Main Determinants of Romanian Emigration. A Regional Perspective

Goschin, Zizi

Bucharest University of Economic Studies

2016

Online at https://mpra.ub.uni-muenchen.de/88829/
MPRA Paper No. 88829, posted 17 Sep 2018 13:40 UTC
Main Determinants of Romanian Emigration. A Regional Perspective

Zizi Goschin
The Bucharest University of Economic Studies, Romania
Institute of National Economy, Romanian Academy

E-mail: zizi.goschin@csie.ase.ro

Abstract
Statistical data reveal considerably higher emigration rates in certain regions of Romania, usually the less developed ones, marked by high unemployment and poverty rates. This suggests that a deeper understanding of Romanian out-migration patterns and the underlying factors of influence can be reached by investigating the phenomenon from a regional perspective. Consequently the paper explores the major causes of Romanian emigration using county level data from the last Census. We are testing some potential push factors, such as low development, small income and high unemployment. The average level of education by county is also considered by using two relevant variables: share of secondary education graduates, as well as university graduates, in stable population aged 10 years and over. We test for spatial dependence in the number of emigrants by county and estimate both classic OLS regression models and spatial models, namely spatial lag and spatial error models. In accordance with previous findings in the literature, our results point to the development level (captured by GDP per capita) as a negative factor of influence for emigration. This means that poorer counties provide more emigrants compared to the richer ones. Surprisingly, the characteristics of the regional labour market, reflected by average wage and unemployment rate, influence the emigration flows in opposite direction than expected. We explained this outcome in the context of highly significant impact of human capital, as we found stronger propensity towards emigration for the persons having higher skills and qualifications, which are seldom affected by unemployment and low income. To sum up, our empirical analysis points to well-educated people from poorer counties as the most likely emigrants. This result should raise awareness on the waste of human capital (“brain drain”) with long-lasting negative consequences for the Romanian society and economy.

Keywords: emigration, push factors, spatial models, county, Romania.

JEL Classification: J61, R11, R15.

1. Introduction
Research on Romanian emigration has been gaining importance in recent years, in the context of the steady increase in the total number of emigrants worldwide. According to Ratha et al. (2013) Romania had the largest growth (4.8 %), among the countries of Europe and Central Asia, in the number of emigrants over 2010-2013. The massive loss of human capital following big emigration flows takes a heavy tool on GDP growth and remittances can’t
counterbalance it. This long-lasting phenomenon demands a thorough analysis of its causes and effects, as well as measures to mitigate it.

International migration commonly originates from major dissatisfaction with the socio-economic and politic situation in Romania. Alongside this overall cause that acts more or less evenly throughout the country, specific socio-economic local characteristics play additional supporting roles, determining significant variation in the regional shares of out-migrants. Statistical data reveal distinct regional patterns in migration, for instance higher emigration propensity in certain regions of Romania, marked by widespread poverty, ineffective labour market and lack of entrepreneurial opportunities. Therefore one can reach a better understanding of the emigration phenomenon and its underlying causes by investigating it from a regional viewpoint.

Contributing to previous empirical analyses that pointed to the main emigration determinants, usually at national level, we aim to identify the major drivers of Romanian emigration from the perspective of the migrant’s region of origin. We are going to test several likely “push” factors for emigration, such as low development (measured by GDP/capita), unsatisfactory income (captured by average wage), poor work prospects (unemployment rate) and education. Our choice of variables puts emphasis on the differences between the living standards of the counties, rather than just differences between earnings. We examine these factors of influence while accounting for spatial autocorrelation (spatial dependence) in order to understand the role of neighboring counties in emigration growth. To this aim spatial lag and spatial error models will be estimated, alongside classical regression. We use the regional migration database issued by the National Institute of Statistics, following the latest census (National Institute of Statistics, 2016).

The reminder of this paper is structured into five sections. Next section presents the main Romanian literature on emigration, section 3 analyses the regional differentials of emigration outflows, section 4 introduces the methodology of analysis, the variables and data, while section 5 presents and discusses the results focusing on the impact of education on the emigration flows. Finally, section 6 outlines the main conclusions and traces the directions for further research.

