GRAVITY MODEL APPLICATION TO ESTIMATE ROMANIAN MIGRATION SIZE

Professor Mariana BĂLAN, PhD.
Institute for Economic Forecasting-NIER, Romanian Academy,
Email: dr.mariana.balan@gmail.com

Professor Valentina VASILE, PhD.
Institute of National Economy -NIER, Romanian Academy,
Email: valentinavasile2009@gmail.com

Simona Maria STĂNESCU, PhD.
Research Institute for Quality of Life, Romanian Academy,
Email: simona_vonica@yahoo.com

Abstract
An impressive number of labour force migration/mobility models, both at national and international level, are presented in the literature. A large number of those either describe international migration versus internal migration or inter-regional migration in Europe or in other parts of the world. Some of the many methods and techniques in use for the analysis and study of migration/labour force mobility are: the Markov chains method, the Harris-Todaro model with its variants and extensions, models for the study of the economic impact of labour mobility in enlarged Europe, models for the study of researchers mobility, models for the study of labour mobility through wage flexibility, models to characterize the impact of labour mobility on macroeconomic indicators. This paper, presents Romanian emigration analyzed using the gravity model.

Keywords: migration, gravity model, migration statistics

JEL Classification: J62

Introduction
European Union enlargement and labour movement liberalization towards EU countries substantially increase the possibility of migration to developed countries. This phenomenon, however, influences economic growth both in the short and long term. Accordingly, there is a need for both measures of migration control and the use of forecasting techniques and models to estimate the economic impact of this phenomenon.

International migration is a complex, multidimensional phenomenon requiring for modeling and forecast to use analysis/methods developed by other disciplines. A group of methods used in forecasting international migration is based on deterministic mathematical models or on other approaches and techniques, which do not involve the uncertainty associated with the migration phenomenon.

Unlike deterministic methods and models, stochastic (probabilistic) instruments for migration...
analysis and forecasting evolved from probability theory. Although some of the models mentioned in
the literature relate to particular theories of migration, such as, for example, econometric
models in economic theory, they do not try to build, from a probability point of view, a general
theoretical framework of population flows. Since they present greater ability to tackle uncertainty
related to migration and forecasting, stochastic models are used more frequently than
deterministic models.

Based on an analogy with physics—Newton's idea of gravity—the gravity model used in
economics was developed by Lowry (1966) and then it was extended to include other economic
variables in order to represent the effects of 'push and pull' of migrants (persons). Sometimes, the
gravity model is used not only to describe geographical mobility based on physical distance,
but also to measure the similarities between the pairs of occupants in a matrix of job changes. This
paper presents, based on the gravitational model, estimates of migration flows out of Romania. The
results allow the forecast of the size of Romanian migration.

1. Theoretical background

The socio-economic, demographic and cultural characteristics of the migrant population are the
determining factors of migration. Some of them operate also as incentives or barriers, or change
their influence type and intensity depending on the time horizon or geographical area considered.
EU enlargement and, in the same time, the free circulation of labour to the EU have substantially
increased the possibility of migration from Romania to developed countries. The analysis of
this phenomenon in Romania compared with flows from other countries reveals the peculiarities
of Romanian migrants.

The analysis of statistical data from the Romanian National Statistics Institute shows the
total number of Romanian emigrants during 2005-2011 for some destination countries: Germany
with 14347 people, Italy with 12357, Canada with 11016, USA with 11016 and Spain with 5626,
followed by Austria, Hungary, Israel and France with about 3500 people. The analysis of the size
and structure of migration flows using various methods allows at least their short-term forecast.
The measurement, analysis and forecast of Romanian migration flows should become a
permanent task, mainly because their severe effects.

The classical migration analysis gravity model considers a flow from location \( i \) to location \( j \) to be
proportional with the population of the origin and destination countries and inversely proportional to
the distance between the two locations:

\[
M_{ij} = a \cdot P_i \cdot P_j \cdot f(d_{ij})
\]

(1)

where: \( a \) is a constant associated with overall mobility
\( P_i \) and \( P_j \) represent the population in the
countries of origin, respective of destination;
\( f(d_{ij}) \) inverse function of the distance between
the two locations.

Since on the ground \( d_{ij} = d_{ji} \) it results that net
migration between the two locations is zero. Therefore, the gravity model, in this formulation
can be interpreted as describing random flows that occur where there is no net migration. For the
empirical analysis case of geographic mobility, when net migration occurs, the gravity model is
typically extended with a generalized attraction factor so that regional economic variables can be
incorporated into the migration equation as determinants (Molho, 1986):

\[
M_{ij} = a \cdot P_i \cdot P_j \cdot f(d_{ij}) \times A_i B_j
\]

(2)

where: \( A_i \) is the "pushing" factor of the home
country and \( B_j \) is the "attraction" factor of the host
country.

