Tendencies in the Romania’s Regional Economic Development during the Period 1991-2004

Tudorel Andrei and Andreea Iluzia Iacob and Liviu Bogdan Vlad

Academy of Economic Studies Bucharest

June 2007

Online at http://mpra.ub.uni-muenchen.de/6886/
MPRA Paper No. 6886, posted 26. January 2008 05:54 UTC
TENDENCIES IN THE ROMANIA’S REGIONAL ECONOMIC DEVELOPMENT DURING THE PERIOD 1991-2004

The objective of this paper represents the analysis of the way the Romania’s economic integration in the EU will influence the regional specialization and industrial activities localization within NUTS (the eight regions of Romania) during the period 1991-2004, using absolute measures (Herfindahl index).

Key-words: regional specialization, geographic concentration, panel data, fixed effect model, random effect model.

1. INTRODUCTION

More and more studies published in the specialized economic magazines investigate the characterization of the specialization degree of a country within an economic area or within the regions of the respective country. The Herfindahl and Krugman specialization indices, used mostly together with other statistical measurements, are primary measures for the descriptive analysis but especially variables within an econometric model.

The following can be listed in this category: studies that investigate the analysis of the specialization degree of West European countries \(^1\), \(^2\), on industry branches; ii) analyses that investigate the characterization of the specialization degree of Romanian regions on sub-branches of activity, etc.

2. DATA SERIES

The specialization degree of a country is decided upon a more comprehensive context, within a well defined economic area. Similarly, the analysis of the specialization of a county or region is achieved within the context of the national economy. Therefore, the following are taken into consideration for calculating the specialization and concentration indices:


- $R_1, R_2, ..., R_m$ are economic entities at whose level the specialization analysis is investigated. In a study they can be: the counties or the economic areas of Romania in a study that investigates the analysis of the economic specialization or concentration of industry on branches; ii) the countries of a certain economic region in the studies of economic geography, etc. The regions $R_1, R_2, ..., R_m$ define an economic area. In this study these will be the eight economic regions of Romania.

- $A_1, A_2, ..., A_n$ are the economic activities that take place within the economic area; they can be branches of the national economy, sub-branches of the industry, etc. In this study they will be the 13 industrial branches (according to NACE classification) of the manufacturing industry.

- The economic variable that helps quantify the volume of activity carried on in an economic region for each economic activity. Thus, $X_{ij}(t)$ = the volume of activity from the region $R_j$ achieved in the activity branch $A_i$ for a certain period of time. The volume of activity from a branch is quantified by the average number of employees in the respective field of activity, the employed persons, the achieved production, etc. This variable must provide a most accurate measurement of the volume of the activity carried on in the economic region, and the data series must be available. In this study, which has as objective the analysis of the specialization and concentration level of industry in Romania per development regions during the period 1991-2004, the variable used is the average number of employees in the 13 industrial branches (according to NACE classification) and the eight regions (NUTS II level) for Romania during the period 1991-2004.

- The one year data in the table below are used for calculating the necessary statistical indices for the characterization of the regional specialization and industry concentration:

<table>
<thead>
<tr>
<th>Industry $i$</th>
<th>Region $j$</th>
<th>Total industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{ij}(t)$</td>
<td>$X_j(t) = \sum_i X_{ij}(t)$</td>
<td>$X_i(t) = \sum_j X_{ij}(t)$</td>
</tr>
<tr>
<td>$X_j(t) = \sum_i X_{ij}(t)$</td>
<td>$X_i(t) = \sum_j X_{ij}(t)$</td>
<td>$X_i(t) = \sum_j X_{ij}(t)$</td>
</tr>
</tbody>
</table>

If the data series are provided for a certain period of time, then the above table multiplies by the number of years of the respective period of time.
3. PROCESSING OF DATA FOR THE CALCULATION OF THE REGIONAL SPECIALIZATION AND CONCENTRATION INDICES

Regional specialization is the characteristic of a region of economic development to specialize in a small number of industries. Regional concentration is the characteristic of an industry to concentrate in a small number of regions (Aiginger 1999(a)). The regional specialization and concentration can be measured using the average number of employees, the employed persons, the production (Gross Added Value) or data on export.