2. Literature review

The literature on Romanian migration displays a variety of approaches, targeting the socio-demographic profile of the emigrants, the socio-economic impact in sending and receiving countries, the loss of human capital (brain-drain), temporary work migration, return migration, etc. (Ghita et al., 2007; Rotila, 2008; Silas and Simina, 2008; Sandu, 2005 and 2010; Rotila, 2008; Vlădescu et al., 2008; Roman and Voicu, 2010; Boboc et al., 2011; Driouchi et al., 2012; Roman and Goschin, 2012 and 2014; Hinks and Davies, 2015).

Empirical studies that have been undertaken so far revealed that the migration is triggered by negative (rejection) factors characteristic for country of origin in opposition to positive (attraction) factors specific for the country of destination. The “push” factors that drive
emigration from Romania have been widely researched, the scholars emphasizing the negative effects of insufficient earnings, high unemployment, lack of business opportunities and disappointing career prospects, discontent related to political instability and wide-spread corruption, poor economic and social environment, negative social climate, etc. (e.g. Stoiciu et al., 2011; Goschin and Roman, 2012 and 2014).

Specific social and personal factors also play a role in the emigration decision: aiming for a better education in famous foreign universities, close links to friends and relatives in Romanian communities abroad, personal problems, the need to provide remittances for the subsistence of the family back home, etc. (Litan, 2009; De Sousa and Duval, 2010; Roman and Ileanu, 2010; Roman, 2011; Goschin, 2013).

Given the scarcity of reliable official statistics, previous studies used international databases (Boboc et al. 2011; Driouchi et al. 2012), recent data from the latest census (Prada et al., 2015; Goschin, 2015) or based on data collected from their own surveys (e.g. Goschin and Roman, 2014).

Much of the latest research on the emigration from Romania was directed towards identifying its main determinants and effects. However, to our best knowledge, very few publications in the literature addressed the issue from a regional point of view (e.g. Ghețău, 2005; Simina, 2008; Prada et al., 2015; Goschin, 2015) the lack of data severely limiting the research (Otoiu and Titan, 2015). From a spatial perspective, Simina (2008) pointed to poor counties in North-Eastern Romania as an important source of international emigrants. Prada et al. (2015) used 2011 Census data and found that regional emigration correlates weakly with the development level because all regions, regardless their wealth, supply significant flows of migrants. In contrast, characteristics of the regional labour markets are very important determinants of emigration propensity. Based also on 2011 Census data, Goschin (2015) showed that long-term and short-term emigration have different territorial distribution, but share some common factors of influence, such as low regional development and high human capital.

Potential spatial heterogeneity in emigration data needs to be included explicitly in the regression models used for regional migration analyses. Consequently, this paper contributes to previous literature on the regional determinants of Romanian out-migration by addressing this issue in the framework of spatial econometrics, in order to capture the interdependence between neighbouring counties.

2. Regional variation in emigration flows

The high number of international emigrants places the Nord-Eastern part of Romania on top position, with Bacau, Suceava, Neamt, Iasi, Galati and Maramures counties as the biggest emigration senders (Figure 1).
This is in accordance with the findings from previous studies (such as Simina, 2008; Sandu et al., 2004) that stressed the higher international migration propensity in the North-East Region, probably drawing on previous experience in internal rural–urban commuting, which was mainly poverty-driven and has been a constant in the region during the communist regime. Since this region continues to be the least developed area of the county, with scarce investments, low business opportunities and high unemployment, massive emigration seems a natural outcome.
At the lower end of the emigration distribution – counties that provide very few emigrants – are Ilfov, Giurgiu, Calarasi and Ialomita counties. All of them are situated in the Southern part of the country. Thus, the spatial structure of Romanian emigration suggests that the highs and lows have a tendency to cluster. The map of emigration at NUTS 3 level (Figure 2) does indeed reveal significant differences among the counties and a clear disposition of emigration levels to be clustered over space, having high values especially in North and North-East and low values in South and West.

**Fig. 2. Emigration clusters, 2011**
Source: processed by the author using OpenGeoda software

The existence of clusters of counties as regards emigration, as indicated by Figure 2, is suggesting the presence of spatial autocorrelation that needs to be tested with spatial statistics methods and if confirmed should be further addressed by using appropriate spatial models.

