed subregion is given.<sup>19</sup> And only the coefficients of localisation which show confidence intervals which are completely above 1 are interpreted as indicators for localised clusters. Given an  $\alpha$  of 5 percent a significant deviation from one could be confirmed with a probability of 95 percent. To use this enhancement of the LQ, the values of the coefficients have

<sup>&</sup>lt;sup>15</sup> Vgl. O'DONOGHUE, D.; GLEAVE, B. (2004), p. 422 und ISSERMAN, A. M. (1977), p. 34 and FIGUEIREDO, O.; GUIMARAES, P; WOODWARD, D. (2007), p. 4 and FINGLETON, B.; IGLIORI, D. C.; MOORE, B. (2004), p. 778.

<sup>&</sup>lt;sup>16</sup> Vgl. O'Donoghue, D.; Gleave, B. (2004), p. 422.

<sup>&</sup>lt;sup>17</sup> Vgl. Moineddin, R.; Beyene, J.; Boyle, E. (2003).

<sup>&</sup>lt;sup>18</sup> See regarding the approximation MOINEDDIN, R.; BEYENE, J.; BOYLE, E. (2003), S. 250-253 and OEHLERT, G. W. (1992), pp. 27-29.

<sup>&</sup>lt;sup>19</sup> Vgl. MOINEDDIN, R.; BEYENE, J.; BOYLE, E. (2003), pp. 250 and 252.

to be normally distributed.<sup>20</sup> A second possibility to add a control for statistical significance when using the LQ is proposed by O'DONOGHUE AND GLEAVE (2004) by using a standardized coefficient (SLQ).<sup>21</sup> Again the underlying assumption for detecting significance is a normal distribution of the values of the coefficient. The transformation of the LQ to the SLQ is done (optionally after logarithmic calculus) by a z-transformation. After the standardization procedure the deviation from the normal distribution indicates significant concentration.<sup>22</sup>

**Decomposing the LQ:** What one has to keep in mind is that a LQ above one can be reached by observing a high number of small firms in a subregion as well as by observing just a few very big firms. There is no control for the concentration on firm level. But HOLMES AND STEVENS (2002) introduced a method for decomposition of the LQ to overcome that obstacle. It is called the decomposition of the LQ. The LQ is deconstructed to its components and put together in a new way. What is needed indeed is a dataset which includes the information on the number of firms (firms per subregion). The composition of the LQ is as follows:<sup>23</sup>

$$Q_{i,r}^x \equiv Q_{i,r}^n \times Q_{i,r}^s$$

 $LQ = Q_{ir}^{x}$ 

with

where

$$Q_{i,r}^{x} = \frac{\frac{E_{i,r}}{E_{r}}}{\frac{E_{i,n}}{E_{n}}}$$
$$Q_{i,r}^{n} = \frac{\frac{n_{i,r}}{E_{r}}}{\frac{n_{i}}{E_{n}}}$$

and

$$Q_{i,r}^{s} = \frac{\frac{E_{i,r}}{n_{i,r}}}{\frac{E_{i,n}}{n_{i}}}$$

<sup>&</sup>lt;sup>20</sup> Vgl. Moineddin, R.; Beyene, J.; Boyle, E. (2003), p. 254.

<sup>&</sup>lt;sup>21</sup> Vgl. O'Donoghue, D.; Gleave, B. (2004), p. 422.

<sup>&</sup>lt;sup>22</sup> Vgl. O'DONOGHUE, D.; GLEAVE, B. (2004), pp. 422-423.

<sup>&</sup>lt;sup>23</sup> Vgl. Holmes, T. J.; Stevens, J. J. (2002), p. 683.

and

 $n_i$ = Number of firms of the industry *i* in total. $n_{i,r}$ = Number of firms of the industry *i* in subregion *r*.

The LQ is a product of  $Q_{i,r}^n$  (a plant location quotient) and  $Q_{i,r}^s$  (a size quotient). Therefore the logarithmic LQ (written as  $q_{i,r}^x$ ) is the sum of  $\log Q_{i,l}^n$ , in following written as  $(q_{i,r}^n)$  and  $\log Q_{i,l}^s$ , in the following written as  $(q_{i,r}^s)$ . In the end the deviation caused by a diverging number of firms and the deviation caused by the size of firms can be observed apart from each other. <sup>24</sup> This offers insight regarding the cause of specific LQ values. By using the decomposition it is possible to analyze if a high value of the LQ is caused by a very high number of firms located in that region or if it is caused by the size of the located firms. It is especially helpful to detect if the two effects have impacts regarding the LQ which work in the opposite direction.