In order to measure the two concepts one can use different statistical measurements such as: entropy, Herfindahl index, Dissimilarity index proposed by Krugman and Gini index.

Annually, the following structures are estimated in order to calculate the above mentioned statistical measures:

i) The distribution of the annual activity volume in a region on industry branches. To this effect, one calculates for each region the distribution of the total volume of the characteristic \( X_j(t) \) on branches of industry according to the formula

\[
S^i_j(t) = \frac{X_i(t)}{X_j(t)}, \quad i = 1, n.
\]

We define for each region a column vector \( SS_j(t) \) defined by

\[
SS_j(t) = (S^i_j(t))_{i=1,n}.
\]

For a certain year we say that a certain region is "highly specialized" if its production is mainly the result of a small number of industry branches. For a certain year the matrix \( SS(t) = (SS_j(t))_{j=1,m} \) is defined. This matrix has the following important properties: on each row or column there is at least one figure other than 0; there are no negative figures in the matrix; the sum of the elements in each column equals 1; the sum of the elements in each row of this matrix doesn’t compulsory equal 1; the sum of the elements of this matrix equals the number of the industry branches, therefore 13 in this application.

ii) The distribution of the production volume of a certain industry on development regions. For each industry branch a row vector of size eight is defined, each element being calculated according to the formula

\[
S^i_j(t) = \frac{X_j(t)}{X_i(t)}, \quad j = 1, m.
\]

We mark this vector by \( SC_i(t) = (S^i_j(t))_{j=1,m} \) row vector. For a certain year we say that a certain industry is "tightly concentrated" if most of the production is achieved in a small number of regions. For the total industry branches the matrix \( SC(t) = (SC_i(t))_{i=1,n} \) is defined for each year. This matrix has the following important properties: on each row or column there is at least one figure other than 0; there are no negative figures in the matrix; the sum of the elements in each row equals 1; the sum of the elements in each column of this matrix doesn’t compulsory equal 1; the sum of the elements of this matrix equals the number of the economic regions, therefore 8 in this application.
iii) The distribution of the total annual volume of the industry on the \( n \) industry branches according to the formula
\[
s_i(t) = \frac{X_i(t)}{\sum X_j(t)}, \quad i = 1, n.
\]
The column vector \( \mathbf{PI}(t) = (s_i(t))_{i=1}^{n} \) is defined.

iv) The distribution of the total annual volume of the industry on developed regions according to the formula
\[
s_j(t) = \frac{X_j(t)}{\sum X_j(t)}, \quad j = 1, m.
\]
The row vector \( \mathbf{PR}(t) = (s_j(t))_{j=1}^{m} \) is defined.

4. HERFINDAHL INDEX

Specialization and concentration can be characterized by some absolute and relative statistical measurements. In the specialized literature several indices have been proposed such as: an absolute measurement (Herfindahl index) and a relative measurement (Dissimilarity index proposed by Krugman, Gini index, entropy) (Aiginger 1999). The calculated indices intend to define some conclusions for each year or for the whole period of time. This is the way the specialization or the concentration indices are calculated: for each year, for certain sub-periods of time or for the whole time. The following analysis of the specialization and concentration of Romanian industry during the period 1991-2004 uses the Herfindahl index.

4.1. General Presentation of Herfindahl Index

Specialization Herfindahl index (SPECH) is calculated for each region as a weighted arithmetical mean of the elements of the structure vector \( \mathbf{SS}'(t) = (s'_j(t))_{j=1}^{m} \) having as weights the elements of this vector:

\[
SPECH_j(t) = \mathbf{SS}'_j(t) \cdot \mathbf{SS}_j(t) \quad j = 1, ..., m
\]

(1)

In case there is no specialization at the regions level, then all the elements of the vector \( \mathbf{SS}'(t) \) are equal to: \( s'_j(t) = 1/n \), where \( n \) represents the number of branches. Under these conditions the index value equals \( 1/n \). If there is maximum specialization, then an element of the vector \( \mathbf{SS}_j(t) \) is equal to 1, and the rest of the elements equal zero. If we apply the above formula, the value of the specialization Herfindahl index equals 1. Generally, in the economic calculations, the index value is included between 0.1 and 0.7. For high values of the index, there is a high level of specialization in the region.

