TESTING THE RELATIONSHIP BETWEEN THE EDUCATIONAL ATTAINMENT LEVELS OF THE POPULATION AND THE REGIONAL TYPOLOGY IN ROMANIA

BABUCEA ANA-GABRIELA,

PROF. PHD., “CONSTANTIN BRÂNCUŞI” UNIVERSITY OF TÂRGU JIU, ROMANIA
e-mail: gabibabucea@gmail.com

Abstract

Considering the obvious relationship between the education level of the population and level of development of the region of residence and its sustainable development was considered interesting to analyze whether the typology of the geographical regions influences the educational attainment level of the population. To highlight the association between geographical territory and educational attainment level of the population, the regions were classified into three different categories, respectively: predominantly urban, intermediate and predominantly rural by respect the Eurostat urban-rural typology, and the educational levels in three categories based on the ISCED11 levels and categories, respectively less than primary, primary and lower secondary education, upper secondary and post-secondary non-tertiary education, and tertiary education. The study aims to determine whether there is a significant relationship between the three categories of regions and the types of training, education predisposition to the resident population. In this sense, was appealed to Chi-Square test, a nonparametric test used to assess the association between two variables, both based on nominal scales, and specific measure of the degree of association. In the study were used centralized data at the national level referring to the population aged 25-64 by educational attainment level, and urban-rural typology, accessed from Eurostat database, for the year 2015, the latest year for which are available official data.

Keywords: Urban-rural typology, Educational attainment level, Nonparametric tests, Chi-square test of independence

Classification JEL: I24, R15

1. INTRODUCTION

The evolution of labor resources and of the active population directly determines the economic and social development of a country, but "education plays a key role in providing individuals with the knowledge, skills and competences needed to participate effectively in society and in the economy. In addition, education may improve people's lives in such areas as health, civic participation, political interest and happiness” and more than that, "having a good education greatly improves the likelihood of finding a job and earning enough money to have a good quality of life. Highly-educated individuals are less affected by unemployment trends, typically because educational attainment makes an individual more attractive in the workforce. Lifetime earnings also increase with each level of education attained”. [18]. Is a real fact that the job market for unqualified workforce shrinks, so more and more people realize that upper secondary education is no longer gives them an edge in their job search, and most young people incline to obtain other qualifications after graduating from high school. Educational attainment can be take into account as a main indicator of economic well-being [12, 23]. Thus, "the analysis of labour resources in rural, and of regional disparities between them could lead to identification of specific problems of social component for each region / county, on which must take actions to improve those levels and to establish priority objectives in the field” [19]. Romania faces after the year 1990 with regional disparities, both in administrative units, counties, but especially in terms of urban-rural regions. Marked by profound structural changes in terms of the distribution of the population at the local level under the impact of internal but especially external migration, Romania faces today, on the
one hand, with a strong downward trend of population, and on the other hand, especially a lack of qualified human resources needed to meet the challenges XXI century. Understanding their specific aspects could help avoid the negative effects these processes cause. Massive emigration in recent years of the population of working age, specialized, added to the population decline come to emphasize these issues, and put a great question mark over the economic future of Romania. Recent studies revealed that in Romania there are a considerable number of children who have never been to school, in time that the percentage of young people without relevant education for the labor market is extremely high. Inequalities arising from the regional profile decreased the interest for education and increased the school dropout rate, so many people are still inactive with minimal chance of entering the labor market [3, 18]. Burja and Burja consider that "early school leaving is an important problem faced by education in the rural areas. This aspect affects quality and competitiveness of the human capital; it generates poverty and social exclusion" [6]. In the last few years school dropout mainly affects the school population in the villages, while most of the higher education graduates live in urban areas (90%) [1]. There is a lot of determinants of an individual’s option to study different levels of education. The family tradition and the availability of chances are some of these main factors. If we discuss in terms of the territories, main factors that affect the education of the inhabitants include besides family support, the infrastructure and administration.