According to NIS data, the largest part (65.4%) of Romanian international emigration is long-term (over one year). Short-term migrants encompass many students, teachers, various professionals with high qualifications, as well as less skilled seasonal workers. The two components of total emigration tend to display different regional patterns and are driven by different reasons (Goschin, 2015).

**3. Variables, models and data**

Our aim is to analyse the Romanian emigration, at county level (NUTS2), focusing on the main regional factors of influence. The selection of variables is based on the literature (both
national and international), in the context of the current data limitations. Official data on Romanian emigration are scarce and only a small part of the emigration flows is currently recorded by the National Institute of Statistics. The latest Census provided better information on emigrants, but their number continue to be strongly underestimated: the Census indicated little over one million emigrants, while many unofficial estimations reach considerably higher values, up to four million persons.

The selected explanatory variables (Table 1) address three types of factors: regional development (proxied by Gross Domestic Product per capita), characteristics of the labour market (captured by average wage and the unemployment rate) and human capital (measured with two education variables: the share of secondary education graduates and the share of university graduates in stable population aged 10 years and over).

Table 1. The variables

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emigr</td>
<td>People who left abroad for less than a year or for a longer period (persons)</td>
<td>2011 Census</td>
</tr>
<tr>
<td>GDPcap</td>
<td>Gross Domestic Product per capita (Ron/inhabitant)</td>
<td>National Institute of Statistics and own calculations</td>
</tr>
<tr>
<td>Unempl</td>
<td>Unemployment rate (%)</td>
<td>National Institute of Statistics</td>
</tr>
<tr>
<td>Wage</td>
<td>Monthly average wage (Ron)</td>
<td>National Institute of Statistics</td>
</tr>
<tr>
<td>Tertiary</td>
<td>Share of university graduates in stable population of 10 and over</td>
<td>2011 Census and own calculations</td>
</tr>
<tr>
<td>Secondary</td>
<td>Share of secondary education graduates in stable population aged 10 years and over</td>
<td>2011 Census and own calculations</td>
</tr>
</tbody>
</table>

The empirical analysis is conducted in the framework of both classic Ordinary Least Squares and spatial regression models, in order to account for potential spatial autocorrelation. The classic linear multifactorial regression has the following specification:

\[
\text{Emigr}_i = \beta_0 + \beta_1 \text{GDPcap}_i + \beta_2 \text{Unempl}_i + \beta_3 \text{Wage}_i + \beta_4 \text{Tertiary}_i + \beta_5 \text{Secondary}_i + \varepsilon_i, \tag{1}
\]

where \(\beta_0\) to \(\beta_5\) are the parameters and \(\varepsilon\) is the error term.

We have to take into account the very likely spatial autocorrelation since previous analysis showed that neighbouring regions tend to have similar emigration patterns. Spatial autocorrelation will be tested using Moran’s \(I\) statistic (Anselin, 2005). If spatial dependence is confirmed, it should be corrected using the appropriate spatial model (Anselin, 2005). The first choice is the spatial autoregressive model (SAR) that includes a spatial lag of the dependent variable:
\[ \text{Emigr}_i = \rho W \text{Emigr}_i + \beta_1 \text{GDPcap}_i + \beta_2 \text{Unempl}_i + \beta_3 \text{Wage}_i + \beta_4 \text{Tertiary}_i + \beta_5 \text{Secondary}_i + \epsilon_i, \]  
\[(2)\]

where \( W \) is the spatial weight matrix capturing the spatial structure of the country, and \( W\text{Emigr}_i \) is the spatial lag of the dependent variable \( \text{Emigr} \).

The second choice is the spatial error model (SEM) which includes the spatial dependence in the error term \( \epsilon_i \):

\[ \epsilon_i = \lambda W \epsilon_i + \nu_i, \]

resulting the following model specification:

\[ \text{Emigr}_i = \beta_0 + \beta_1 \text{GDPcap}_i + \beta_2 \text{Unempl}_i + \beta_3 \text{Wage}_i + \beta_4 \text{Tertiary}_i + \beta_5 \text{Secondary}_i + (\lambda W \epsilon_i + \nu_i), \]
\[(3)\]

This model indicates that the error terms across different spatial units are correlated.

Finally, the choice of the best model for our data will be based on Lagrange multiplier test for spatial error and lag. Failing to reject the null hypothesis of spatial randomness indicates that classical OLS regression should be used instead of a spatial model.

