COLLEGE STUDENT MOBILITY IN THE UNITED STATES IN FALL 1968: GRAPH-THEORETIC STRUCTURING OF TRANSACTION FLOWS

Paul B. Slater, West Virginia University

1. Introduction

In 1960 I. Richard Savage, a statistician, and Karl W. Deutsch, a political scientist, published a study of the 1928 trade flows among North Atlantic nations (Savage and Deutsch, 1960). They employed an origin-destination independence model to control for the substantial differences among countries in the sizes of their total imports and exports. Deviations -- measured by relative acceptance coefficients -- from this indifference or null model were interpreted as indicators of the magnitudes of trade linkages. It was suggested that these values be employed as dependent variables in additional analyses which would seek to explain variation in trading ties on the basis of distance, political relations, etc. Subsequently, Leo A. Goodman, a statistician and sociologist, proposed that the model should be estimated in such a manner that zero values would be assigned to flows among nations with no trading relations (Goodman, 1963, 1964). Due to the difficulties of obtaining data on trade within countries and developing a theory to explain such commerce, Savage and Deutsch had themselves nullified the diagonals of their flow table and indifference model.

A wide variety of transaction flow data -- for example, diplomatic and tourist exchanges, international mail deliveries, and newspaper and periodical circulations -- have been analyzed by these methods. Most of the applications and discussions of the null model have been reported in the political science literature (Lijphart, 1964; Puchala, 1970; Clark and Welch, 1972; Hughes, 1972; Chadwick and Deutsch, 1973; Clark, 1973; Mann and Chadwick, 1973; Seligson, 1973), while a few have appeared in geographical sources (Soja, 1968; Smith, 1970a, b; Williams and Zelinsky, 1970; Britton, 1971; Bradford, 1973).

The major innovation and methodological development in this area since the work of Savage, Deutsch, and Goodman has been accomplished by a political scientist, Steven J. Brams (1966a, 1966b, 1968a, 1968b, 1969). He used graph theoretic algorithms to cluster and hierarchically structure political units on the basis of salient linkages; that is, relative acceptance coefficients greater than some specified threshold level. [Nystuen and Dacey (1961), Filani (1972), Rouget (1972), Campbell (1972, 1974), and Taliaferro and Remmers (1973) also employed graph theory to study transaction flows. They did not, however, focus on deviations from an origin-destination independence model as Slater (1974) did in studying interindustrial flow tables. Fisher (1969) employed a "lockstep progressive merger procedure", while Ghosh (1970) and Roy (1971) minimized a contiguity index in structuring interindustry tables.]

A 51 x 51 matrix classifying the entire Fall 1968 American college student population by state (and District of Columbia) of residence and enrollment is the subject of the transaction flow analysis reported here. These data were compiled through the intensive efforts of the National Center for Educational Statistics (Wade, 1970, 1971). 2,445 of the 2,495 institutions of higher learning to which questionnaires were sent responded. The residence-enrollment patterns of the 50 non-responding institutions were imputed on the basis of a 1963 survey. Information was collected and published on a more disaggregated basis (sex, type of college, and degree program) than that studied here. Data for individual institutions have not been released. Post-1968 surveys have not been conducted and are not, at present, planned due to both the expense involved and recent legal decisions which have allowed many students to become in-state residents and thus avoid discriminatory fees levied against nonresident students.

The approach of Brams is paralleled to a considerable extent in this investigation of these flow data. However, in addition to an origin-destination independence model, models of symmetry and quasi-symmetry are also estimated. (Non-diagonal cell estimates in these two models are independent of the omission or inclusion of the diagonal entries.) Departures from these two models reflect imbalances of flows between pairs of states. The quasi-symmetry analysis controls for the fact that total outflows from states are unequal to total inflows, while the symmetry study does not. Residuals -- not relative acceptance coefficients -- are utilized in all three analyses, as measures of deviation, due to their greater resistance to fluctuations in the ratios of small observed to expected values. Questions concerning the distribution of the sizes of "consignments" do not arise here, as they can in trade flow studies (Savage and Deutsch, 1960; Goodman, 1963). Unlike each monetary unit expended in trade, each individual student is an essentially independent unit.