The residence area determines gaps in participation in education, especially in lower secondary education and post-obligatory education, affecting completely young population. These gaps are amplified by the certain "general conditions" that place the urban areas in a favorable position towards rural areas. Rural areas have still a "low degree of technical and municipal equipment, the poor and poor quality of the social and services infrastructure, poor development of the communication systems, poor transport network, especially road transport, severe poverty of some population groups". In addition, other certain "specific problems" of the educational system manifest in rural areas: "the inadequate condition of the buildings and the logistics under the conditions in which resources available for their rehabilitation are lower, or the insufficient number of qualified teachers" [13].

The economic and occupational disparities are highlighted both inter and intra-regionally. "Considering the important share of rural areas in the national territory, their prospects depend largely on the provision of services of general interest as the educational services are because there is an increase in the poverty rate and the exclusion of certain groups of rural inhabitants from certain services and utilities of general interest" [1]. To resolve these problems, Cirnu considers that "a responsible approach to the field should start from the quality of education and training of all the human resources involved, a real education process to represent a real long term investment for the success of regional development policy but also of other policies aimed at achieving economic and social cohesion [7]. Investing in education is a means to stop intergenerational transmission of disadvantage, if educational policies are sensitive to concentrations of disadvantage and integrated with social policies [9]. Therefore, important changes in economic policies in education, training programs for workforce retraining are necessary, contextual variables describing local and national policies and measures as well as special institutional features should be considered [8, 15, 11]. It needs to be correlate with the addressed population and their distinct types of regions.

The European Commission introduced territorial typologies based on population size, geographic location, and population density to monitor situations and trends in specific regions and areas. The Urban-rural typology classifies the NUTS 3 level regions of the EU into three different categories: predominantly urban, intermediate, and predominantly rural [5]. The importance of rural areas in the Romanian regional profiles in terms of the new European Union typology of 'predominantly rural', 'intermediate' or 'predominantly urban' regions, are evident when we discuss in term of territorial development.

According to European Commission, the aim of this new typology based on a variation of
the OECD methodology is to provide a consistent basis for the description of these three distinct types of regions in statistical analyses regarding aspects of quality of life. The methodology is based on the share of the rural population of NUTS level 3 regions (counties) and identifies three typologies of regions [24]:

- Predominantly urban regions, (PUR): with a rural population less than 20% of the total population.
- Intermediate regions, (INR): with a rural population between 20% and 50% of the total population.
- Predominantly rural regions, (PRR): with a rural population more than 50% of the total population.

The extremely small NUTS 3 regions, smaller than 500 km², are combined with one or more of their neighbors. The methodology considered the size of the cities in the region, too.

Therefore, a predominantly rural region becomes intermediate if it contains a city of more than 200000 inhabitants representing at least 25% of the regional population. In the same reason, an intermediate region becomes predominantly urban if it contains a city of more than 500000 inhabitants representing at least 25% of the regional population. The urban-rural typology of NUTS 3 regions is applying in Romania at the county level.

Romania has over 50% of the territory in predominantly rural areas, inhabited by almost 50% of the population. [17]. In the table below are presented the main indicators of Romania's urban-rural typology, at the year 2015. Note that almost 30% of the territory are intermediate regions, inhabited by over than 30% of population.

<table>
<thead>
<tr>
<th>Urban/rural typology</th>
<th>Territory</th>
<th>Total population</th>
<th>Total population aged 25-64</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Km²</td>
<td>%</td>
<td>Number</td>
</tr>
<tr>
<td>Predominantly urban regions</td>
<td>6.587.00</td>
<td>2.76</td>
<td>2.592.376</td>
</tr>
<tr>
<td>Intermediate regions</td>
<td>70.126.60</td>
<td>29.42</td>
<td>6.565.929</td>
</tr>
<tr>
<td>Predominantly rural regions</td>
<td>161.677.70</td>
<td>67.82</td>
<td>10.712.342</td>
</tr>
<tr>
<td>Total</td>
<td>238.391</td>
<td>100</td>
<td>19.870.647</td>
</tr>
</tbody>
</table>

Source: Calculated by the author based on data from EUROSTAT: Area of the regions by urban-rural typology [urt_d3area], Population on 1 January by five-year age group, sex and urban-rural typology [urt_pjangrp3] Extracted on 04.02.2017.

