GEOGRAFICKÝ ČASOPIS

52 2000 1

*Anton Bezák**

INTERREGIONAL MIGRATION IN SLOVAKIA: SOME TESTS OF SPATIAL INTERACTION MODELS

A. Bezák: Interregional migration in Slovakia: some tests of spatial interaction models. Geografický časopis, 52, 2000, 1, 3 figs., 5 tabs., 29 refs.

In this paper observed migration flows between administrative districts in Slovakia during the 1984-1995 period are used to test a set of 13 alternative spatial interactions models. The results obtained indicate that the doubly constrained spatial interaction models with the power or Tanner distance functions provide the best reproduction of observed flows, but the overall performance level of all the models tested is not satisfactory. In order to improve the overall model fit, the doubly constrained spatial interaction model with zone-specific distance decay parameters is examined in some detail. Both origin- and destination-specific parameter versions of the model are calibrated and the spatial variation in the propensity to migrate over distance is analysed. It is shown that the variation is very little in space and stable in time.

Key words: spatial interaction models, calibrating and testing, performance level, interregional migration flows, propensity to migrate over distance, Slovakia

INTRODUCTION

Spatial interaction models have played an important role in urban and regional analysis. They are used to predict spatial choices reflected in flows of people, goods or information between origins and destinations at a variety of different spatial scales. The theoretical framework of spatial interaction modelling was established over thirty years ago by British geographer A. G. Wilson (1967). Since his derivation

* Geografický ústav SAV, Štefánikova 49, 814 73 Bratislava

of a family of spatial interaction models using entropy-maximizing techniques, there has been a large volume of applications of this modelling approach in various fields of human geography and regional science. Population migration represents a classic form of spatial interaction with a very long history of investigation through the use of mathematical models. It is not surprising, therefore, that applications of spatial interaction models in migration studies are numerous. Examples can be found in Stillwell (1978, 1986, 1991), Plane (1982, 1984), Tobler (1983), Willekens (1983), Ledent (1985), Fotheringham (1986), Ishikawa (1987), Pooler (1987, 1993), Duley and Rees (1991), Rees (1994) and Sandraps (1995).

In Slovak geographical literature several interesting attempts have been made to contribute to the spatial interaction theme. A series of papers written by Paulov (1991, 1993, 1996) is noteworthy in this respect. Despite the considerable interest in the theoretical properties of spatial interaction models, little attention has been given to the empirical testing and validation of various models in the Slovak context. A partial exception is the paper by Paulov and Poláčik (1979) in which four spatial interaction models were fitted to data on migration between ten administrative regions of former Czechoslovakia.

This paper is intended to demonstrate the application of spatial interaction models for exploring interregional migration in Slovakia. Using data on migration flows between 36 old administrative districts for the 1984-95 period, 13 alternative spatial interaction models are calibrated and compared in terms of their ability to replicate observed data. Particular attention is given to parameter estimates that can provide insights into the migration system under investigation.

The next section of the paper describes a set of spatial interaction models selected for application. After this, the data set and regional units used in the analysis are briefly discussed. In the following section the results of model calibrations are presented and their empirical implications are examined. The last section contains some tentative conclusions.

SPATIAL INTERACTION MODELS

The theoretical basis for the family of spatial interaction models and their derivation using entropy-maximizing techniques are well known (cf. Wilson 1967, 1970; Fotheringham and O'Kelly 1989, pp. 15-31; Paulov 1993, pp. 37-44). For this reason only some brief comments on the models selected will be given and discussion will be centred around model equations, calibration procedures and goodness-of-fit statistics.

Following Stillwell (1991), the general form of the spatial interaction model in a migration context can be expressed verbally as follows:

Migration from region i to region j = Scaling constant or Balancing factors
× Origin outmigration or Propulsiveness factor
× Destination inmigration or Attractiveness factor
× Distance function [1]

Different versions of the model are then defined according to the information that is known about migration over some given period of time. The scaling constant or balancing factors operate to ensure that migration flow predictions are consistent with known information.

The basic spatial interaction model is derived for the situation where migration flows are constrained to known origin region outmigration totals and known destination region inmigration totals. This so-called *doubly constrained spatial interaction model* has the form

$$M_{ij} = A_i \, B_j \, O_i \, D_j \exp(-\beta \, c_{ij}) \tag{2}$$

where M_{ij} is the number of migrations from origin region i to destination region j, O_i is the total outmigration from region i, D_j is the total inmigration to region j, c_{ij} is the distance between regions i and j, and β is a distance decay parameter representing the frictional effect of distance on migration or the propensity to migrate over distance. A high parameter value indicates that distance has a stronger frictional effect on migration, whereas a low parameter value suggests that migrants are less affected by the distance over which they move. Note that the subscripts i and j in equation [2] vary from 1 to n, where n is the number of regions considered.

The other terms on the right-hand side of the equation [2] are the balancing factors defined as

$$A_i = \left[\sum_{j=1}^{n} B_j \, D_j \exp(-\beta \, c_{ij}) \right]^{-1} \tag{3}$$

$$B_j = \left[\sum_{i=1}^{n} A_i \, O_i \exp(-\beta \, c_{ij}) \right]^{-1} \tag{4}$$

which are introduced to ensure that the following constraints on total outmigration and inmigration are satisfied

$$O_i = \sum_{j=1}^{n} M_{ij} \tag{5}$$

$$D_j = \sum_{i=1}^{n} M_{ij} \, . \tag{6}$$

In addition, a migration length constraint is imposed on the migrations. This constraint has the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} \, c_{ij} = C \tag{7}$$

and requires that the model reproduces some observed total migration length, C.