If the specialization index value multiplies by \( n \), then \( NSPECH_j(t) \) is obtained, with a value included between 1 and \( n \), and called the equivalent number of branches in a region.
Tendencies in the Romania’s Regional Economic Development During the Period 1991-2004

For each year, $SPECH_j(t) \quad j=1,...,m$, indices are calculated, where $m$ represents the number of regions in the country. The $SPECH(t) = (SPEC_j(t))_{j=1,m}$ vector of the specialization factors is defined. For this vector, using the elements of the vector $PR(t) = (s_j(t))_{j=1,m}$ as weights, one can calculate an average index of the specialization, the average quadratic deviation and the uniformity factor for the analysis of the variation degree of time specialization indices.

For example: The average specialization Herfindahl index is determined according to the formula:

$$SPECHM(t) = SPECH(t) \cdot PR'(t)$$

$$= \sum_{j=1}^{m} SPEC_j(t) \cdot s_j(t).$$

Similarly, Herfindahl index of concentration (CONCH) for each industry branch is calculated as weighted arithmetical mean of the elements of the structure vector, $SC_i(t) = (s_{ij}(t))_{j=1,m}$, having as weights the elements of this vector:

$$CONCH_i(t) = SC_i'(t) \cdot SC_i(t) \quad i = 1,...,n$$

In case there is no concentration of industry, then all the elements of the vector $SC(t)$ equal 1/m, and the index value equals 1/m. If there is maximum concentration of industry, then an element of this vector equal 1 and the others equal 0. Under these conditions the index value equals 1. The vector of the concentration Herfindahl indices $CONCH(t) = (CONCH_i(t))_{i=1,n}$, is defined. For this vector, using as weights the elements of the vector $PI'(t) = (s_j(t))_{j=1,m}$ one can calculate a concentration average index, the average quadratic deviation and the uniformity factor for the analysis of the variation degree of time concentration indices.

If the value of the concentration index multiplies by $m$, then $NCONCH_i(t)$ is obtained, with a value included between 1 and $m$. The new index is called the equivalent number of regions in which an industry is concentrated.

For example, the concentration Herfindahl average index is calculated according to the formula:

$$CONCHM(t) = CONCH(t) \cdot PI'(t)$$

$$= \sum_{i=1}^{n} CONCH_i(t) \cdot s_i(t).$$

We consider a case in which all the regions have the same economic dimension and the industry branches have equal dimensions. Under these conditions $X_j(t) = X_j(t)/n$, and $X_j(t) = X_j(t)/m$. Thus, the two specialization and concentration Herfindahl average indices are calculated according to the formulas below:
\[ SPECHM(t) = m \sum_{i,j} (X_{ij}(t) / X_i(t))^2 \]
\[ CONCHM(t) = n \sum_{i,j} (X_{ij}(t) / X_j(t))^2. \]  

(5)

The two indices may be rewritten in case the regions and industries are not of equal dimensions:
\[ SPECHM(t) = \sum_{i,j} (X_{ij}(t) / X_i(t))^2 / \sum_j (s_{ij}(t))^2 \]
\[ CONCHM(t) = \sum_{i,j} (X_{ij}(t) / X_j(t))^2 / \sum_i (s_{ij}(t))^2. \]  

(6)

The denominator of the first index is, in fact, a measurement of the regional specialization written as \( CONCRO(t) \), while with the second, the denominator is an index of the concentration of industries written as \( SPECRO(t) \). Out of the two formulas the result is that:
\[ SPECHM(t) \cdot CONCRO(t) = CONCHM(t) \cdot SPECRO(t). \]  

(7)

The distribution of the employed persons on sub-branches for each region has been used for the characterization of the degree of regional specialization and of industrial concentration. The regions taken into consideration in the study are listed in the following table. \( N_{ij} \) represents the employed persons in the \( j \) branch of industry of the economic region \( i \). The 1991-2004 specialization and concentration indices, as well as the equivalent number of branches or regions, are calculated using the above formulas and placing \( N_{ij} \) instead of \( X_{ij}(t) \). The results are listed in the table below.