The authors also propose a method to calculate the extent to which the both effects (firm size and number of firms) explain the variation of the LQ. By using  $\beta^s = \frac{\text{cov}(q_{i,r}^s, q_{i,r}^x)}{\text{var}(q_{i,r}^x)}$  and

 $\beta^{n} = \frac{\operatorname{cov}(q_{i,r}^{n}, q_{i,r}^{x})}{\operatorname{var}(q_{i,r}^{x})} \text{ (where } \beta^{s} + \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1 \text{) HOLMES UND STEVENS (2002) show the dependence of } \beta^{n} = 1$ 

 $q_{i,r}^{s}$  and  $q_{i,r}^{x}$ , and  $q_{i,r}^{n}$  and  $q_{i,r}^{x}$  respectively. If  $\beta^{s}$  equals zero the variation of the LQ depends totally on the number of firms ( $\beta^{n} = 1$ ) and is independent from the size of the firms and vice versa.<sup>25</sup>

**Identifying mass effects by horizontally clustering the LQ:** The LQ allows for a comparison of industries with different numbers regarding employment offering a relative result. What gets lost is information about the absolute size of an industry. The relative shares of the different industries might vary a lot in absolute terms however (thinking of automotive versus nanotechnology for example). The method of horizontal clustering proposed by FINGLETON, IGLIORI AND MOORE (2004) therefore takes into account the absolute industry size and allows to find out the number of how many employees lead to a LQ above (or below) the average. The LQ has to be changed in the following way:<sup>26</sup>

<sup>&</sup>lt;sup>24</sup> Vgl. Holmes, T. J.; Stevens, J. J. (2002), p. 683.

<sup>&</sup>lt;sup>25</sup> Vgl. Holmes, T. J.; Stevens, J. J. (2002), pp. 683-684.

<sup>&</sup>lt;sup>26</sup> Vgl. FINGLETON, B.; IGLIORI, D. C.; MOORE, B. (2004), p. 779 and POLENSKE, K. R. (2007), pp. 45-47.

Starting point is the LQ:

$$LQ_{i} = \frac{\frac{E_{i,r}}{E_{r}}}{\frac{E_{i,n}}{E_{n}}}$$

This is changed to:

$$LQ_{i} = \frac{\frac{E_{i,r}}{E_{i,n}}}{\frac{E_{r}}{E_{n}}}$$

At this point the  $E_{i,r}$  (the number of employees of industry i in subregion r) is replaced by  $\hat{E}_{i,r}$ .  $\hat{E}_{i,r}$  ist he expected value, assuming the subregion should have the employment share regardin the observed industry that can be observed on the national level. In this case the LQ would equal 1:

$$LQ_{i} = \frac{\frac{E_{i,r}}{E_{i,n}}}{\frac{E_{r}}{E_{n}}} = 1$$

The result from horizontal clustering is the difference between the observed value of  $E_{i,r}$  and the expected value of  $\hat{E}_{i,r}$  and therefore shows the in absolute terms the difference of the observed number of employees of an industry in a subregion and the expected number of employees in the subregion.<sup>27</sup>

**Combining methods:** Comparing the results of decomposing the LQ and horizontal clustering one can see that the latter does only take into account the employment effect and ignores the number of firms. The decomposition showed that the results of the LQ can be improved by considering the different effects of the firm size and the number of firms in a subregion simultaneously. In the following it is time to propose a combination of both methodological

<sup>&</sup>lt;sup>27</sup> Vgl. Fingleton, B.; Igliori, D. C.; Moore, B. (2004), p. 780.

enhancements of the LQ. Doing so means to use the horizontal clustering for the quotients regarding firm size and number of firms each:

Regarding

$$Q_{i,r}^{n} = \frac{\frac{n_{i,r}}{E_{r}}}{\frac{n_{i}}{E_{n}}}$$

 $n_{i,r}$  is replaced by the expected value  $\hat{n}_{i,r}$  :

$$Q_{i,r}^{n} = \frac{\frac{\hat{n}_{i,r}}{E_{r}}}{\frac{n_{i}}{E_{n}}} = 1$$

The difference between  $n_{i,r}$  and  $\hat{n}_{i,r}$  is the difference between the observed number of firms in the subregion and the expected number of firms considering the number on national level.