Other members of the spatial interaction model family can be derived by using appropriate combinations of the constraints [5], [6] and [7] or by modifying some of them (Wilson 1971). If data are available only on outmigration from each origin region, then the constraint [6] must be dropped. In this situation the so-called *production constrained spatial interaction model* is formulated, in which the unknown total D_j is replaced by a term W_j representing the attractiveness of the destination region j as a recipient of inmigrations. Alternatively, if data are available only on inmigration to each destination region, then the constraint [5] is dropped and the *attraction constrained spatial interaction model* is formulated. In this model the unknown total O_i is replaced by a term V_i representing the propulsiveness of the origin region i as a generator of outmigrations. These two models are usually referred to as *singly constrained spatial interaction models*. Furthermore, the modification of the migration length constraint [7] can lead to the spatial interaction models with different distance functions. For example, if c_{ij} is replaced by ln c_{ij} in the equation [7], then the negative exponential function, $\exp(-\beta c_{ij})$, is replaced by the negative power function, $c_{ij}^{-\beta}$.

In this paper thirteen spatial interaction models were selected in order to compare the levels of performance which they achieve. The models are denoted by Roman numerals I to XIII and their equations (including ones for the balancing factors) are given in Table 1.

Model I is the basic production constrained spatial interaction model. Model II is similar to model I but an additional parameter, α, is attached to variable W_j to represent some measure of how the destination attractiveness varies. Models III and IV are two other versions of the production constrained spatial interaction model in which the exponential distance function was replaced by the power function. Model III is related to model I and model IV to model II. Exactly the same principle of model specification was applied to produce four attraction constrained spatial interaction models denoted by numerals V to VIII. Note that an additional parameter, μ, representing some measure of the origin propulsiveness variation is now attached to variable V_i in both models VI and VIII.

Several doubly constrained spatial interaction models were also examined. Model IX is the basic doubly constrained spatial interaction model with the negative exponential distance function. Two other versions of this model were produced by replacing the exponential distance function by the power distance function (model X) or Tanner distance function (model XI).

As the last alternative, the doubly constrained *spatial interaction model with zone-specific distance decay parameters* can be investigated. The reason is that the distance decay parameters associated with the above described models are representative of the whole matrix of interregional migration flows. In this respect models I to XI have generalized interregional parameters indicating the overall propensity to migrate over distance. It is clear, however, that the propensity to migrate is likely to vary spatially. According to Stillwell (1978), the extent of this variation can be identified by the use of spatial interaction models which incorporate zone-specific decay parameters. In this study both origin-specific (model XII) and destination-specific (model XIII) parameter versions of model X were utilized.

The process of assigning optimal values to model parameters is generally known as calibration. A variety of alternative methods exists for calibrating spatial interaction models (cf. Batty and Mackie 1972, Openshaw 1976, Baxter 1983, Sen 1986, Fotheringham and O'Kelly 1989, pp. 43-65). The most commonly used method of calibration is maximum likelihood estimation. This method is based on finding the maximum of a likelihood function which expresses the probability of different parameter values producing an observed result. In general, the calibration problem is reduced to a problem involving the solution of as many maximum-likelihood equations as there are unknown parameters. For purposes of this study a special program based on the Newton-Raphson iterative routine for solving systems of non-linear equations was written in Visual Basic 6.0 and used for calibrating all spatial interaction models considered.

To complete this discussion, it should be noted that a modified calibration procedure, described by Stillwell (1978, 1991), was employed for calibrating the models XII and XIII with zone-specific parameters. The procedure consists of three different stages. In the first stage, a generalized distance decay parameter is determined by the Newton-Raphson routine. The second stage involves repeating the same procedure for each origin or destination region using generalized parameter as a starting value and the balancing factors defined in the first stage. Thus, a zone-specific parameter is determined in each case, which satisfies its corresponding zonal mean-migration-length convergence criterion. In the final stage the balancing factors and model equation are recalculated with the complete set of optimum zone-specific parameters in order to ensure that the overall constraints are satisfied. Exactly the same procedure is used to calibrate both origin- and destination-specific parameters, but in the case of the latter the original migration matrix has to be transposed and the row and columns totals exchanged.

The ability of the models considered to replicate the pattern of observed migrations can be assessed in a number of different ways. In this study three goodness-of-fit statistics were used to measure the degree of agreement between observed M_{ij} and predicted $\hat{M}_{ij}$ migrations. The first of them, the Standardized Root Mean Square Error (SRMSE) is defined as

$$\mathrm{SRMSE} = \left[\sum_{i=1}^{n} \sum_{j=1}^{n} (M_{ij} - \hat{M}_{ij})^2 / (n^2 - n) \right]^{1/2} \Bigg/ \left[\sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} / (n^2 - n) \right] \qquad [8]$$

The statistic has a lower limit of zero indicating perfectly accurate predictions but its upper limit is variable and depends on the distribution of the observed migrations, although in practice it is often 1.0. The Index Of Dissimilarity (IOD) measures the percentage of the total number of the predicted migrations that would have to be reallocated between regions to replicate exactly the matrix of the observed migrations. It is calculated as

$$\mathrm{IOD} = 50/M \sum_{i=1}^{n} \sum_{j=1}^{n} | M_{ij} - \hat{M}_{ij} | \qquad [9]$$

where M is the total number of migrations. The statistic has a minimum of zero when