**Table 1: Indices for the Characterization of Specialization and Concentration**

<table>
<thead>
<tr>
<th>Year</th>
<th>Indices of Specialization or Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( SPECHM )</td>
</tr>
<tr>
<td>1991</td>
<td>0.1270</td>
</tr>
<tr>
<td>1992</td>
<td>0.1225</td>
</tr>
<tr>
<td>1993</td>
<td>0.1212</td>
</tr>
<tr>
<td>1994</td>
<td>0.1189</td>
</tr>
<tr>
<td>1995</td>
<td>0.1182</td>
</tr>
<tr>
<td>1996</td>
<td>0.1186</td>
</tr>
<tr>
<td>1997</td>
<td>0.1140</td>
</tr>
<tr>
<td>1998</td>
<td>0.1181</td>
</tr>
<tr>
<td>1999</td>
<td>0.1214</td>
</tr>
<tr>
<td>2000</td>
<td>0.1240</td>
</tr>
<tr>
<td>2001</td>
<td>0.1294</td>
</tr>
<tr>
<td>2002</td>
<td>0.1296</td>
</tr>
<tr>
<td>2003</td>
<td>0.1280</td>
</tr>
<tr>
<td>2004</td>
<td>0.1261</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equivalent Number of Branches or Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( NSPECHM )</td>
</tr>
<tr>
<td>1.0159</td>
</tr>
<tr>
<td>0.9802</td>
</tr>
<tr>
<td>0.9693</td>
</tr>
</tbody>
</table>
Tendencies in the Romania’s Regional Economic Development During the Period 1991-2004

\[
\begin{array}{cccc}
0.9511 & 2.0111 & 1.0623 & 1.3691 \\
0.9459 & 2.0144 & 1.0569 & 1.3551 \\
0.9485 & 1.9931 & 1.0584 & 1.3657 \\
0.9116 & 1.9861 & 1.0569 & 1.3118 \\
0.9449 & 1.9806 & 1.0508 & 1.3711 \\
0.9715 & 1.9919 & 1.0526 & 1.3983 \\
0.9923 & 1.9704 & 1.0449 & 1.4447 \\
1.0351 & 1.9932 & 1.0468 & 1.5052 \\
1.0368 & 1.9893 & 1.0476 & 1.5107 \\
1.0242 & 1.9754 & 1.0410 & 1.4985 \\
1.0087 & 1.9598 & 1.0326 & 1.4762 \\
\end{array}
\]

*Data Source: Data Processing by INS*

4.2. The Analysis of the Acquired Results

For the interpretation of obtained results on specialization and concentration analysis, the starting point is the following general characteristics related to the employees number dynamics within the 13 branches during 1991-2004:

- Total number of employees decreased with 11696487 persons, accounting for a diminution by 53.2%; the most significant annual decrease was recorded in 1992, by 13.8% in relation with the figures for the previous year, being followed by the one of 1999, which accounted for 13.0%; significant decreases in employees number also occurred in the years 1993 (-5.8% as compared to previous year), 1994 (-6.4%), 1995 (-9.5%), 1998 (-6.1%), 2000 (-6.0%) and 2004 (-5.7%); for the whole period, insignificant increases were recorded only in 2001 (1.9% as against previous year) and 2002 (0.2%).

- At the level of the eight regions, the decreases in employees’ number are illustrated by the graph below. The most significant decreases were at the level of regions 1 (North-East Region) and 8 (Bucharest-Ilfov Region), each accounting for about 60%, while the lowest one was in region 3 (South-Muntenia Region), about 40%.
In order to observe the dependence between decreases in employees number by branch during 1991-2004 and the employees number of 1991, three linear regression models were defined:

- **Model I** where the percentage diminution in employees number by branch is a linear function of the branch weight in total industry in 1991;
- **Model II** where the percentage decrease in the employees number by branch is an exponential function of the branch weight in total industry in 1991;
- **Model III** where the absolute decrease in the employees number during this period is a linear function of employees’ number in each branch in 1991.