The same procedure can be done regarding the size of firms:

$$Q_{i,r}^{s} = \frac{\frac{E_{i,r}}{n_{i,r}}}{\frac{E_{i,n}}{n_{i}}}$$

Replacing  $E_{i,r}$  by  $\hat{E}_{i,r}$ :

$$Q_{i,r}^{s} = \frac{\frac{\hat{E}_{i,r}}{n_{i,r}}}{\frac{E_{i,n}}{n_{i}}} = 1$$

The difference between  $E_{i,r}$  and  $\hat{E}_{i,r}$  is the difference between the observed number of employment and the expected number of employment of the industry. By multiplying the average employment of a firm with the difference between  $n_{i,r}$  and  $\hat{n}_{i,r}$  and adding the result to the difference between  $E_{i,r}$  and  $\hat{E}_{i,r}$  the result would equal the result from horizontal clustering of the normal LQ.

By decomposing the LQ before – and therefore combining the methods of HOLMES UND STEVENS (2002) and FINGLETON, IGLIORI AND MOORE (2004) – the information from horizontal clustering can be improved. Whereas every deviation is explained by an absolute number regarding the employment, the method of applying the decomposition method beforehand provides information on the number of employees and the number of firms in absolute terms. Taking into account the distribution on national level one can see if a subregion suffers from an insufficient stock of firms or too small firms or what combination of both. Additionally the horizontal clustering provides information about a net total in terms of employment. The decomposition showed that firm size and the number of firms in a region might have an impact on the LQ that works in a different direction. The combination of the methods takes this into account and gives information about opposing trends in absolute terms. One can see that considering existing enhancements like methods for testing for statistical significance or including the firm size structure the LQ is an easy to use but informative measure.

#### Application of the collected enhancements and concluding remarks

In this last section the LQ and its enhancements are applied using data of the German biotechnology industry. The data that is used has been taken from the BIOCOM Year- and Addressbook 2005. The data (e.g. address, phone number, mail address, homepage, contact, date of foundation, number of employees) is given voluntarily by the firms, therefore sometimes data is missing. About 512 firms (regarded as the core firms of the biotechnology industry) are counted for the year 2005. 488 of these give information about their location and the number of employees. Altogether 17359 employees are counted. Data on the overall employment (26178266, regarding employees which are subject to social insurance contribution) for the German districts has been taken from the destatis online database (the German statistical office).

The following table shows the LQ values on federal state level.<sup>28</sup> Using the value of 1 as a threshold for indicating a cluster or not, eight of the sixteen federal states show a regional industry employment share above the average.

<sup>&</sup>lt;sup>28</sup> Ignoring the possible (and probable) bias caused by the modifiable areal unit problem. (Cf. OPENSHAW, 1984 and also see STROTEBECK, 2009 for the actual results regarding the biotechnology industry).

Federal state	LQ
Baden-Württemberg	0,897
Bavaria	1,421
Berlin	1,723
Brandenburg	1,262
Bremen	0,240
Hamburg	1,542
Hessen	1,334
Mecklenburg-Western	
Pomerania	1,717
Lower Saxony	0,711
North Rhine-Westfalia	0,731
Rhineland-Palatinate	0,317
Saarland	0,256
Saxony	0,407
Saxony-Anhalt	1,267
Schleswig-Holstein	1,844
Thuringia	0,370

#### Table 1. Location quotient on federal state level

Using the method proposed by MOINEDDIN, BEYENE UND BOYLE (2003) the values can be tested regarding their significance.<sup>29</sup> The following results can be displayed:

