SPATIAL AGGREGATION AND INTRANSITIVITY IN U.S. MIGRATION STREAMS

Michael F. Goodchild
and
Andrew W. Grandfield
Department of Geography
The University of Western Ontario
London, Canada

ABSTRACT

Earlier papers have drawn attention to the inconsistency between observed migration flows between U.S. states and the predictions of a broad class of spatial interaction models, in that observed flows show frequent intransitivities while the models predict perfect transitivity. The paper compares the properties of flows between the real states with those of county flows aggregated to alternative sets of regions generated using various stochastic processes and satisfying certain objective functions. Intransitivity is affected by the variance in the number of counties per state, but is otherwise largely independent of the particular boundaries chosen. The set of 'states' which minimizes intransitivity is found to consist of strongly heterogeneous regions which lack compactness.

INTRODUCTION

Smith and Clayton (1) showed that the flows predicted by a broad class of spatially separable interaction models must be transitive. More specifically, consider models of the form

$$x_{ij} = E_j A_j c_{ij} \qquad (1)$$

where x_{ij} is the flow between origin place i and destination place j, E_i is the propensity of the origin place to generate trips, A_j is the propensity of the destination to attract trips, c_{ij} is a measure associated with the pair (i, j) and $c_{ij} = c_{ji}$. Note that the set of origins is identical to the set of destinations and the matrix x_{ij} is therefore square. The general form includes both constrained and unconstrained versions of the gravity model.

The transitivity property requires that for all triples (i, j, k), if $x_{ij} > x_{ji}$ and $x_{jk} > x_{kj}$ then $x_{ik} > x_{ki}$. It thus provides a very simple test of the empirical validity of spatial interaction models. An analysis of U.S. state-to-state migration flows for 1935-1970 showed that intransitivity was very common, and a Null Hypothesis that observed migration flows were sampled from a transitive population was decisively rejected, implying a fundamental and qualitative inability of models of this type to predict migration flows.

Goodchild and Smith (2) tried to reconcile the observed disagreement between model and reality by proposing two alternative mechanisms by which intransitivity could result despite the transitivity of the model. First, intransitivity could be produced by aggregation, either spatially through the use of large heterogeneous reporting zones, temporally through the longitudinal accumulation of varying trip patterns, or vertically through the superimposition of different social and economic groups. Goodchild and Smith simulated a broad range of aggregation conditions and were unable to replicate the relative abundance of intransitivities found in the empirical data. Second, intransitivity could be the result of a real asymmetry in the c_{ij} term. The simulations were less conclusive in this case.

Smith and Slater (3) proposed that intransitivities could be modelled using a fundamentally different approach to spatial interaction modelling based on Tversky's elimination by aspects. The model which they developed had the desired properties but proved to be extremely difficult to calibrate (see also 4) for all but the smallest data sets.

All of the above research was carried out on state-to-state tables, which provide a very coarse view of migration patterns. The work to be described in this paper was prompted by the availability of a county-level table giving flows of migrants between all pairs of counties in the U.S. from 1965 to 1970 (see also 5). The data was further disaggregated by age and sex, but the present paper describes only the analyses of total flow.

If flow intransitivity is an indicator of the suitability of spatial interaction models, full transitivity is also indicative of a simple structure of flow asymmetry: it is possible to order the zones so that net flow is always in the direction predicted by that order. More specifically, let ρ_i denote the position of zone i in the ordering of zones. Then there exists an ordering of the n zones ρ_i, i = 1,n such that

$$\rho_i < \rho_j \text{ implies } x_{\rho_i\rho_j} > x_{\rho_j\rho_i} \text{ for all } i \neq j \quad (2)$$

The paper discusses an algorithm for aggregating counties into a given number of reporting units in order to minimize flow intransitivity, and discusses the properties of the zones so produced. The results are compared to those obtained for reporting units designed to satisfy a number of alternative criteria.

DATA SETS

The list of U.S. counties is subject to some variation, so a standard set of 3073 was adopted for this study, representing the area of the coterminous 48 states. The set of states used included the District of Columbia, which was represented by one county, for a total of 49.

The analyses described in this paper are of the 36,109,810 migrants of all ages who moved across county lines between 1965 and 1970. The county table was aggregated to the 49 states and the resulting flows were found to contain 2151 intransitive triples. No attempt is made in this paper to distinguish between intransitivities on grounds of statistical significance, because of the methodological problems involved (1),(2).