The estimates for the three models are presented in the table below:

<table>
<thead>
<tr>
<th>Table 2: Regression models for the analysis of decreases in the activity volume by branch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model I</strong></td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>Slope coefficient</td>
</tr>
<tr>
<td>F-statistic</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
</tbody>
</table>

* *, **, *** Statistically significant at the 1-percent level, at the 5-percent level, respectively at the 10-percent level.

The above presented data prove that the decreased in the employees number, both in terms of absolute and relative figures, was more important at the industrial branches level which were dominant in early 90s. The most significant decrease, accounting for 79.7% was recorded for branch I10 (Electric and Optical Equipment Branch) that held 18.28% of the total number of employees from industry. As for the branch I2 (Textile and textile products Branch), which held the highest weight within industry, accounting for 19.8%, the diminution was by 43.2%. The most insignificant decrease, by 14.2%, was recorded for branch I4 (Wood Processing (excluding furniture) Branch), which held, however, only 2.74% of the number of employees in industry in 1991. These results make proof of a poor adaptation of Romanian industry to the competitiveness of an open market.

**5. AN ECONOMETRIC MODEL FOR SPECIALIZATION AND CONCENTRATION ANALYSIS OVER TIME**

In order to follow up the concentration and specialization dynamics during the transition period, the following two regression models are taken into account, both for specializations and for concentration.

In terms of specializations:

**I. Common Intercept (Pooled OLS) Model:**
Tendencies in the Romania’s Regional Economic Development During the Period 1991-2004

\[
\log(SPECH_j(t)) = b + a \cdot t + \varepsilon_{jt} \tag{8}
\]

where \( j = 1, m, \) and \( t = 1991, 2004. \) \( R_{CC}^2 \) represents the determination coefficient of the common intercept (Pooled OLS) model.

II. Fixed Effect Model:

\[
\log(SPECH_j(t)) = b_j + a \cdot t + \varepsilon_{jt} \tag{9}
\]

\( j = 1, m, t = 1991, 2004, \) and \( b_j \) is estimated for each region. \( R_{EF}^2 \) represents the determination coefficient of the fixed effect model. In the case of this model, it is assumed that regions are different in relation with industrial specialization degree through the constant terms.

In order to decide upon the equivalence of the two models, a \( F \) test is to be done. It is thus checked if there are significant differences between the two models related to the constant term. In this sense, the null hypothesis of equal constant terms is tested for the second model:

\[ H_0: a_1 = a_2 = \ldots = a_m \]

In this case, the test statistics is defined:

\[
F = \frac{(R_{EF}^2 - R_{CC}^2)/(m-1)}{(1-R_{EF}^2)/(T \cdot m - m - 2)} \rightarrow F(m-1, T \cdot m - m - 2) \tag{10}
\]

If the calculated value of statistics is greater than the critical value, the null hypothesis is rejected. Under these conditions, it could be stated that the fixed effect model should be preferred to the one with common intercept.

III. Random Effect Model

If, in the case of the model above, it is considered that the constant could be split into a determinist component and a random one, where \( b_j = b + u_j, \) then the random effect model is defined through:

\[
\log(SPECH_j(t)) = b + a \cdot t + (u_j + \varepsilon_{jt}) \tag{11}
\]

\( j = 1, m, t = 1991, 2004. \) \( R_{EA}^2 \) represents the determination coefficient of the random effect model. In case of this model, it is assumed that regions are different in terms of industrial specialization degree by the random errors series.
In order to decide upon the equivalence of the two models, with fixed and random effects, the Hausman\(^3\) test is to be done. This test envisages the check, for the two models, of estimators’ efficiency and inconsistency. As for the next paragraph, \(\mathbf{a}_{EF}\) defines the estimators’ vector for the first model and \(\mathbf{a}_{RE}\) defines the estimators’ vector for the random effect model.