Based on the remark of Blanden, J., McNally even if "the inequality of educational achievements across the country can be measured in much the same way as income inequality, an approach more closely related to a concept of inequality of opportunity, is to consider inequalities between groups – for example, comparing achievements for children from different regions, socio-economic groups and migration status – because they show how education/economic outcomes are linked to characteristics of individuals, other than their effort or inherent abilities" [4].

Considering the obvious relationship between the education level of the population and level of development of the region of residence and its sustainable development was considered interesting to analyze the influences of the typology of the geographical regions the educational attainment level of the population in Romania, because of high levels of percentages of territories and population of predominantly rural and intermediate regions.

2. DATA OVERVIEW AND METHODOLOGY

Under the urban-rural typology Eurostat methodology the geographical territory, the 42 NUTS-3 regions of the Romania (counties), are classified into three different categories, respectively: predominantly urban, intermediate and predominantly rural.

The study aims to answer the question: Is there a statistically significant relationship between educational attainment and the urban-rural typology of the residence of the Romanian
population aged 25-65 years? In this regard were used centralized data at the national level referring to the population aged 25-64, as population able to be involved in the labor market, grouped by educational attainment level, and urban-rural typology.

The three considered categories classify the 9 ISCED levels (from 0 to 8) and categories of educations as in the table below.

Table no. 2. - Educational attainment level groups based on the classification of educational programmes by ISCED levels and categories

<table>
<thead>
<tr>
<th>Group</th>
<th>ISCED level</th>
<th>Category</th>
</tr>
</thead>
</table>
| Less than primary, primary and lower secondary education | 0 - Early childhood education ('less than primary' for educational attainment) | 0.1 early childhood development  
0.2 pre-primary education |
| | 1 - Primary education | |
| | 2 - Lower secondary education | 2.1 general education  
2.2 vocational education |
| Upper secondary and post-secondary non-tertiary education | 3 - Upper secondary education | 3.1 general education  
3.2 vocational education |
| | 4 - Post-secondary non-tertiary education | 4.1 general education  
4.2 vocational education |
| Tertiary education | 5 - Short cycle tertiary education | 5.1 general education  
5.2 vocational education |
| | 6 - Bachelor’s or equivalent level | 6.1 long first degree  
6.2 first degree  
6.3 further degree |
| | 7 - Master’s or equivalent level | 7.1 long first degree  
7.2 second degree  
7.3 further degree |
| | 8 - Doctoral or equivalent level | |

Source: International Standard Classification of Education (ISCED) 2011

Data accessed from Eurostat database, refer to the year 2015, the latest year for which are available official data. (See Table no. 3)

Table no. 3. - Population aged 25-64 by educational attainment level, and urban-rural typology, at the end of 2015

<table>
<thead>
<tr>
<th>Year</th>
<th>Urban/rural typology</th>
<th>Total population aged 25-64</th>
<th>Less than primary, primary and lower secondary education (ISCED levels 0-2)</th>
<th>Upper secondary and post-secondary non-tertiary education (ISCED levels 3 and 4)</th>
<th>Tertiary education (ISCED levels 5-8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
</tr>
<tr>
<td>2015</td>
<td>Predominantly urban regions</td>
<td>1,597,141</td>
<td>100</td>
<td>269,917</td>
<td>16,9</td>
</tr>
<tr>
<td></td>
<td>Intermediate regions</td>
<td>3,755,836</td>
<td>100</td>
<td>878,866</td>
<td>23,4</td>
</tr>
<tr>
<td></td>
<td>Predominantly rural regions</td>
<td>5,851,070</td>
<td>100</td>
<td>1,772,874</td>
<td>30,3</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>11,204,047</td>
<td>100</td>
<td>2,921,657</td>
<td>26,1</td>
</tr>
</tbody>
</table>