Tab. 1. Spatial interaction models used in the analysis

I	$M_{ij} = A_i\, O_i\, W_j \exp(-\beta\, c_{ij})$	$A_i = \left[\sum_{j=1}^{n} W_j \exp(-\beta\, c_{ij}) \right]^{-1}$
II	$M_{ij} = A_i\, O_i\, W_j^{\alpha} \exp(-\beta\, c_{ij})$	$A_i = \left[\sum_{j=1}^{n} W_j^{\alpha} \exp(-\beta\, c_{ij}) \right]^{-1}$
III	$M_{ij} = A_i\, O_i\, W_j\, c_{ij}^{-\gamma}$	$A_i = \left[\sum_{j=1}^{n} W_j\, c_{ij}^{-\gamma} \right]^{-1}$
IV	$M_{ij} = A_i\, O_i\, W_j^{\alpha}\, c_{ij}^{-\gamma}$	$A_i = \left[\sum_{j=1}^{n} W_j^{\alpha}\, c_{ij}^{-\gamma} \right]^{-1}$
V	$M_{ij} = B_j\, V_i\, D_j \exp(-\beta\, c_{ij})$	$B_j = \left[\sum_{i=1}^{n} V_i \exp(-\beta\, c_{ij}) \right]^{-1}$
VI	$M_{ij} = B_j\, V_i^{\mu}\, D_j \exp(-\beta\, c_{ij})$	$B_j = \left[\sum_{i=1}^{n} V_i^{\mu} \exp(-\beta\, c_{ij}) \right]^{-1}$
VII	$M_{ij} = B_j\, V_i\, D_j\, c_{ij}^{-\gamma}$	$B_j = \left[\sum_{i=1}^{n} V_i\, c_{ij}^{-\gamma} \right]^{-1}$
VIII	$M_{ij} = B_j\, V_i^{\mu}\, D_j\, c_{ij}^{-\gamma}$	$B_j = \left[\sum_{i=1}^{n} V_i^{\mu}\, c_{ij}^{-\gamma} \right]^{-1}$
IX	$M_{ij} = A_i\, B_j\, O_i\, D_j \exp(-\beta\, c_{ij})$	$A_i = \left[\sum_{j=1}^{n} B_j\, D_j \exp(-\beta\, c_{ij}) \right]^{-1}$

Tab. 1. (continued)

		$B_j = \left[\sum_{i=1}^{n} A_i O_i \exp(-\beta c_{ij}) \right]^{-1}$
X	$M_{ij} = A_i B_j O_i D_j c_{ij}^{-\gamma}$	$A_i = \left[\sum_{j=1}^{n} B_j D_j c_{ij}^{-\gamma} \right]^{-1}$
		$B_j = \left[\sum_{i=1}^{n} A_i O_i c_{ij}^{-\gamma} \right]^{-1}$
XI	$M_{ij} = A_i B_j O_i D_j c_{ij}^{-\gamma} \exp(-\beta c_{ij})$	$A_i = \left[\sum_{j=1}^{n} B_j D_j c_{ij}^{-\gamma} \exp(-\beta c_{ij}) \right]^{-1}$
		$B_j = \left[\sum_{i=1}^{n} A_i O_i c_{ij}^{-\gamma} \exp(-\beta c_{ij}) \right]^{-1}$
XII	$M_{ij} = A_i B_j O_i D_j c_{ij}^{-\gamma_i}$	$A_i = \left[\sum_{j=1}^{n} B_j D_j c_{ij}^{-\gamma_i} \right]^{-1}$
		$B_j = \left[\sum_{i=1}^{n} A_i O_i c_{ij}^{-\gamma_i} \right]^{-1}$
XIII	$M_{ij} = A_i B_j O_i D_j c_{ij}^{-\gamma_j}$	$A_i = \left[\sum_{j=1}^{n} B_j D_j c_{ij}^{-\gamma_j} \right]^{-1}$
		$B_j = \left[\sum_{i=1}^{n} A_i O_i c_{ij}^{-\gamma_j} \right]^{-1}$

predictions are perfect and a maximum of 100 in the reverse case. The final statistic is the coefficient of determination (R^2) measuring the proportion of total variation in the matrix of observed migrations which is statistically explained by the matrix of predicted migrations. The statistic R^2 ranges between 0 and 1. Zero indicates no correspondence between the observed and projected migrations, one indicates perfect correspondence. A detailed discussion of various statistics for comparing observed and predicted spatial interaction matrices can be found in Knudsen and Fotheringham (1986).

DATA

The spatial interaction models presented in the preceding section were fitted to data on migrations between the old administrative districts in Slovakia over the 1984-95 period. The primary data used in this study are the data on the total number of persons leaving a given administrative district for another district, which are reported annually by the Statistical Office of the Slovak Republic. Note that these data are counts of moves rather than of transitions. If a person makes several moves across district boundaries during a given year, it appears in the data set as many times as the person moves. In order to facilitate comparisons among districts, each of the two urban districts of Bratislava and Košice was amalgamated with its rural counterpart into one metropolitan district, so that the number of districts used in the analysis was reduced from thirty-eight to thirty-six. The 36 district system is shown in Figure 1.

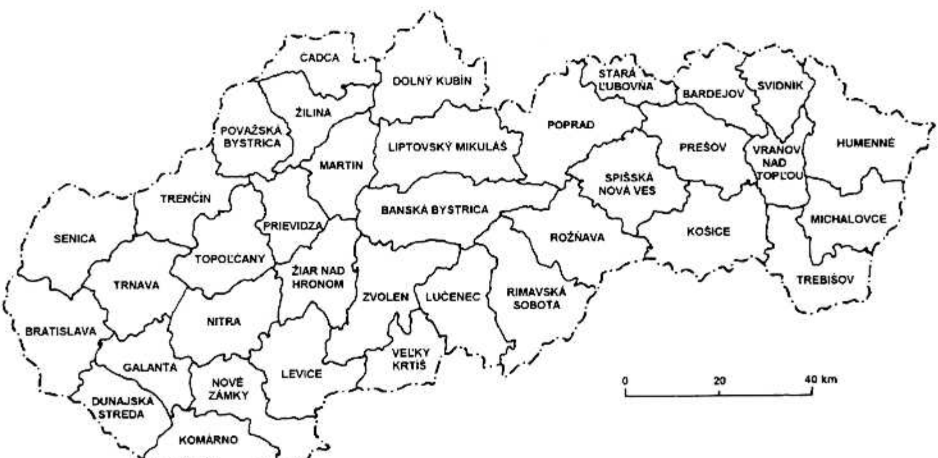

Fig. 1. The 36 district system used for the analysis of interregional migration flows.