In order to estimate the coefficients of all these models we exploit the new database on the regional number of Romanian emigrants, issued by the National Institute of Statistics, resulting from the latest Romanian Census, unfolded in October 2011 (NIS, 2015). The data reflect the post-crisis situation of Romanian emigration, therefore we expect significant differences from the findings in previous research. We also used other NIS databases for regional GDP, wage and unemployment rate and calculated some derivative indicators (GDP per inhabitant, share of secondary education graduates, and university graduates, in stable population aged 10 years and over). All data are recorded at county level (NUTS3).

4. Results and discussion

The coefficients of the equations specified in (1), (2) and (3) have been estimated in OpenGeoda, and the results are presented in Table 2.

The classic OLS regression model indicates regional development (captured by Gross Domestic Product per capita) as a negative factor of influence for emigration. This means that poorer counties, having lower standard of living, provide more emigrants compared to the developed ones. This result is consistent with the international literature and our own expectations.

Table 2 The results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classic OLS model</th>
<th>Spatial lag model</th>
<th>Spatial error model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Prob</td>
<td>Coefficient</td>
</tr>
<tr>
<td>( W_{\text{Emigr}} ) (spatial lag)</td>
<td></td>
<td>0.2717</td>
<td>0.1213</td>
</tr>
</tbody>
</table>
Surprisingly, the characteristics of the regional labour market, reflected by the variables average wage and unemployment rate, influence the emigration flows in opposite than expected ways: unemployment is highly significant and correlates negatively with the number of migrants, while average wage is less significant and seems to exert a positive influence on the number of regional emigrants. This outcome might be explained by taking into account another significant factor of influence, namely education. The number of emigrants from a county correlates positively with the number of tertiary educated persons, indicating that persons having higher skills and qualifications are more prone to emigrate, despite having employment opportunities on the local labour market. This suggests that the first wave of preponderant low and medium skilled temporary workers is increasingly giving way to higher educated emigrants, able to access better jobs abroad. Since such persons are well suited for good jobs on the local market as well and are less affected by unemployment, this could explain the unexpected signs of the variables unemployment and wage. It is also noteworthy that this change in the education level of the emigrants has a very negative impact on the Romanian economy, as it implies higher human capital loss (brain-drain) with long-term diminishing effects on potential economic growth. The other education variable (share of secondary education graduates in stable population) is statistically insignificant. Therefore our empirical analysis points to well-educated people from poorer counties as most likely to emigrate.

Since the results from the classic regression model are insignificant if spatial dependence in endogenous variable exists, we tested for potential spatial autocorrelation by employing a basic indicator in spatial statistics, namely Moran’s I. The value of Moran’s I and the permutation test presented in Appendix 1 clearly indicates that spatial autocorrelation is statistically significant for the variable number of emigrants by county. This implies that the overall parameters estimated for the entire country cannot adequately describe the emigration phenomenon at any given location (county). The presence of spatial dependency implies the...
need to use spatial regression models that acknowledge correlation among counties either as spatial lag or spatial error (Anselin, 2005). Consequently, the coefficients of the equations specified in (2) and (3) have been estimated in OpenGeoda, using the maximum likelihood estimation method, and the results are presented in Table 2.

The results from spatial lag model (Table 2) are unsatisfactory, based on the comparison with the classic model by means of Likelihood Ratio Test: the corresponding probability is well over the 5% threshold, therefore the autoregressive specification is inappropriate for our data and the classic OLS model performs better than the spatial lag model. Moreover, the estimated coefficients for the spatial lag variables of the model (W_emigr) bear high probability, therefore are statistically insignificant. Consequently, the spatial lag model is rejected. This means that the magnitude of emigration flow from each county is not correlated with the emigration from the neighboring counties.

Based on the same criteria - Likelihood Ratio Test, we accept the spatial error model as a better alternative to the classic OLS model (Table 2). Given the values of the Likelihood Ratio Test, Log likelihood and R-squared, and the high significance of the spatial error coefficient LAMBDA, the best specification for the dependent variable emigration is the spatial error model. This outcome may be motivated by omitted variable bias. It shows that there is significant spatial dependence in emigration among counties, but the correlation between neighbors is determined by variables not included in the model (errors).