2. A Quasi-Independence Model

The multiplicative model

$$x_{ij} = r_i c_j, \tag{1}$$

where x_{ij} is the number of Fall 1968 American students resident in state i enrolled in all institutions of higher learning in state j was estimated. The diagonal entries, x_{ii}, are typically omitted from transaction flow analyses. Their values are often unavailable or, if available, of dominant magnitudes that strongly perturb estimation procedures. Since a line (loop) from a point to itself is not an admitted concept in the theory of directed graphs, the diagonal entries are, in any case, irrelevant in directed graph (digraph) studies. Iterative techniques which are necessary for the estimation of (1), if any of the ij-cells are disregarded (Goodman, 1968), were employed in this study due to the omission of the diagonal cells. (The extents to which students attend colleges in their states of residence are not, therefore, revealed in the results presented here.) (1) is, in such cases, referred to as a quasi-independence model.

One of several asymptotically equivalent formulas for the residuals from this model

$$(x_{ij} - r_i c_j)/(r_i c_j)^{\frac{1}{2}} \tag{2}$$

is employed as a measure of deviation, not relative acceptance coefficients

$$(x_{ij} - r_i c_j)/r_i c_j \tag{3}$$

nor absolute deviations

$$x_{ij} - r_i c_j \tag{4}$$

[The solution of a set of 102 (51 + 51) normal equations could have been used to adjust the standardized residuals (2) for their asymptotic variances (Haberman, 1973). However, considerable cost would have been involved in computing the adjustments. In addition, modifications would have been small, since no state had more than thirteen percent of all the Fall 1968 out-of-state students resident or enrolled in it. For exploratory purposes, therefore, the standardized residuals appeared adequate.]

Residuals incorporate a square root scale. They are, consequently, in a sense intermediate between relative acceptance coefficients (logarithmic scale) and absolute deviations (original scale). Brams could conceivably have employed residuals in his studies. He expressed dissatisfaction with the exclusive usage of either relative acceptance coefficients or absolute deviations, since the former were biased against nations with large trade flows, while the latter, on the other hand, gave undue importance to relatively small changes in large trade flows. Therefore, his standards for classifying linkages as salient required that both these measures had to be greater than certain levels (Brams, 1966a).

There are 2,550 (51 x 51 - 51) residuals from the quasi-independence model (1). A salient linkage criterion which sets all residuals greater than some chosen threshold to 1 and all other residuals to 0 can, of course, be established in many ways. It would be of interest to vary thresholds and observe how interrelationships revealed by the application of graph-theoretic procedures are altered (Gotlieb and Kumar, 1968; Doreian, 1969; Auguston and Minker, 1970; Osteen, 1974). The stability of relationships could be evaluated by doing so, as in clustering by fuzzy sets (Zadeh, 1965).

After preliminary investigation of the matrix of residuals, it was decided to set the 150 largest residuals to 1 and the remaining 2,400 to 0. The 151st largest residual, 46.3, was found with the use of a partial sorting algorithm (Chambers, 1971). If only a few order statistics are required, this method yields the desired information at less expense than if complete sorts are performed. This choice was made to strike a balance between richness of detail, significance of conclusions, and ease of computation and presentation. The results obtained are displayed in Figure 1. (The states of Alaska and Hawaii are represented by circles. Readers unfamiliar with the names and locations of the states should refer to a map which yields such information.) If the ij-entry of the Boolean adjacency matrix obtained by use of 46.3 as a threshold is 1, an arrow is directed from state i to state j. This indicates that state i sent substantially more students to state j than would be expected if the null model held. There is, of course, extensive departure from (1) because factors such as distance, tuition differentials, and quality of educational facilities are not accounted for in it. A summary measure of the degree to which it fails to hold is given by the percentage discrepancy statistic

$$(\sum_{ij}\sum 50|x_{ij} - r_i c_j|)/\sum\sum x_{ij} \text{ for all } x_{ij},\ i \neq j \tag{5}$$

The numerical value of this statistic is the percentage of the 1,104,622 students attending out-of-state colleges in the Fall of 1968 that would have had to have been redistributed in order to achieve quasi-independence (Brams, 1966a). It was computed to be 40.7. Since the maximum possible value is 50, state of residence and state of enrollment are strongly associated. Most of the 150 salient linkages are, in fact, between contiguous pairs of states, while the remainder occur between relatively proximate pairs. (It should be realized that there are certainly substantial within-state effects present in the data. Residents of one section of a state might have different preferences for out-of-state colleges than those of another.)