RANDOM AGGREGATION

In this section we examine the Null Hypothesis that the observed level of intransitivity in interstate flows would result from random aggregation of counties into reporting units, using several alternative stochastic aggregation processes.

The number of ways in which 3073 units can be combined into 49 regions constrained by contiguity is enormous, and depends on the precise set of adjacencies among the basic units. Garfinckel and Nemhauser (6) and Goodchild and Hosage (7) have described methods for producing all possible aggregations of zones into a prescribed number of regions, but both methods are limited to small numbers of zones because of the explosive growth in the number of solutions. Rossiter and Johnston (8) and Minciardi et al (9) describe a random aggregation procedure, but it is clear that the method does not give equal likelihood to each possible aggregation. In the absence of any such ideal procedure, the method described here is a modification of Rossiter and Johnston's.

49 counties were selected randomly as cores. The algorithm then proceeded to examine each core in turn, and to add to it an adjacent unallocated county, randomly chosen from the list of adjacencies. After one cycle of the process each region would thus consist of two adjacent counties. In the last few cycles unallocated counties could only be added in the case of a few regions, leading to unequal numbers of counties per region in the final solution.

Nine different runs of the process were made, in each case generating a aystem of 49 regions. Flows aggregated to these regions gave a mean number of intransitivities of 1703, much less than the number for the real set of 49 states, and with a standard deviation of 152. The observed value of 2151 is thus 2.95 standard deviations above the mean. It appears that the real set of states produces significantly more intransitivities than a random aggregation of counties into regions. Nevertheless the observed number is much less than the expected value of 4606 (10, p. 157) in a random 49 by 49 tournament, in which the net flow between two states is independently assigned in either direction with equal likelihood.

Although the simulation produced the correct number of reporting units, with unequal numbers of counties in each, the variation in the number of counties per reporting unit was much less in the simulations than in reality. It is possible, then, that the difference in intransitivity was attributable to this rather than to any spatial structure existing in the real states. The random accretion process was therefore modified to produce a wider variance in counties per region. Instead of examining each region in turn in each cycle, the probability of a region being chosen for possible addition of an adjacent county was made dependent on the number of counties already allocated, as follows:

$$p_i = c_i^b / \sum_k c_k^b \quad (3)$$

where p_i is the probability of region i being chosen, c_i is the number of counties already allocated to region i and b is a parameter. The variance in counties per region in the final solution depends directly on b.

The results are shown in Figure 1. Contrary to the previous conclusion, it is evident that the real number of intransitivities is not significantly different from the numbers for simulations with similar variance in counties per region. The experiment implies that the boundaries of the existing states have no particular effect on the abundance of intransitivities relative to random reporting units. If the asymmetries of the county table were assigned randomly and independently, then we would expect any spatial aggregation of the county table to 49 groups to produce a random 49 by 49 tournament and thus 4606 intransitivities. The experimental aggregations produced many fewer because both the magnitudes of the county flows and the directions of asymmetry are strongly correlated.

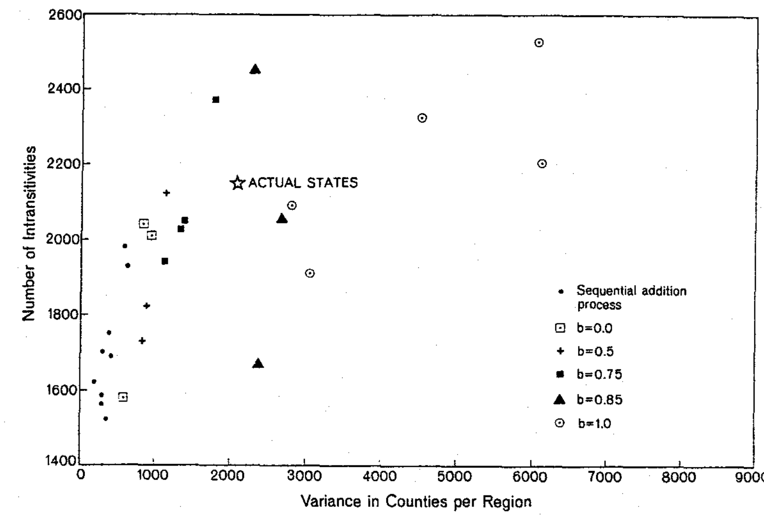

Figure 1 Number of intransitivities plotted against variance in counties per region for simulated sets of reporting units and for actual states.