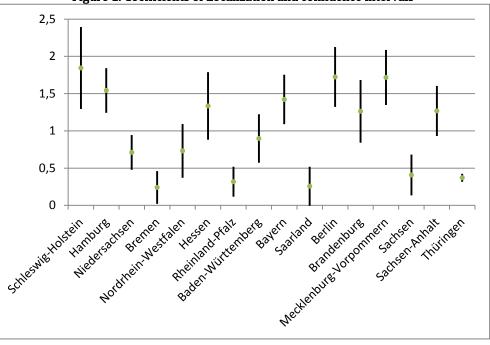


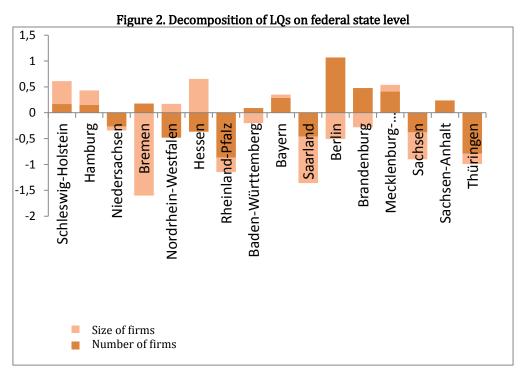
Figure 1. Coefficients of Localization and confidence intervals

As mentioned before, the regions with the whole confidence interval being above zero are of special interest. This is the case for five of the sixteen federal states (Schleswig-Holstein,

<sup>&</sup>lt;sup>29</sup> But it has to be mentioned that the Shapiro-Wilk test, which has been used to test the values about being normally distributed is undetermined with p = 0,098. The normal distribution is therefore not proved and the results are more for demonstrative purpose.

Hamburg, Bavaria, Berlin and Mecklenburg-Western Pomerania). The LQ values have been calculated in district level also. For the SLQ method of O'DONOGHUE AND GLEAVE the logarithmic and standardized values are used. On district level the values are normally distributed. Looking for significant values, only three districts (Ebersberg in Bavaria, Anhalt-Zerbst in Saxony-Anhalt and Greifswald in Mecklenburg-western Pomerania) show a significant localization of the biotechnology industry.<sup>30</sup>

In the next step the LQ gets decomposed, following Holmes and Stevens (2002). Looking at a figure, which displays the logarithmic values for the size and the plants quotients, it gets clear which effect works in what way towards the result measured by the LQ. Effects which foster agglomeration are displayed by the values above the zero baseline and effects which hinder agglomeration are displayed by values below the line. For ten of the sixteen federal states both quotients (plants and size) show in the same direction. But in six cases the direction of the effects regarding the number of plants and the size of them is different. In North Rhine-Westfalia for example the number of biotech firms has a negative impact on the value of the LQ, whereas the size of the firms has a positive impact. In Baden-Württemberg exactly the opposite can be seen. In both federal states the negative impact outweighs the positive one and in the result the LQ is below one. By using all the results of the federal states the explanatory power of both effects can be calculated. In this case size and number of firms explain about a half of the LQ value each ( $\beta^s = 0.5305$  and  $\beta^n = 0.4695$ ).



<sup>&</sup>lt;sup>30</sup> The used threshold is taken from O'DONOGHUE AND GLEAVE (2004). It is the deviation of 1.96 or more from the mean of the LQ values.

The core of the industry shows about 17000 employees. The method of horizontal clustering (FINGLETON, IGLIORI UND MOORE, 2004) is now used to show deviations from the overall employment in absolute terms on federal state level (see column: Deviation regarding employment in total). Additionally the first two columns in the following table display the results one can get when combining the methods of horizontal clustering and decomposing.

	Deviation regarding the number of firms	Deviation regarding employment (existing firms)	Deviation regarding employment in total
Schleswig- Holstein	3	337	431
Hamburg	2	186	265
Lower Saxony	-10	-87	-442
Bremen	1	-170	-136
North Rhine- Westfalia	-40	422	-985
Hessen	-12	888	462
Rhineland- Palatinate	-12	-78	-520
Baden- Württemberg	7	-492	-253
Bavaria	26	253	1192
Saarland	-2	-84	-168
Berlin	36	-778	506
Brandenburg	8	-182	102
Mecklenburg- Western	5	70	237
Pomerania Saxony	-8	-245	-523
Saxony-Anhalt	4	0	128
Thuringia	-7	-39	-296