For purposes of this paper, the annual data were consolidated into four three-year sets covering the periods 1984-86, 1987-89, 1990-92, and 1993-95. The corresponding matrices of interregional flows, each reflecting an annual average over these four periods, were used in the analysis. The main reason here was one of avoiding the problems of sparse matrices which arise with data for single-year migration flows between districts. In all cases migration within districts was ignored since the emphasis of the model calibration was on exploring interregional migration. Consequently, all diagonal elements in the migration matrices were set to zero.

Distances between districts were measured as the road distances between the district centres. Corresponding values expressed in kilometres were taken from the military road atlas (Kontra 1995). Finally, the population of the district at the beginning of each three-year period was used as the sole measure of both origin propulsiveness and destination attractiveness. All data used were obtained from the records of the current registration of population, as reported by the Statistical Office of the Slovak Republic.

RESULTS

Spatial interaction models with generalized distance decay parameters

The first eleven spatial interaction models with generalized distance decay parameters were calibrated using data for all the four periods considered. The results of the calibration for the 1993-95 period are summarized in Table 2 which contains the parameter values obtained for each model and the corresponding values of all three goodness-of-fit statistics. Of these indicators the IOD is the more readily interpretable in terms of overall model performance. As shown, its magnitude varies in the range from 16.8 to 25.8 %. This means that approximately a fifth to quarter of the predicted migrations was misallocated by the spatial interaction models calibrated. The values of the SRMSE are also relatively large and in the case of model I the SRMSE is even greater than 1.0 indicating that the average error is greater than the mean observed flow. Moreover, the coefficient of determination reaches the value of 0.8 only in two cases. These results clearly suggest that the degree of correspondence between the predicted and observed migrations is not very high and substantial errors occur in model predictions.

Tab. 2. Parameters and goodness-of-fit statistics for spatial interaction models fitted to migration data for the 1993-95 period

Model	Parameter				IOD	SRMSE	R^2
	μ	α	γ	β			
I	.	.	.	0.0105	25.8	1.006	0.725
II	.	1.249	.	0.0105	25.0	0.997	0.726
III	.	.	1.350	.	20.8	0.931	0.761
IV	.	1.267	1.367	.	19.9	0.926	0.771
V	.	.	.	0.0105	24.1	0.953	0.751
VI	1.035	.	.	0.0105	24.1	0.953	0.752
VII	.	.	1.370	.	19.0	0.879	0.798
VIII	1.045	.	1.372	.	18.9	0.880	0.799
IX	.	.	.	0.0109	22.9	0.907	0.776
X	.	.	1.450	.	17.0	0.726	0.861
XI	.	.	1.756	-0.0027	16.8	0.727	0.860

As can be seen from Table 2, the doubly constrained spatial interaction models with the complete system of constraints are, in general, more accurate in replicating migrations than both types of the singly constrained spatial interaction models. In addition, the predictions based on the attraction constrained models V-VIII are slightly more accurate than those generated by the corresponding production constrained models I-IV. A possible reason for the superior performance of the attraction constrained models is that the population size of a district is probably a more adequate measure of propulsiveness than of attractiveness. As pointed out by Fotheringham and O'Kelly (1989, p. 58), many other social, cultural and economic variables determine the attractiveness of a region for migration. Such variables do not seem to be as important in determining the number of people leaving particular regions.

Another interesting point is that modifications of the singly constrained spatial interaction models by introducing some additional parameters, attached to variables V_i or W_j, are relevant only in the case of the production constrained spatial interaction models, where they moderately improve the model performance. Remember that population size was used as composite measure to represent the attractiveness of the destination region as a recipient of inmigrations. Consequently, the values of 1.249 and 1.267 attached to the variable W_j in models II and IV respectively indicate increasing inmigration rates for districts with larger populations. On the contrary, the addition of the parameter μ to the attraction constrained spatial interaction models adds little or nothing to the performance level of the models. These results again confirm the above mentioned assumption that destination attractiveness appears to be a more complicated migration characteristic than origin propulsiveness.

As regards the form of the distance function utilized, it is evident that the negative power function provides in any case a better performance than the negative exponential function. From the behavioural point of view it means that the migrants perceive the distance logarithmically in evaluating alternative destination choices (cf. Wilson 1970, p. 35). Furthermore, it is clear that the use of Tanner distance function in the doubly constrained spatial interaction model does not bring any significant effect in comparison with the negative power distance function. In addition, it should be noted that the interpretation of two parameters in Tanner function is not obvious (see the sign of the parameter γ). Consequently, the doubly constrained spatial interaction model X with the negative power distance function is preferred as the most accurate model for replicating the observed migrations between Slovak districts. For this reason much of the subsequent discussion will be devoted to this type of spatial interaction model.

The most general way of identifying the sources of error in model predictions is to compare observed and predicted migration totals for different distance categories. In Table 3 the predicted totals were computed using the doubly constrained spatial interaction model with the power function. Data from this table clearly show that the model underpredicts both the number of short distance (1 - 100 kilometres) and long distance (greater than 250 kilometres) migrations, whereas medium-distance migrations (101 - 250 kilometres) are overpredicted. Note that similar patterns of under- and overprediction were generated also by two other doubly constrained spatial interaction models (models IX and XI) but the differences between the observed and predicted migration totals were substantially larger in the case of the model with the exponential distance function.