Source: Calculated by the author based on data from EUROSTAT Population on 1 January by five-year age group, sex and urban-rural typology [urt_pjangrp3], and Labor market statistics by urban-rural typology [urt_edat_lfse4], Extracted on 04.02.2017

To determine whether there is an association between educational attainment levels of the population and the urban-rural typology regarding their residence, was used "Chi-Square (denoted $\chi^2$) test of independence" also named as "test of association", which is a single-sample nonparametric test, because the two variables considered were categorical and ordinal, too.

For a valid result as the Chi-Square test return, it is advisable to check whether such test can be used. Therefore, it is necessary to verify the following required assumptions for the data considered:

- categorical variable;
- independence of observations, which means that there is no relationship between any of the cases;
- the groups of the categorical variable must be mutually exclusive;
- there must be at least 5 expected frequencies in each group of your categorical
variable. The use of Chi-Square test is restricted to large samples [19].

In Chi-Square test, firstly needs to formulate the hypotheses. The null hypothesis, H0, regards the independence of the two variables, in time that the alternative hypothesis, H1, formulates the statistical association of the considered variables, respectively:

**Null hypothesis H0:** The values of two categorical variables A and B are independent/are not associated;

**Alternative hypothesis H1:** The values of two categorical variables A and B are not independent/are associated.

The test statistic for the Chi-Square test of independence is computed using the formula 1:

\[ \chi^2_{\text{computed}} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij}-e_{ij})^2}{e_{ij}}, \]  

where:
- \( o_{ij} \) are the observed frequency of the ith row and jth column,
- \( e_{ij} \) are the expected frequency of the ith row and jth column,
- \( r \) is the number of rows in the contingency table, and
- \( c \) is the number of columns in the contingency table.

The degree of freedom of a contingency table with \( r \) rows and \( c \) columns is computed using the following formula 2.

\[ df = (r - 1) \times (c - 1) \]  

The decision on acceptance the hypothesis H0 is such:

a) If \( \chi^2_{\text{computed}} \leq \chi^2_{\text{critical}}(\alpha, df) \), then accept the null hypothesis, with \( \alpha \) set at 5% (0.05), or 1% (0.01).

b) Else, the null hypothesis is rejected.

The critical value, \( \chi^2_{\text{critical}}(\alpha, df) \), is from the table at the significance level (\( \alpha \)).

To highlight the association between the two variables, we need to prove the alternative hypothesis to be true. If a statistical significance of the association is identified, the degree of association is evaluated only indirectly. To evaluate in a direct way the degree of association must be used a specific measure of association based on the computed Chi-Square value [14]. Cramer’s V is the most popular measure of association because it is good for all type of table size. Cramer’s V is based on Pearson's Chi-Square statistic, so it gives a value between 0 and +1.

Cramer’s V is computed with the formula 3:

\[ V = \sqrt{\frac{\chi^2_{\text{computed}}}{n \times \min(c-1,r-1)}}, \]  

where \( n \) is the grand total of observations. For processing of data was applied SPSS v20.00 Nonparametric Tests procedure.

**3. RESULTS AND DISCUSSION**

The Chi-Square test statistic is based on the data from the table in the following figure, which is a crosstabulation table refers to the educational attainment level of the Romanian population aged 25-64, at the end of the year 2015 classified as urban-rural typology.
The distribution of the population by the urban-rural typology graphical represented in the figure no. 2 is correlated with Romania’s territorial profile. The 42 counties are grouped as it follows: 25 counties are predominantly rural regions, making up 59% of all, 15 counties are intermediate regions representing 36%, and only 2 regions are predominantly urban, being 5%.