#### Table 2. Results after combining decomposition and horizontal clustering

Looking at Berlin as an example the better insight by using a combination of the methods becomes clear. The horizontal clustering showed that Berlin has about 500 employees more in the core section of the biotech industry regarding the average on national level. But this number is a bit misleading when looking at the composition of the biotech in Berlin. The decomposition method showed that the size of firms has a negative impact on the LQ value. This fact should be considered when starting the horizontal clustering. When doing so, one can see that the firms in Berlin are relatively small compared to the national average. The calculated deficit in this case would be about 780 employees less than expected. But Berlin has an over proportional amount of firms. In absolute terms about 36 firms more than expected. The interaction of these two effects lead to the result of about 500 employees more than expected. In North Rhine-Westfalia on the opposite the firms are relatively big sized which leads to a high expected value regarding the employment in the region. But the number of firms is much lower than expected and this effect outweighs the other effect by far. Looking at Bavaria as an example of a federal state where both effects show in the same direction one can say that the over proportional employment in the core of the biotech industry is a result of a lot and at the same time bigger sized firms. The federal states with a high LQ can be grouped. In Schleswig-Holstein and Hamburg the LQ is mainly based on the size of the firms. In Bavaria, Berlin and Mecklenburg-Western Pomerania the number of firms is crucial.

The LQ is an easy to use measure which has, in its core version, a few drawbacks. For example a correction of the firm size distribution is not implemented, as well as a test of statistical significance of the LQ values. But, and this is what this paper wanted to show, there are enhancements which help the LQ to comply many of the presented conditions. Testing for statistical significance can be provided by the methods proposed by MOINEDDIN, BEYENE AND BOYLE (2003) or O'DONOGHUE AND GLEAVE (2004). Controlling for firm size can be done by decomposing (HOLMES AND STEVENS, 2002) the LQ. Even more possibilities can be added by horizontal clustering (FINGLETON, IGLIORI AND MOORE, 2004) and the combination of the decomposition and the horizontal clustering simultaneously.

### References

- ALECKE, B.; UNTIEDT, G. (2008): Die räumliche Konzentration von Industrie und Dienstleistungen in Deutschland – Neue empirische Evidenz mit dem Ellison-Glaeser-Index, Jahrbuch für Regionalwissenschaft, Vol. 28, 2008, pp. 61-92.
- AMITI, M. (1997): Specialisation patterns in Europe, Discussion Paper No. 63, Centre for Economic Performance, September 1997.
- AUDRETSCH, D. B.; FELDMAN, M. P. (1996): R&D Spillovers and the Geography of Innovation and Producation, The American Economic Review, Vol. 86, No. 3, Juni 1996, pp. 630-640.
- BIOCOM AG (2005): BioTechnologie Das Jahr- und Adressbuch 2005, Mietzsch, A. (Hrsg.), 19. Jahrgang, Berlin 2005.
- BRÜLHART, M.; TRAGER, R. (2005): An account of geographic concentration patterns in Europe, Regional Science and Urban Economics, Vol. 35, 2005, S. 597-624.
- DARK, S. J., BRAM, D. (2007): The modifiable unit problem (MAUP) in physical geography, Progress in Physical Geography, Vol. 51, Nr. 5, 2007, pp. 471-479.
- DEVEREUX, M. P.; GRIFFITH, R.; SIMPSON, H. (1999): The Geographic Distribution of Production Activity in the UK, Institute for Fiscal Studies, Working Paper 26/99, London, September 1999.
- DURANTON, G.; OVERMAN, H. G. (2005): Testing for Localization Using Micro-Geographic Data, Review of Economic Studies, Vol. 72, 2005, pp. 1077-1106.
- ELLISON, G.; GLAESER, E. L. (1997): Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach, The Journal of Political Economy, Vol. 105, No. 5, Oktober 1997, pp. 889-927.
- FESER, E. J. (2000): On the Ellison-Glaeser geographic concentration index, Department of Urban& RegionalPlanning,2000,publishedontheinternet:http://www.urban.uiuc.edu/faculty/feser/PUBS/EG\_index.pdf, at 14. August 2008.
- FIGUEIREDO, O.; GUIMARAES, P; WOODWARD, D. (2007): Localization Economies and Establishment Scale: A Dartboard Approach, FEP Working Papers, Nr. 247, September 2007.
- FINGLETON, B.; IGLIORI, D. C.; MOORE, B. (2004): Employment Growth of Small High-technology Firms and the Role of Horizontal Clustering: Evidence from Computing Services and R&D in Great Britain, 1991-2000, Urban Studies, Vol. 41, No. 4, April 2004, pp. 773-799.
- FRATESI, U. (2008): Issues in the measurement of localization, Environment and Planning A, Vol. 40, 2008, pp. 733-758.