Tab. 3. Observed and predicted migration totals according to distance moved, 1993-95

Distance of move (in kilometres)	Observed		Predicted by model X	
	total number	percentage	total number	percentage
1-50	24136	26.80	22674	25.18
51-100	34501	38.31	34385	38.18
101-150	11777	13.08	14290	15.87
151-200	7067	7.85	7548	8.38
201-250	4750	5.27	4792	5.32
251-300	2447	2.72	2288	2.54
301-350	2157	2.39	1904	2.11
351-400	1062	1.18	845	0.94
401-450	1733	1.92	1113	1.24
451-500	433	0.48	224	0.25

The last problem to be discussed is the question: how do the calibration results change in time? In brief, when the results for the 1993-95 period are compared with those obtained for the other three periods considered, the above mentioned generalizations are fully confirmed. Nevertheless, the selected results of the calibration of the doubly constrained spatial interaction model with the power distance function and some other data of interest are presented in Table 4. The second column contains the total number of migrations between districts in the given time interval. As shown, this total was steady declining throughout the whole period under investigation. The third and fourth columns contain the overall mean length of migration (measured in kilometres), $\bar{c}$, and the distance decay parameter, γ, respectively. It can be seen that with the exception of the 1987-89 period the mean length of migration has increased over time. As expected, the value of parameter γ shows a pattern of decline as the mean migration length increases. These results indicate that in Slovakia over the period under investigation, at the interdistrict scale, and on the average, fewer and fewer persons are migrating over greater and greater distances. In spite of all these changes the performance of the model remains rather stabilized, as indicated by all three goodness-of-fit statistics.

Tab. 4. The results of the calibration of model X for the four periods investigated

Period	Total number of migrations	$\bar{c}$	γ	IOD	SRMSE	R^2
1984-86	122 725	103.4	1.5315	16.6	0.703	0.886
1987-89	114 008	102.1	1.5402	16.9	0.720	0.869
1990-92	103 407	103.8	1.5075	16.7	0.717	0.865
1993-95	90 063	107.0	1.4501	17.0	0.726	0.861

Spatial interaction models with zone-specific distance decay parameters

The distance decay parameters calibrated in the previous section are generalized interregional parameters indicating the overall propensity to migrate over distance. It is clear, however, that the propensity to migrate over distance is likely to vary spatially so that each generalized parameter value can mask spatial variation that may occur between different origins or destinations. As already noted, the extent of this variation can be identified by the calibration of spatial interaction models which incorporate zone-specific decay parameters. The second reason for the use of the zone-specific parameters is that this form of parameter disaggregation can improve the overall performance of the model.

Table 5 contains the calibration results for three alternative versions of the doubly constrained spatial interaction model with the power distance function fitted to data for the 1993-95 period. The values of three goodness-of-fit statistics for both origin- and destination-specific parameter calibration are presented and compared with those for the generalized parameter calibration of the same model. As shown, the overall percentage of migrations misallocated by the model falls from 17.0 to 15.3 when origin-specific parameters are introduced, and to 14.0 when destination-specific parameters are utilized. A similar pattern of change in two other statistics can be observed. The SRMSE falls from 0.726 to 0.579, whereas the R^2 coefficient rises from 0.861 to 0.908. Thus, it is obvious that introduction of the zone-specific parameters improves the model performance, particularly when the destination-specific parameter calibration is used.

Tab. 5. Goodness-of-fit-statistics associated with alternative doubly constrained spatial interaction models, 1993-95

Model	Distance decay parameter(s)	IDO	SRMSE	R^2
X	generalized	17.0	0.726	0.861
XII	origin-specific	15.3	0.644	0.888
XIII	destination-specific	14.0	0.579	0.908

Regional variation in origin-specific parameter values is shown in Figure 2. The values range from 1.16 for the district of Trnava to 2.31 for the district of Komárno, with the mean value of 1.55. The variation is very small, the coefficient of variation being only 18.3 %. It is not surprising to find that outmigrations from some peripheral districts (for example, the districts of Komárno, Nové Zámky, Levice, Veľký Krtíš, Stará Ľubovňa and Čadca) are associated with high parameter values indicating low propensity to migrate over distance, whereas outmigrations from some centralized districts with low parameter values are less impeded by the distance moved. In fact, the pattern of variation in parameter values is rather less predictable. It is worth mentioning, for instance, that outmigrants from the majority of the East-Slovak districts appear to be unaffected by the distance over which they move, whereas outmigrations from some more accessible West-Slovak districts (for example, the districts of Trenčín and Topoľčany) have relatively high parameter values. Note also

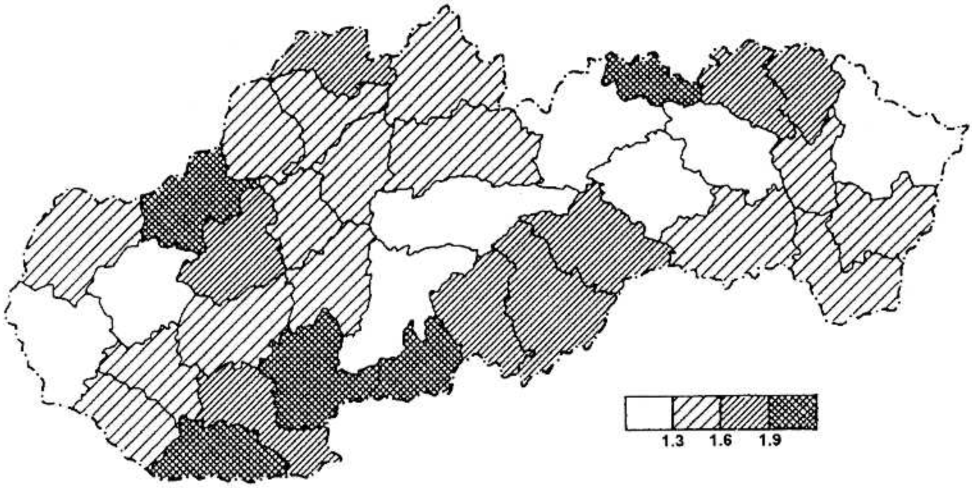

Fig. 2. Spatial variation of the origin-specific distance decay parameters, 1993-95.

that of the eight districts associated with the lowest parameter values, only the districts of Banská Bystrica and Zvolen can be characterized as ones with the "central" location.