Both \( A \) and \( B \) are positive and are functions of
explanatory variables for each location. To
maintain consistency with the original equation
(1), both variables should be calibrated.

In their work "The Deregulation of People
Flows in China: Did the Structure of Migration
Change?", Bao Shuming, Orn B. Bodvarsson, Jack
W. Hou and Zhao Yaohua developed several
versions of the traditional gravity model for inter-
regional migration applied for the study of
international or inter-provincial migration.

The uniqueness of this version resides in the
inclusion of variables related to the economic,
political and social aspects of the country/
province. As dependent variable is considered the
logarithm of gross emigration rate (log($M_{ij}$)), calculated as the volume of emigration from country $i$ to country $j$ divided by total migration for country $i$:

$$logM_{ij} = \alpha_0 + \alpha_1 \log D_{ij} + \alpha_2 \log NETWORK + \alpha_3 \log FDI_{ij} + \alpha_4 \log FAI_{ij} + \alpha_5 \log Y_{ij} + \alpha_6 \log E_i + \alpha_7 \log U_i +$$

$$+ \alpha_8 \log \text{MANEMP}_{ij} + \alpha_9 \log \text{URBAN}_{ij} +$$

$$+ \alpha_{10} \log \text{MINORITY}_{ij} + \alpha_{11} \log \text{WARM}_{ij} + \varepsilon_{ij}$$

where:

- $D_{ij}$ = rail distance (in km) between the capital of country $i$ to that of country $j$;
- $NETWORK =$ size of the migrant community residing in country $j$, who migrated from country $i$, measured as the rate of past migration flows;
- $FDI_{ij}$ = ratio of real FDI per capita in country $j$ and actual foreign direct investment per capita in country $i$;
- $FAI_{ij}$ = ratio of real fixed assets domestic investment per capita in country $j$ and real fixed assets domestic investment per capita in country $i$;
- $Y_{ij}$ = ratio between real per capita income in country $j$ and real per capita income in country $i$;
- $E_i, E_j =$ level of education in country $i$ and, respectively $j$;
- $U_i, U_j =$ unemployment rates during the week preceding implementation of the census in country $i$ and $j$ respectively $i$;
- MANEMP$_{ij}$ = ratio between employment in the manufacturing sector share for country $j$ and employment in the manufacturing sector share in country $i$;
- URBAN$_{ij}$ = ratio between urban population share for country $j$ and urban population share for country $i$;
- MINORITY$_{ij}$ = ratio between the minority population share in country $j$ and the minority population share in country $i$;
- WARM$_{ij}$ = ratio of the annual average temperature in the capital of country $j$ and the annual average temperature in the capital of country $i$;
- $\varepsilon_{ij} =$ error term

**Model assumptions:**

- Migration rate may be positively or negatively correlated with any type of investment expenditure in the country of destination compared to that in the origin country;
- The ratio between employment in industry share for the destination country and at home is included in the model as an industry control mix;
- Destination countries that are relatively more urbanized, have different jobs available, different standards of living, etc. all with influence on migration flows;

2. Study of Romanian emigration development using the gravity model

For the study of Romanian emigration, we used a simplified gravitational model, because the data provided by Eurostat statistics, National Statistical Institute of Romania (Tempo database on-line), or Statistical Yearbooks of some EU countries where Romanian migration is significant—are not complete for all exogenous variables used in the gravity model. As a solution, the analysis of Romanian migrant flows for 1995-2010 was reported only for some EU countries and also for EU-27.

Exogenous variables considered in the model were:

- GDP for country $i$;
- Employment rate for country $i$;
- Unemployment rate for country $i$;
- Foreign direct investment for country $i$;
- Number of Romanian emigrants in each country studied and for EU-27;
- Distance (on road) between Romania and the various destinations of Romanian emigrants.

The general form of the model used to analyze Romanian emigration to some EU countries, USA and Canada with the considered variables is:

$$Emigr_{ij} = \alpha_0 + \alpha_1 \log \text{PIB}_{i} + \alpha_2 \log \text{PIB}_{Rom} + \alpha_3 \log \text{Rsomaj}_{ij} + \alpha_4 \log \text{Rsomaj}_{Rom} + \alpha_5 \log \text{Rocup}_{ij} +$$

$$+ \alpha_6 \log \text{Rocup}_{Rom} + \alpha_7 \log \text{dist}_{Rom} + \alpha_8 \log (\text{FDI}_{ij}) + \varepsilon_{ij}$$
Estimates of Romanian emigration to Spain, the UK, Italy, Germany, USA, Canada, France, Austria, and Hungary have revealed a determination report \( R^2 \) with values above 75%, indicating that the exogenous variables considered have a significant impact on the dimension of Romanian migration phenomena. Thus, for Spain, the dependence degree of the endogenous variable of all the independent variables is 98.75%, for the UK 94.5%, for Italy 78%, for Germany 96%, for USA 93.5%, for Canada 89%, for France 86.3%, for Austria 97.5%, and for Hungary 92.4%.