By exponentiating a 51 x 51 Boolean matrix with 201 non-zero entries (150 for the saliences plus 51 on the main diagonal), the reachability matrix of the digraph formed by the 150 links can be obtained (Harary et al., 1965). [The power to which it is necessary to raise the matrix is termed the diameter of the graph. Algorithms for the multiplication of sparse Boolean matrices were employed for this purpose (IBM, 1970). These were much more efficient than conventional matrix multiplication routines would have been. Graph algorithmic programming languages that are also of value in computing reachability matrices have been developed (Rheinboldt et al., 1972).] One state (A) is reachable from another (B) if a sequence of directed links exists from B to A. Strong components -- which may be interpreted as regions -- of the graph can be easily determined from the reachability matrix. These are maximal sets, any member of which is reachable from any other.

The graph formed by the 150 salient linkages was found to possess five strong components. Two of them were sets consisting of single states -- Alaska and Montana. The other three were non-trivial. One was a western region composed of Arizona, California, Hawaii, Idaho, Nevada, Oregon, Utah, and Washington. Another was a northeastern region comprised of Connecticut, Delaware, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The fifth region was a large diffuse one comprised of the remaining thirty (midwestern and southeastern) states plus the District of Columbia. Virginia, a member of this region, was reachable from another member, Colorado, in a minimal path of nine links. (The diameter of the graph was thirteen.)

If the threshold for saliences were raised, more regions (strong components) with fewer states would result. By lowering the threshold, on the other hand, coarser regional distinctions would appear. When, during preliminary investigations, the largest 330 residuals were deemed salient, every state was reachable from every other one. Therefore, there was only one strong component or region, the entire nation.

One of the most interesting aspects of regionalizing in this fashion is isolating the links, if any, between the strong components. Since an acyclic digraph is obtained by condensing a digraph with respect to its strong components, no two strong components can be mutually salient. For example, Montana and Alaska -- both strong components -- are each salient to the same member of the western region -- Washington. However, no member of the western region is salient to either of these two states, thus preventing their incorporation into it. Wyoming, a constituent of the largest region, is salient to its neighboring states of Montana and Utah, but neither Montana nor any western state is salient to Wyoming. Pennsylvania, which is in the northeastern region, is salient to two states that it borders, Ohio and West Virginia. However, the only states salient to Pennsylvania are other northeastern ones -- Delaware, New Jersey, and New York.

Strong components may contain cliques; that is, maximal complete subgraphs. Every member of a clique is salient to every other member. There are five three-state cliques in the 150-link digraph. These are {Idaho, Oregon, Washington}, {California, Oregon, Washington}, {Connecticut, Massachusetts, New York}, {Alabama, Mississippi, Tennessee}, and {Alabama, Georgia, Tennessee}. Clearly, cliques need not be mutually exclusive, in the manner of strong components. [Gotlieb and Kumar (1968) combine overlapping cliques to obtain "diffuse" classes.] Therefore, the latter are more appropriate for regionalization purposes. [Taliaferro (1973) regionalized the counties of Kansas on the basis of weak components; that is, maximal sets of counties all mutually reachable by sequences of undirected links. The weak components of a digraph are the strong components of the digraph's symmetric closure (Harary et al., 1965; Harary and Miller, 1970).]

3. A Symmetry Model

A quasi-independence model can be estimated for any rectangular table of counts. The study of square tables of counts -- where row and column classifications coincide -- is richer, however, in that models of symmetry and quasi-symmetry can also be estimated (Ireland et al., 1969). Under an hypothesis of symmetry, the estimate of the ij or ji cell of a square table is simply

$$(x_{ij} + x_{ji}) / 2 \tag{6}$$

These values were computed for the Fall 1968 student flow table. The associated 2,550 residuals were then calculated, and the 150 largest were again deemed salient. (The 151st largest residual was 12.1. 21.8 percent of the students would have had to have been redistributed for symmetry to hold.) These saliences are exhibited in Figure 2. An arrow directed from state i to state j indicates that state i sent substantially more students to state j than it received from state j. For several reasons, however, Figure 2 would be unacceptably complex if arrows were used to represent all 150 saliences. One reason is that contiguity is not as prominent a factor as in Figure 1, while another is that there can be no two-headed arrows in Figure 2. Also, a very high density of emitted saliences was found in the neighboring states of Connecticut, New Jersey, New York, and Pennsylvania. Therefore, symbols are used to represent most of the saliences directed from these four states.