A similar pattern of regional variation is displayed in Figure 3, which depicts the set of parameter values resulting from the destination-specific parameter calibration of the model. The parameter values vary from 1.07 (the district of Trnava again) to 2.41 (the district of Stará Ľubovňa) with the mean value of 1.58 and the coefficient of variation of 21.2 %. The close similarity of Figures 2 and 3 is confirmed by the correlation coefficient of 0.931, indicating a high degree of association between origin and destination parameter sets. The destination parameter values are again

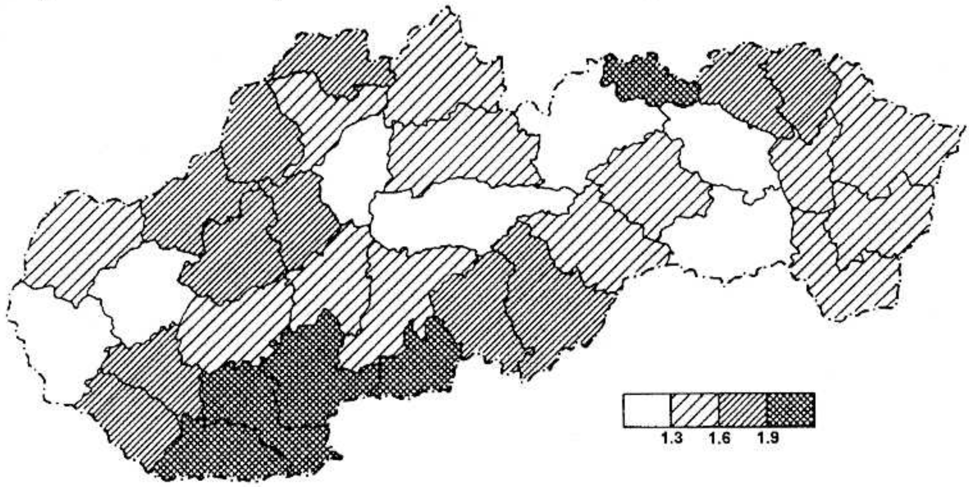

Fig. 3. Spatial variation of the destination-specific distance decay parameters, 1993-95.

higher for some peripheral districts located along both the northern and especially the southern boundary of Slovakia. But the same is true, surprisingly, also for the four West-Slovak districts of Trenčín, Topoľčany, Považská Bystrica and Prievidza with a good accessibility level. On the contrary, inmigrants to the majority of the districts in the eastern part of Slovakia are characterized again by a high propensity to migrate over distance. On the other side, the districts with the low parameter values are scattered over the whole country. It is interesting, however, that the lowest parameter values are associated with the districts centred on the largest Slovak cities (namely, the districts of Bratislava, Trnava, Martin, Banská Bystrica, Poprad, Košice and Prešov), which appear to attract inmigrants who tend to be less affected than most by the distance moved.

As indicated in Figures 2 and 3, the pattern of variation in both origin- and destination-specific parameter values is so complex that it cannot be explained only by the influences of relative location or accessibility of the districts. Obviously many other factors can play significant roles in this respect. Two of them have been mentioned: first, the "size" of the district measured by the population of its centre, and second, the high propensity to migrate associated with migrants from or to the East-Slovak districts. In addition, it is worth commenting the high parameter values for the districts in southern part of Slovakia with the mixed ethnic structure of the population. These values can be probably explained by the low propensity of the Hungarian subpopulation to migrate over long distances. Evidently further investigation of the spatial variation in specific parameter values is needed before definitive conclusions can be reached.

In order to examine changes in the parameter value variation over time, the doubly constrained spatial interaction model with the power distance function was fitted to data for the 1984-86 period and both origin- and destination-specific parameters were calibrated. The results obtained correspond closely with those derived for the 1993-95 period. It should be noted that the correlation coefficients between the two periods for both origin- and destination-specific parameter sets were 0.951 and 0.933 respectively. Thus, the regional variation in the propensity to migrate seems to be stabilized over time.

CONCLUSIONS

In this study various alternative models of spatial interaction were examined to demonstrate their performance for exploring interregional migration in Slovakia. The results of the analysis lead to the following conclusions.

First, the doubly constrained spatial interaction models are more accurate in replicating migrations than both types of singly constrained spatial interaction models, whereas the predictions based on the attraction constrained models are more accurate than those generated by the production constrained models. Moreover, the modifications of the singly constrained spatial interaction models by introducing the additional parameters, attached to variables representing the origin propulsiveness or destination attractiveness, are relevant only in the case of the production constrained spatial interaction models.

Second, spatial interaction models incorporating the negative power distance function were found to be preferable to models with the negative exponential distance function. Although the Tanner distance function may also be appropriate, it can be concluded that the doubly constrained spatial interaction model with the negative power distance function most accurately replicates the observed migrations between Slovak districts. It should be noted, however, that the degree of correspondence between the predicted and observed migrations is not very high and substantial errors occur in model predictions.

Third, the calibration of the doubly constrained spatial interaction model with zone-specific parameters improves substantially the goodness-of-fit of the model, especially when the destination-specific parameters are introduced. It is interesting, however, that the variation both in origin- and destination-specific parameter values is small. In addition, there is a little evidence to suggest that the spatial pattern of both parameter sets can be explained only by regional assessibility.

Finally, it is worthwhile commenting on the temporal stability of the results obtained. At the overall level, the performance of the models is rather stable during the whole period under investigation. On the other side, the general trend of decline in the frictional effect of distance is confirmed. In spite of this, the regional variation in the propensity to migrate over distance semms to be stabilized in time.