The main conclusion of our study is that regional variation in emigration rates is explained mostly in terms of tertiary education, which is the strongest and most significant factor of influence in all model specifications. The higher the percentage of tertiary graduates in population, the higher the regional number of emigrants. This result can be understood in a larger international context. The number of graduates from Romanian universities has significantly increased over the last decade, while the amount of jobs available for them remains limited in this post-crisis environment and the wage-gap between Romania and Western Europe or United States is large enough to justify the costs and risks of leaving the country. These findings are confirmed by official statistics on education that indicate the decrease in resident population having upper secondary and tertiary education, despite the increase in education attainment at these levels. The difference between the number of university graduates and actual resident population is likely to be represented by Romanian emigrants abroad. Among CEE countries, Romania has the highest emigration rate of the tertiary educated (Brücker 2013): 20.36% in 2010 compared to 9.17% in 2000, while the emigration rate of secondary educated is comparatively low, reaching only 3.88% in 2010 (from 1.35% in 2000). Other recent studies found a significant rise in the education level of Romanian emigrants in the post-crisis environment (e.g. Alexe, 2011), warning that temporary migration of high-skilled professionals could become permanent.

5. Conclusions
In this paper we envisaged the main factors of influence on emigration from Romania, from a spatial perspective. The empirical analysis was conducted at county level, using data from the 2011 population census, in the framework of both classic Ordinary Least Squares and spatial regression models, in order to account for spatial autocorrelation.

We found that Romanian emigration is triggered by low regional development and high (tertiary) education. Tertiary education seems to be the most significant explanatory variable in all model specifications. Emigration of higher educated persons signifies waste of human capital (“brain drain”) with long-lasting negative consequences: scant supply of educated workforce in the future, public funds spent for emigrants’ education are lost, potential loss of national income, work shortage in critical areas (e.g. physicians, IT specialists). Besides, many short-term emigrants are students, who might be tempted to remain abroad after graduation in order to take advantage of better jobs and higher wages. Since highly skilled emigration is more likely to become permanent than the low skilled one (Faini, 2003), the potential human capital loss is larger and definitive.

Temporary emigration of tertiary educated could have positive impact on the Romanian economy only if it does not turn into a permanent one. Highly educated returnees, benefitting from the skill improvement and experience acquired abroad, could be very productive employees or efficient entrepreneurs, contributing to economic growth in Romania. Therefore a strategy focused on increasing the rate of return and reducing the emigration rate of the highly educated should be a priority in the future.

References


Anselin, Luc: Exploring Spatial Data with GeoDaTM: A Workbook, Spatial Analysis Laboratory, Department of Geography University of Illinois, Urbana, 2005 


Rotila, V. (2005) „The impact of the migration of health care workers on the countries involved: the Romanian situation”, South-East Europe Review, 1, pp. 53-77


Appendix 1. Moran’s $I* \text{ scatter-plot for: county emigration}$

*Moran’s $I$ is computed as:

$$MI = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where $i$ and $j$ indicate the regions and $w_{ij}$ are the spatial weights ($w_{ij}$ is 1 if regions $i$ and $j$ are neighbours, and 0 otherwise).

Appendix 2. Diagnostics for spatial dependence for the emigration models

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : Judete.gal (row-standardized weights)

<table>
<thead>
<tr>
<th>TEST</th>
<th>MI/DF</th>
<th>VALUE</th>
<th>PROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moran’s I (error)</td>
<td>0.1941</td>
<td>2.6823</td>
<td>0.00731</td>
</tr>
<tr>
<td>Lagrange Multiplier (lag)</td>
<td>1</td>
<td>3.0758</td>
<td>0.07946</td>
</tr>
<tr>
<td>Robust LM (lag)</td>
<td>1</td>
<td>0.1343</td>
<td>0.71400</td>
</tr>
<tr>
<td>Lagrange Multiplier (error)</td>
<td>1</td>
<td>3.5549</td>
<td>0.05937</td>
</tr>
<tr>
<td>Robust LM (error)</td>
<td>1</td>
<td>0.6133</td>
<td>0.43353</td>
</tr>
<tr>
<td>Lagrange Multiplier (SARMA)</td>
<td>2</td>
<td>3.6892</td>
<td>0.15809</td>
</tr>
</tbody>
</table>