CREATION OF MINIMUM-INTRANSITIVITY STATES

The design of reporting units clearly has a marked influence on asymmetry and intransitivity. In this section we examine the feasibility of redesigning state boundaries so as to produce fully transitive flows, which would be desirable on two counts: they would be consistent with a very simple model of the directions of asymmetry, and would not be qualitatively inconsistent with spatially separable interaction models.

An algorithm was devised to take the existing state configuration and attempt to reduce intransitivity in flows by swapping counties. To preserve contiguity, a county may be reassigned to another state only if it is adjacent to a county which is currently assigned to that state, in other words if part of the county's boundary is also part of the boundary of the state to which it will be assigned. It is also necessary to check that removal of the county will not split the state to which it currently belongs.

In each major cycle, the algorithm examined each of the 8557 county adjacencies in turn. An adjacency between counties belonging to two different states created two potential swaps, one for each county. Let S_i denote the state to which county i belongs, and let i, j be the pair of adjacent counties, $S_i \neq S_j$. The swaps would thus reallocate i from S_i to S_j and j from S_j to S_i. Each swap affects the flows in two rows and two columns of the state flow table; the first swap will in general decrease all inflows and outflows involving S_i and increase all flows to and from S_j, including the diagonal terms. The effect on intransitivity can be computed very efficiently by converting the state flow table into a (0, 1) tournament matrix and computing row sums (10), and fortunately only the rows and columns in the tournament matrix corresponding to S_i and S_j are affected by a swap.

The first pass of the algorithm produced 86 swaps, and reduced the number of intransitivities from 2151 to 1260. Only 23 swaps were made on the second, and after pass 5 no further changes occurred, with 1094 intransitivities left in the state flow matrix. At this stage the algorithm was modified to include a second objective, as follows. We noted earlier that the row sums of the state tournament matrix define an ordering of the states such that the direction of net flow is completely consistent with this ordering when flows are fully transitive. More specifically, define the tournament matrix $\underset{\sim}{T}$ such that

$$T_{pq} = 1 \text{ if } x_{pq} > x_{qp}, \text{ else } T_{pq} = 0 \quad (4)$$

where x_{pq} is the flow from state p to state q. Assign the value of 0 to the diagonal terms, $T_{pp} = 0$. Then the row sums $\sum_q T_{pq} = N_p$ will be a permutation of the integers 0 through $n - 1$.

Now consider the case where the flows are not fully transitive. The row sums will no longe be a permutation: there will be ties and missing integers. Some net flow directions will be consistent with the row sums, i.e.,

$$x_{pq} \geqslant x_{qp} \text{ when } N_p > N_q \quad (5)$$

but some will be inconsistent. In the original state flows there were 247 such inconsistent flows out of a total of 1176 state pairs. Define the total inconsistent flow E as the sum of the net flows between all pairs of states for which the direction is inconsistent, i.e.

$$E = \sum_{(p,q) \in I} (x_{qp} - x_{pq}) \quad (6)$$

where I is the set of ordered pairs (p, q) such that

$$x_{pq} < x_{qp} \text{ and } N_p > N_q \quad (7)$$

The total inconsistent flow in the original state table was 273,922. This is only 0.76% of the total flow and suggests that it is the relatively small flows which are responsible for intransitivity: the directions of the large flows in the system tend to be consistent with a simple ranking of states.

In the first 5 passes of the algorithm swaps were made only if they reduced total intransitivity. In subsequent passes a secondary objective of reducing total inconsistent flow was added to allow for the possibility that several swaps might be needed before a reduction in intransitivity could occur. Swaps were made if either intransitivity or total inconsistent flow was reduced, with the first condition taking precedence. Note that a swap which reduces intransitivity can also produce an increase in total inconsistent flow. Unfortunately it is necessary to examine the entire state flow matrix to evaluate the change in inconsistent flow which would be produced by every hypothetical swap.

19 passes were needed to produce convergence of the algorithm. The final solution contained 387 intransitivities and a total inconsistent flow of 85,298 or 0.24%, generated by 59 pairs of states. Subsequent passes produced no swaps according to either objective.

