Calibration and Testing the Gravity Model on Migration Data

S u m m a r y

J. Paulov - Š. Poláčik

Notation and Abbreviations:

T_{ij}^{*} observed migration flows;

T_{ij}: predicted migration flows;

$[T_{ij}^{*}]_{i,j=1}^{m,n}$ matrix of values T_{ij}^{*};

$[T_{ij}]_{i,j=1}^{m,n}$ matrix of values T_{ij} ;

$M_i^{(1)} M_j^{(2)}, O_i, D_j,$: mass measures;

c_{ij}: generalized measure of distance;

K, A_i, B_j, α, β: parameters of gravity model;

a, b: coefficients of regression equation $T_{ij} = a + bT_{ij}^{*}$;

r: coefficient of correlation between $T_{ij}^{*}$and T_{ij} ;

$\sum_i \sum_j (T_{ij} - T_{ij}^{*})^2$= s; C'($\alpha_i$) — C' = z': difference between the left-hand and right-hand side of equation (21);

C(β_i) — C = z: difference between the left-hand and right-hand side of equation (22);

MNŠ: least squares method;

IM: iteration method;

MNV: maximum likelihood method;

FV: physical distances (km);

EV: economic distances (railway tariffs, Kcs);

StC: Middle-Czech region; JC: South-Czech region; ZC: West-Czech region; SC: North-Czech region; VC: East-Czech region; JM: South-Moravian region; SM: North- Moravian region; ZS: West-Slovakian region; SS: Middle-Slovakian region; VS: East- Slovakian region.

The study is an attempt at calibration and testing the gravity model on migration, data among the Czechoslovak regions in individual years of the interval 1961 - 1965.

In the first part of the study there are traced the theoretical fundaments of the gravity model. There are mentioned its several forms, from the primary ones, based on analogy with the Newtonian mechanics, up to the forms based on entropy maximizing principle.

The second part of the study is devoted to proper calibration and testing the gravity model. As the subject of calibration and testing we have chosen these four following forms: (5), (8), (15), (18). In each of these

forms we have worked with physical distances (km) as well as with economic distances (railway tariffs, Kcs). The distances among regions were considered as distances among regional seats, only in cases, when the seat of a region had a too eccentric position we have chosen another town with approximately centric position (there are 3 such cases).

The inner-regional distances were considered as radii of circles, whose surface is equal to the surface of relevant region.

For the estimation of parameters K, A_i, B_j, α, β we have employed 3 methods:

(a) least square method [equations (19), (20)],
(b) iteration method (in the estimation of parameters A_i, B_j),
(c) maximum likelihood method [equations (21), (22)].

As ,,testing" criteria (goodness-of-fit statistics) we have employed the coefficients of regression equation

$T_{ij} = a + bT_{ij}^*$; and $\sum_i \sum_j (T_{ij} - T_{ij}^*)^2$ The migration matrices have a dimension 10 x 10, in the case of models (5) and (18) however, we work with aggregated triangle matrices with a number of elements 55. The numeric results of calibration and testing are shown in the table enclosure. From the five years of observation we have chosen the year 1965. The table enclosure involves also the number of inhabitants of individual regions for 1965 (employed as the mass measures $M_i^{(1)} M_j^{(2)}$ and matrices of physical and economic distances.

The results of calibration and testing can be concisely summarized as follows:

(A)In all the discussed forms of the gravity model the parameter α or β indicates that the ,.deterrent" influence of the distance upon the migration among the Czechoslovak regions is relatively small. A significant mark of both parameters is their considerable stability during the studied period which indicates a certain stationarity of the migration process.

(B) Parameters A_i and B_j, called balance factors, which are the measures of accessibility and competition of interacting territorial units indicate that for migration in general the Czech and Moravian regions are better accessible than the Slovak ones. An exception is only the South-Czech region which is „similar" to the Slovak regions.

©The residuals show a great variability (some values are relatively high). We presuppose that the size of residuals besides factors which are not explicitly considered in the model is influenced by only approximately defined inter-regional and inner- regional distances and the fact that we work also with the inner-regional migration which is not so well comparable with the inter-regional migrations. In this study however, we do not deal with the analysis of the residuals.

(D) The correlation coefficient (r) considered as a descriptive measure of dependence is in all followed eases relatively high. Not even in one case it decreases below the value 0.75 whereby its maximum value is 0.98. To this effect all the predicted migration matrices agree relatively adequately with basic trend in the real migration data.

(E) The values of the regression coefficient ,,a" show that those forms of gravity model in which the estimation of parameters was realized by the least squares method underestimate the data. In other cases it is prevailingly the overestimation of data. The parameter ,,b" in all the followed cases is about 1.0 which indicates again that the fundamental trend in the migration data is kept in the modeling process. The optimal values of both regression coefficients were reached in the model form (8) and the employment of economic distances.

(F) Those forms of gravity model in which the parameters A_i, B_j substitute the parameter K predict the migration data apparently closer to reality.

(G) The maximum likelihood method in the estimation of parameters secures in the same models in general better conformity of the predicted data with the real data.

(H) The employment of physical distances brings in general better results in power forms of the gravity model than in exponential forms and paralelly in that form of the gravity model in which the constraints (12) and (13) are not required.

(1) The predominance of the exponential forms of the gravity model compared with the power forms in general becomes apparent in the employment of economic distances and at the same time of the validity of constraints (12) and (13).

Translation P. Misseje