The predominant upper secondary and post-secondary non-tertiary (ISCED levels 3 and 4) educational level of the predominantly rural population could indicate a possible disinterest to invest in their education, even if that would allow them to have got a job according to a higher qualification with higher wages. The inability to adapt the education system to the labor market requirements could be another cause [19].

The following hypotheses were formulated to evaluate the two considered variables association:

- **Null hypothesis, \( H_0 \):** The educational attainment level of the Romanian population aged 25-64 (less than primary, primary and lower secondary education (levels 0-2), upper secondary and post-secondary non-tertiary education (levels 3 and 4), and tertiary education
(levels 5-8) is independent on the type of residence region (predominantly urban, intermediate, and predominantly rural);

- **Alternative hypothesis, (H1):** The educational attainment level of the Romanian population aged 25-64 (less than primary, primary and lower secondary education (levels 0-2), upper secondary and post-secondary non-tertiary education (levels 3 and 4), and tertiary education (levels 5-8)) is related on the type of residence region (predominantly urban, intermediate, and predominantly rural);

On the assumption that there is no relationship between the regional typology and educational attainment of population aged 25-64, were computed the expected frequencies for each cell using SPSS nonparametric tests procedure.

The SPSS output shows in the figure no. 3, named 3x3 contingency table, presents the expected frequencies for each cell. Based on the contingency table the Chi-Square statistic is calculated.

![Figure no. 3.- SPSS output – The 3x3 contingency table for calculated Chi-Square statistic](image)

The computed Chi-Square Statistic is 417444.302 and 0 cells (0%) have expected count less than 5 as can see in figure no. 4. The p-value in Chi-Square output is \( p = 0.000 \). This means that the relationship between education level and region typology is significant.

Therefore, the hypothesis \( H_0 \) must be rejected and conclude that the two variables considered, respectively the territorial typology and educational attainment of the population, are not "independent". They are statistically associated. The differences between the three regional typologies of Romania on educational attainment categories by level are significant and not due to random sampling variation.

![Figure no. 4. – SPSS output – Chi-Square Test result](image)
However, the test, as all tests of significance, only elucidate that there is a relationship, the relationship has statistical significance, and in other words, it is not due to chance. Nevertheless, cannot say that regional typology necessarily causes the education attainment level, because the Chi-Square test of independence only assesses the association between the two variables, and cannot offer any suggestions about causation, but it does indicate that the reason for the association is worth investigating [21].

Using SPSS was computed Cramer’s V, as a measure of the strength of association among the levels of the row and column variables. With only a 0,136 value, a small association between considered categorical variables was confirmed by the Cramer’s V coefficient.

<table>
<thead>
<tr>
<th>Symmetriic Measures</th>
<th>Value</th>
<th>Approx. Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal by Nominal Phi</td>
<td>0.183</td>
<td>0.000</td>
</tr>
<tr>
<td>Cramer’s V</td>
<td>0.136</td>
<td>0.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>11204947</td>
<td></td>
</tr>
</tbody>
</table>

Figure no. 5. – SPSS output – Cramer’s V value for measure the strength of association

It means that educational attainment is small associated with the urban-rural typology of the residence of the Romanian population aged 25-65 years.

3. CONCLUSIONS

This study can help regions design their strategies to reducing economic disparities requires a more careful approach to the quality of human resources that can which can create jobs and provide services contributing to building social cohesion and reducing inequalities between geographic areas and different communities. A highly skilled workforce is essential for a competitive and sustainable economy. The Chi-Square test confirms the relationship between the education level of the population and level of development of the region of residence, so the typology of the geographical regions influences the educational attainment level of the population, but in a small measure. The fact that more than half of the population resides in rural areas makes the approach to education and training human resources from this area needed a special attention, especially if we consider that the family, living standards, and the sociocultural influence factors are so important.

BIBLIOGRAPHY