Concluding this paper one point have to be stressed. It is well-known that the overall performance of spatial interaction models in general and estimated distance decay parameters in particular are dependent upon the relative size of the interaction matrices, presence or absence of intraregional flows and level of spatial aggregation. For this reason, no comparison to other similar studies was given in this paper. It can be shown, however, that the results obtained are not principally different from those reported in some papers quoted in the introduction. Nevertheless, further research on spatial interaction modelling of migration in Slovakia is evidently needed.

ACKNOWLEDGEMENT

The author is grateful to M. Žirko from the Statistical Office of the Slovak Republik for providing unpublished migration data used in this study and to Martin C. Styan for correcting the original English version of the paper.

The research reported in this paper was supported, in part, by the Scientific Agency of the Ministry of Education of the Slovak Republik and the Slovak Academy of Sciences under grant 2/7501/20.

REFERENCES

BATTY, M., MACKIE, S. (1972). The calibration of gravity, entropy, and related models of spatial interaction. *Environment and Planning*, 4, 205-233.

BAXTER, M. (1983). Estimation and inference in spatial interaction models. *Progress in Human Geography,* 7, 40-59.

DULEY, C., RESS, P. (1991). Incorporating migration into simulation models. In Stillwell, J., Congdon, P., eds. *Migration models: macro and micro approaches*. London (Belhaven Press), pp. 228-261.

FOTHERINGHAM, A. S. (1986). Modelling hierarchical destination choice. *Environment and Planning A*, 18, 401-418.

FOTHERINGHAM, A. S., O'KELLY, M. E. (1989). *Spatial interaction models: formulations and applications*. Dordrecht (Kluwer).

ISHIKAWA, Y. (1987). An empirical study of the competing destinations model using Japanese interaction data. *Environment and Planning A*, 19, 1359-1373.

KNUDSEN, D. C., FOTHERINGHAM, A. S. (1986). Matrix comparison, goodness-of-fit, and spatial interaction modeling. *International Regional Science Review*, 10, 127-147.

KONTRA, P., ed. (1995). *Podrobný autoatlas - Slovenská republika, 1:100 000*. Harmanec (VKÚ).

LEDENT, J. (1985). The doubly constrained model of spatial interaction: a more general formulation. *Environment and Planning A*, 17, 253-262.

OPENSHAW, S. (1976). An empirical study of some spatial interaction models. *Environment and Planning A*, 8, 23-41.

PAULOV, J. (1991). The zone-size-dependent entropy formula and spatial interaction modelling: a note on some implications. *Environment and Planning A*, 23, 557-570.

PAULOV, J. (1993). *Entropia v urbánnej a regionálnej analýze: konceptuálny rámec a základy aplikácie (Entropy in urban and regional analysis: conceptual framework and the basis of an application)*. Geographia Slovaca, 2. Bratislava (Geografický ústav SAV).

PAULOV, J. (1996). Towards comparison of two entropy formulae. *Wirtschaftsuniversität Wien, Institut für Wirtschafts- und Sozialgeographie, WSG-Discussion Paper*, 51.

PAULOV, J., POLÁČIK, Š. (1979). Kalibrácia a testovanie gravitačného modelu na migračných dátach (Calibration and testing the gravity model on migration data). *Acta Facultatis Rerum Naturalium Universitatis Comenianae, Geographica*, 17, 209-232.

PLANE, D. A. (1982). An information theoretic approach to the estimation of migration flows. *Journal of Regional Science*, 22, 441-456.

PLANE, D. A. (1984). Migration space: doubly constrained gravity model mapping of relative interstate separation. *Annals of the Association of American Geographers*, 74, 244-256.

POOLER, J. (1987). Modelling interprovincial migration using entropy-maximizing methods. *Canadian Geographer*, 31, 57-64.

POOLER, J. (1993). Structural spatial interaction. *Professional Geographer*, 45, 297-305.

REES, P. (1994). Estimating and projecting the populations of urban communities. *Environment and Planning A*, 26, 1671-1697.

SANDRAPS, S. (1995). Influence de facteurs sociaux et géographiques sur l'interaction spatiale. *L'Espace Géographique*, 24, 11-24.

SEN, A. (1986). Maximum likelihood estimation of gravity model parameters. *Journal of Regional Science*, 26, 461-474.

STILLWELL, J. C. H. (1978). Interzonal migration: some historical tests of spatial interaction models. *Environment and Planning A*, 10, 1187-1200.

STILLWELL, J. (1986). The analysis and projection of interregional migration in the United Kingdom. In Woods, R., Rees, P., eds. *Population structures and models: developments in spatial demography*. London (Allen and Unwin), pp. 160-202.

STILLWELL, J. (1991). Spatial interaction models and the propensity to migrate over distance. In Stillwell, J., Congdon, P., eds. *Migration models: macro and micro approaches*. London (Belhaven Press), pp. 34-56.

TOBLER, W. R. (1983). An alternative formulation for spatial interaction modeling. *Environment and Planning A*, 15, 693-707.

WILLEKENS, F. J. (1983). Specification and calibration of spatial interaction models: a contigency-table perspective and an application to intra-urban migration in Rotterdam. *Tijdschrift voor Economische en Sociale Geografie*, 74, 239-252.
WILSON, A. G. (1967). A statistical theory of spatial distribution models. *Transportation Research*, 1, 253-269.
WILSON, A. G. (1970). *Entropy in urban and regional modelling*. London (Pion).
WILSON, A. G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning*, 3, 1-32.