CHARACTERISTICS OF MINIMUM-INTRANSITIVITY STATES

The solution is shown in Figure 2. The reporting units clearly make little sense from a spatial point of view: they are extremely contorted so that the solution appears to be trying to maximize the length of state boundaries. Some regions are close to their original form, particularly 'Michigan', 'Maine' and 'Florida', because in these cases there are relatively few potential swaps. Others, such as those in the areas originally occupied by Oklahoma, Texas, Louisiana and Arkansas, have developed long interdigitated corridors. This suggests that the objective of minimal intransitivity is incompatible with the contiguity constraint: that the counties which would have to be aggregated to produce a fully transitive set of flows are not spatially clustered.

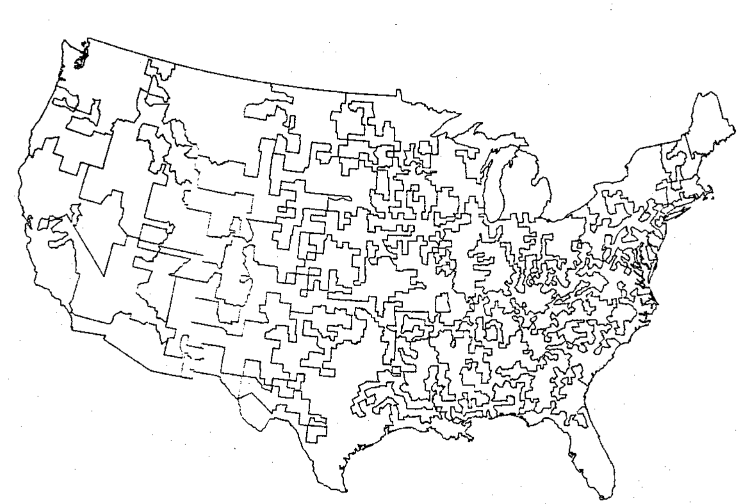

Figure 2 49 contiguous reporting units minimizing intransitivity and inconsistent flow.

CONCLUSIONS

The various pieces of evidence accumulated in this paper point to the conclusion that intransitivity is a fundamental property of migration flows, and not an artifact of the reporting units used. Randomly generated regions are found to show similar levels of intransitivity provided they are constrained to be contiguous and to have similar variance in the number of counties per region. Regions designed to show minimum intransitivity lack spatial compactness, suggesting that this objective is incompatible with contiguity constraints. This suggests strongly that the need in aggregate migration analysis is for new models which predict flows with an appropriate level of intransitivity. The goodness of fit of any model depends on the reporting regions used, but in this case manipulation of those regions is not capable of making spatially separable interaction models compatible with reality.

REFERENCES

1. Smith, T.R. and C. Clayton. Transitivity, spatially separable utility functions, and United States migration streams, 1935-1970. Environment and Planning A 10, 1978, 399-414.
2. Goodchild, M.F. and T.R. Smith. Intransitivity, the spatial interaction model, and U.S. migration streams. Environment and Planning A 12, 1980, 1131-1144.
3. Smith, T.R. and P.B. Slater. Analysis of U.S. interstate labor flows using interaction models with choice sets constrained by information flow. International Regional Science Review 6, 1981, 15-31.
4. Puchalski, M. Calibrating and evaluating choice set constrained spatial interaction models. Unpublished M.A. dissertation, Department of Geography, University of Western Ontario, 1983.
5. Slater, P.B. Measuring migration fields of U.S. counties. Geographical Analysis 16, 1984, 65-73.
6. Garfinckel, R.S. and G.L. Nemhauser. Optimal political districting by implicit enumeration techniques. Management Science, 16, 1970, B495-B508.
7. Goodchild, M.F. and C.M. Hosage. On enumerating all feasible solutions to polygon aggregation problems. Modelling and Simulation 14 (Proceedings of the Fourteenth Annual Pittsburg conference), 1983, 591-595.
8. Rossiter, D.J. and R.J. Johnston. Program GROUP: the identification of all possible solutions to a constituency-delimitation problem. Environment and Planning A 13, 1981, 231-238.
9. Minciardi, R., Puliafito, P.P. and R. Zoppoli. A districting procedure for social organizations. European Journal of Operational Research 8, 1981, 47-57.
10. Kendall, M.G. Rank Correlation Methods, 4th Edition, Griffin, London, 1970.