The lack of adequate statistical databases limit migration analysis and therefore, for some countries, the estimates were made with fewer exogenous variables leading to lower values of the coefficient of determination (the case of Italy).

Because the data sample has less than 30 observations, we used the Student variable \( t_{\alpha, \nu} \), whose values were taken from the Student distribution table (according to the value determined for \( \alpha \) and the number of degrees of freedom, \( \nu \)). For the cases considered, its value is 1.7459 (with a significance level of 5%).

The analysis of the calculated values for each of the determined estimators indicates that for all \( t_{calc}>1.7459 \) as they are significantly different from zero. For the considered model, for each estimator its probability indicates that it can be accepted as unbiased, consistent and efficient.

Durbin-Watson statistic, which is one of the most common procedures used to identify first-order errors autocorrelation in linear regression models, has values between 2.21 and 2.80, which indicates no autocorrelation of errors.

F test, measuring how well the independent variables explain the evolution of the dependent variable, for each considered regression, has values below the F distribution (Fisher-Snedecor) corresponding to the 9 degrees of freedom (39.55) and the significance level of 5%. Therefore, emigration evolution is stable over time, and the model, with all its drawbacks, can be used for forecasting.

Eurostat statistics have enabled an estimate of the emigration from Romania to European Union countries. The gravity model led to the estimators presented in table 1:

| Table 1 |
|-----------------------|-----------------------|-----------------------|
| Dependent Variable: EMIGR |
| Method: Least Squares |
| Sample(adjusted): 1995 2010 |
| Included observations: 16 after adjusting endpoints |
| EMIGR=C(1)+C(2)*GDPEU+C(3)*GDPRO+C(4)*RUNEMPUE+C(5)*RUNEMPRO+C(6)*REMPEU(-3) +C(7)*REMPRO |
| Coefficient | Std. Error | t-Statistic | Prob. |
| C(1) | -5724.5 | 186266.3 | -3.073452 | 0.0372 |
| C(2) | 8.100699 | 2.751530 | 2.944071 | 0.0422 |
| C(3) | -12.49417 | 6.346712 | 1.968605 | 0.0120 |
| C(4) | -4957.699 | 1106.996 | 4.478514 | 0.0110 |
| C(5) | 1014.022 | 844.0468 | 1.901382 | 0.0295 |
| C(6) | 6914.966 | 2504.218 | 2.761328 | 0.0508 |
| C(7) | -1703.000 | 435.2528 | 3.912669 | 0.0174 |
| R-squared | 0.915846 | Mean dependent var 11635.67 |
| Adjusted R-squared | 0.768577 | S.D. dependent var 2842.375 |
| S.E. of regression | 1367.367 | Akaike info criterion 17.51388 |
| Sum squared resid | 7487865. | Schwarz criterion 17.83715 |
| Log likelihood | -97.08329 | F-statistic 6.218851 |
| Durbin-Watson stat | 2.562207 | Prob(F-statistic) 0.048314 |
Durbin-Watson statistic has the value 2.56 which indicates no autocorrelation of errors (table 1).

F test, that measures how well the independent variables explain the evolution of the dependent variable, has a value of 6.22 (table 1), lower than the F distribution (Fisher-Snedecor) corresponding to 7 degrees of freedom and the 5% significance level \( F_{7,0.05} = 3.97 \). Therefore, emigration evolution is stable over time, and the model, with all its drawbacks, can be used for forecasting.

**Conclusion**

Using gravity model to estimate the number of emigrants in general and Romanians in particular, can lead to significant results from an econometric point of view. The variables considered for these estimates describe with a good percentage the variation of the number of Romanian emigrants in each of the countries surveyed. However, these estimates can be substantially improved if:

1. the data would be available for a longer period of time;
2. the information about the number of Romanian emigrants would be more detailed;
3. other exogenous variables, such as R&D&I costs would be introduced;
4. qualitative variables, which very often have great importance in the decision to migrate, would be introduced;
5. "education and training", etc. were considered as exogenous variables

**BIBLIOGRAPHY**