FUJITA, M.; THISSE, J.-F. (2002): Economics of Agglomeration, University Press, Cambridge 2002.

- GORDON, I. R.; MCCANN, P. (2000): Industrial Clusters: Complexes, Agglomeration and/or Social Networks?, in: Urban Studies, Vol. 37, No. 3, pp. 513-532.
- HOLMES, T. J.; STEVENS, J. J. (2002): Geographic Concentration and Establishment Scale, The Review of Economics and Statistics, Vol. 84 (4), November 2002, pp. 682-690.
- ISSERMAN, A. M. (1977): The Location Quotient Approach to Estimating Regional Economic Impacts, Journal of the American Planning Association, 43:1, Januar 1977, pp. 33-41.

- KRUGMAN, P. (1991): Geography and Trade, Leuven University Press, Leuven, Belgium and The MIT Press, Cambridge, Massachusetts, London, England, 1991.
- MARSHALL, A. (1895): Principles of Economics, Vol. 1, Aufl. 3, Macmillan and Co., London und New York 1985.
- MARTIN, R.; SUNLEY P. (2003): Deconstructing clusters: chaotic concept or policy panacea?, Journal of Economic Geography, Vol. 3, 2003, pp. 5-35.
- MAUREL, F.; SEDILLOT, B. (1999): A measure of the geographic concentration in French manufacturing industries, Regional Science and Urban Economics, Vol. 29, 1999, pp. 575-604.
- MOINEDDIN, R.; BEYENE, J.; BOYLE, E. (2003): On the Location Quotient Confidence Interval, Geographical Analysis, Vol. 35, No. 3, Juli 2003, Ohio State University.
- O'DONOGHUE, D.; GLEAVE, B. (2004): A Note on Methods for Measuring Industrial Agglomeration, Regional Studies, Vol. 38. 4, Juni 2004, pp. 419-427.
- OEHLERT, G. W. (1992): A Note on the Delta Method, The American Statistician, Vol. 46, No. 1, pp. 27-29.
- OPENSHAW, S. (1984): The Modifiable Areal Unit Problem, Concepts and Techniques in Modern Geography, Nr. 38, Norwich, Geo Books, 1984.
- OPENSHAW, S.; TAYLOR, P. J. (1979): A million or so correlation coefficients: three experiments on the modifiable areal unit problem, in: Wrigley, N., editor, Statistical Applications in the Spatial Sciences, London: Pion, 1979, pp. 127-144.
- POLENSKE, K. R. (2007): The Economic Geography of Innovation, Cambridge University Press, Cambridge, 2007.
- PORTER, M. E. (1998): Clusters and the new economics of competition, in Harvard Business Review, Vol. 76, Issue 6, Nov/Dec 1998, pp. 77-90.
- PORTER, M. E. (1990): The competitive advantage of nations, free press, New York 1990.
- REICHELT, R. (2008): Deutschland und seine neue ökonomische Geographie, Diss., Freie Universität Berlin, Gutachter K. Jaeger und Zweitgutachter U. Baßeler, Berlin 2008.
- SHELBURNE, R. C.; BEDNARZIK, R. W. (1993): Geographic concentration of trade-sensitive employment, Monthly Labor Review, Vol. 116, pp. 3-13.
- STROTEBECK, F. (2009): Agglomerationsmessung und die Veränderung von Messwerten aufgrund des modifiable areal unit problems, GfR (Hrsg.), Seminarbericht 52, Heidelberg.
- VON THÜNEN, J. H. (1875): Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie, 3. Aufl. Zweiter Theil, II. Abteilung, Verlag von Wiegandt, Hempel & Paren, Berlin 1875.