Anton B e z á k

INTERREGIONÁLNE MIGRÁCIE NA SLOVENSKU: NIEKOĽKO TESTOV PRIESTOROVO-INTERAKČNÝCH MODELOV

Modely priestorovej interakcie sa všeobecne pokladajú za jeden z najvýznamnejších a najúčinnejších nástrojov urbánnej a regionálnej analýzy. V tejto štúdii sme sa prostredníctvom testovania predikčných vlastností niekoľkých interakčných modelov pokúsili demonštrovať možnosti ich aplikácie pri analýze interregionálnych migrácií na Slovensku.

Na testovanie sme vybrali 13 alternatívnych modelov, označených rímskymi číslicami I až XIII. Ich rovnice (vrátane rovníc pre korešpondujúce vyrovnávacie koeficienty) sú uvedené v tabuľke 1. Prvé štyri modely patria k typu východiskovo ohraničených a ďalšie štyri k typu cieľovo ohraničených interakčných modelov. V rámci oboch skupín sa rozlišujú dva subtypy modelov s exponenciálnou a dva subtypy modelov s mocninovou funkciou vzdialenosti. Pri každom subtype sa navyše uvažuje model s dodatočným parametrom, ktorý je priradený buď k miere atraktivity cieľového regiónu, alebo k miere propulzivity východiskového regiónu. Ďalej nasledujú tri obojstranne ohraničené interakčné modely, ktoré sa navzájom odlišujú tým, že prvý z nich obsahuje exponenciálnu, druhý mocninovú a tretí Tannerovu funkciu vzdialenosti. Posledná dvojica predstavuje osobitnú kategóriu interakčných modelov s regionálne špecifikovanými parametrami vzdialenosti. V prvom prípade ide o model s východiskovo špecifikovanými, v druhom s cieľovo špecifikovanými parametrami.

Parametre jednotlivých modelov sme odhadli metódou maximálnej vierohodnosti, pričom na riešenie vierohodnostných rovníc sme použili Newtonovu-Raphsonovu iteračnú metódu ako súčasť špeciálneho programu v jazyku Visual Basic 6.0. Osobitný postup, opísaný v anglickom texte, si vyžiadala kalibrácia interakčných modelov s regionálne špecifikovanými parametrami. Na zhodnotenie účinnosti každého modelu sme použili tri štatistiky zhody, ktoré merajú stupeň súladu medzi pozorovanými a modelom predikovanými migračnými tokmi. Sú to: štandardizovaná stredná kvadratická chyba, index nepodobnosti a koeficient determinácie.

Všetky uvažované modely sa aplikovali na dáta o migračných tokoch medzi starými okresmi na Slovensku v období 1984-1995. V záujme regionálnej komparability sa mestské okresy Bratislava a Košice spojili s korešpondujúcimi vidieckymi okresmi do spoločných "metropolitných" okresov. Migračné toky vo vnútri okresov sa nebrali do úvahy a dáta za jednotlivé roky sa zlúčili do štyroch trojročných období 1984-1986, 1987-1989, 1990-1992 a 1993-1995. Za mieru propulzivity východiskového regiónu a súčasne aj za mieru atraktivity cieľového regiónu sme zvolili počet obyvateľov príslušného okresu na začiatku každého trojročného obdobia. Vzdialenosti medzi okresmi sme definovali ako cestné vzdialenosti medzi okresnými mestami.

Výsledky získané kalibráciou interakčných modelov sú zhrnuté v tabuľkách 2 až 5. Z uvedených dát vyplýva niekoľko závažných poznatkov.

Obojstranne ohraničené interakčné modely reprodukujú pozorované migračné toky presnejšie ako obidva typy jednostranne ohraničených interakčných modelov. Účinnosť cieľovo ohraničených modelov je pritom o poznanie vyššia ako účinnosť východiskovo ohraničených modelov. Rozšírenie jednostranne ohraničených interakčných modelov o dodatočný parameter α, resp. μ má význam iba v prípade mier atraktivity vo východiskovo ohraničených interakčných modeloch, kde prispieva k miernemu zvýšeniu účinnosti modelu.

Pokiaľ ide o tvar funkcie vzdialenosti, je zrejmé, že interakčné modely s mocninovou funkciou vzdialenosti sú podstatne účinnejšie ako korešpondujúce interakčné modely s exponenciálnou funkciou vzdialenosti. Keďže zavedenie Tannerovej funkcie do obojstranne ohraničeného modelu nemá v porovnaní s mocninovou funkciou žiadny podstatný efekt, možno konštatovať, že obojstranne ohraničený interakčný model s mocninovou funkciou vzdialenosti najlepšie reprodukuje migračné toky medzi slovenskými okresmi v pozorovanom období. Podotýkame však, že stupeň zhody pozorovaných a predikovaných tokov nie je veľmi vysoký a v predikciách sa vyskytujú značné chyby.

Účinnosť najlepšieho modelu sa dá zvýšiť kalibráciou modelu s regionálne špecifikovanými parametrami vzdialenosti. Osobitne to platí pre model s cieľovo špecifikovanými parametrami, ktorý nesprávne alokuje iba 14 % predikovaných tokov. Je zaujímavé, že priestorová variabilita hodnôt východiskovo i cieľovo špecifikovaných parametrov je veľmi malá. Navyše sa nedá predpokladať, že priestorová variabilita parametrov je spôsobená iba regionálnymi diferenciami v polohe a dostupnosti jednotlivých okresov.

Posledný poznatok sa týka časovej stability získaných výsledkov. Vo všeobecnosti je možné prehlásiť, že účinnosť všetkých modelov prejavuje počas celého skúmaného obdobia pomerne vysoký stupeň stability. Na druhej strane sa potvrdil pokles vplyvu vzdialenosti na migráciu, čo znamená, že migranti prekonávajú postupne väčšiu a väčšiu vzdialenosť. Napriek tomu sa zdá, že regionálne diferencie v migračnej propenzite sú z časového hľadiska veľmi stabilné.