The two hypotheses of the test are defined as follows:

\[ H_0 : \text{the two estimators } \mathbf{a}_{CC} \text{ and } \mathbf{a}_{EF} \text{ are consistent, but } \mathbf{a}_{CC} \text{ is inefficient;} \]

\[ H_1 : \mathbf{a}_{CC} \text{ is consistent and efficient, but } \mathbf{a}_{EF} \text{ is inconsistent.} \]

The test statistics is defined based on the formula below:

\[
H = (\mathbf{a}_{EF} - \mathbf{a}_{CC})'(\text{var}(\mathbf{a}_{EF}) - \text{var}(\mathbf{a}_{CC}))^{-1}(\mathbf{a}_{EF} - \mathbf{a}_{CC}) \rightarrow \chi^2(2) \quad (12)
\]

If the statistics value exceeds the critical value, the null hypothesis is rejected, assuming that the second model is much more appropriate.

The same manner is applied for defining the three regression models for the variable quantifying the industry concentration level in relation with the employees’ number \(CONCH_i(t), \ i = 1, ..., n, t = 1991, 2004\). For these models, only the results of estimates are presented.

Within the table below, the characteristics of the three regression models defined for the variable quantifying the regional specialization are presented.

**Table 3: Regression characteristics for the specialization analysis by region**

<table>
<thead>
<tr>
<th></th>
<th>Pooled OLS</th>
<th>Fixed Effect Model</th>
<th>Random Effect Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.932526(^{*})</td>
<td>-</td>
<td>-0.932526(^{*})</td>
</tr>
<tr>
<td>Slope coefficient</td>
<td>0.002261(^{**})</td>
<td>0.002261(^{**})</td>
<td>0.002261(^{**})</td>
</tr>
<tr>
<td>(b_1)</td>
<td>-</td>
<td>-0.862331</td>
<td>0.065868</td>
</tr>
<tr>
<td>(b_2)</td>
<td>-</td>
<td>-0.876381</td>
<td>0.052684</td>
</tr>
<tr>
<td>(b_3)</td>
<td>-</td>
<td>-0.943424</td>
<td>-0.010226</td>
</tr>
<tr>
<td>(b_4)</td>
<td>-</td>
<td>-0.979146</td>
<td>-0.043746</td>
</tr>
<tr>
<td>(b_5)</td>
<td>-</td>
<td>-0.936210</td>
<td>-0.003457</td>
</tr>
<tr>
<td>(b_6)</td>
<td>-</td>
<td>-0.942096</td>
<td>-0.008980</td>
</tr>
<tr>
<td>(b_7)</td>
<td>-</td>
<td>-0.960310</td>
<td>-0.026071</td>
</tr>
<tr>
<td>(b_8)</td>
<td>-</td>
<td>-0.960310</td>
<td>-0.026071</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.26167</td>
<td>0.499337</td>
<td>0.470172</td>
</tr>
</tbody>
</table>

\(^3\) For a more detailed presentation of this test, we recommend: Hausman, J., *Specific Tests in Econometrics*, Econometrica, 46, 1251-1271, 1978.
In order to decide upon choosing one of the three models of regional specialization analysis, the two above mentioned tests are to be done:

a) For the first two models $F$ statistics is used, where statistics value is 13.77. From the table of $F$ statistics, for a significance level of 5%, 2.72 is determined. It is thus rejected the null hypothesis, accepting for the regional specialization analysis the fixed effect model.

b) In order to choose between fixed effect model and random effect model, Hausman test is to be done for calculating the statistics, case where 42579.11 is obtained.
As a consequence of applying the two tests, the most performing model for regional specialization analysis was found to be the random effects model. Under these conditions, the following conclusions could be worded:

i) At the beginning of the transition period, no significant specialization existed at regional level;

ii) During the analyzed period, a specialization effect existed only in regions 1 (North-East Region) and 2 (South-East Region). As for the other regions, a more or less significant decrease in this phenomenon was identified.

For the analysis of industry concentration by sub-branches, the parameters of the three regression models presented above were estimated, obtaining the results shown in Table 4. The data from the above table point out the inexistence of a rule for changing industry concentration during the transition period.

REFERENCES


Tendencies in the Romania’s Regional Economic Development During the Period 1991-2004