In addition to transmitting to 32 states plus the District of Columbia, the matrix of residuals revealed that all of the other fifty units received more students from New Jersey than it sent there. Even to New York, which itself transmitted to 29 units (including Pennsylvania, which transmitted to 19), New Jersey sent 24,534 students, while receiving only 6,689 in return. Overall, 16,831 out-of-state students were enrolled in New Jersey's institutions of higher education, while 116,536 New Jerseyites were attending colleges outside their home state.

Another interesting aspect of Figure 2 is the large number of saliences incident to Utah. This phenomenon is probably in large part attributable to Mormon students enrolling in Brigham Young University. The catalog for this school, like many others, lists the states and foreign areas from which the students come. As of August 18, 1969, 9,513 of 27,277 students were from Utah. 16,348 were from the other 49 states plus the District of Columbia. In the eastern portion of the nation, Tennessee has a dominant role similar to Utah's in the west. Tennessee and Utah -- in strong contrast to New Jersey -- have large net inflows of students. (Tennessee received 31,763 and sent 12,973, while the comparable data for Utah were 19,108 and 3,047.) The quasi-symmetry model controls for such differences between inflows and outflows.

Several of the one-way linkages under the quasi-independence model (Figure 1) are shown in Figure 2 to also be strongly asymmetrical. (For example, Wyoming → Utah; Montana → Washington; Florida → Tennessee.) Also, some two-way linkages exhibit the same property. (Illinois → Indiana; Florida → Georgia; Idaho → Utah, for instance.) As an illustration of the magnitude of the 150 asymmetries, a few -- not selected on any systematic basis -- are listed.

Ill.	sent	12,052	to Wisc.	and	recd.	2,428.
Iowa	"	2,428	" Neb.	"	"	725.
Maryland	"	1,903	" N. C.	"	"	193.
Pa.	"	6,080	" W. Va.	"	"	428.
New York	"	1,083	" Arizona	"	"	109.
Florida	"	997	" Miss.	"	"	154.

There are no cycles (intransitivities) in Figure 2. If one unit can be reached by a series of directed links from another unit, there is, therefore, no return path. Hence, there are no non-trivial strong components. The points in such an acyclic digraph can be hierarchically ordered so that a point is assigned to a kth level if, and only if, the longest path to the point is of length k (Brams, 1968a, b, 1969; Rouget, 1972). A point at any level of the hierarchy is either salient to a point at the next higher level or has a salience from a point at the immediately preceding level. Table 1 presents the hierarchical arrangement -- based on seven levels -- for the digraph of Figure 2. (Since Arkansas and North Dakota are isolated, that is, have no saliences associated with them, they are omitted from the table.) No state is salient to New Jersey (6), while Maine (0), for instance, is not salient to any other state. Maryland (3) is salient to Ohio (2), which in turn is salient to Kentucky (1). Two states, not at adjacent levels, can be salient. However, the salience must be directed from the higher level to the lower level.

4. A Quasi-Symmetry Model

The marginal (row and column) sums of the cell estimates in the quasi-independence model (1) and the sums of symmetrically located pairs of cell estimates in the symmetry model (6) are identical to the corresponding sums in the flow table. By iteratively constraining a 51 x 51 table, the entries of which are all initially equal, to possess first one and then the other of these two properties, convergence to estimates under an hypothesis of quasi-symmetry can be obtained. Cell estimates in such a model are the closest possible -- in the sense of minimum discrimination information -- to being symmetric, given that the marginal sums are inhomogeneous; that is, the sums of rows and columns corresponding to the same classifications are unequal (Ireland et al., 1969). The 150 largest residuals from the quasi-symmetry model are presented in Figure 3. The 151st largest residual was 6.9. 11.5 percent of the students would have had to have been redistributed for quasi-symmetry to hold. Since symmetry is equivalent to quasi-symmetry plus marginal homogeneity, the difference, 10.3, of the two percentage discrepancy statistics, 21.8 and 11.5, is the percentage of departure from (6) attributable to marginal inhomogeneities.

New Jersey is salient to nine states in Figure 3, a considerable decline from the 33 saliences in Figure 2. By controlling for the large disparity between the number of students leaving New Jersey and the number entering, most of the asymmetries under (6) were diminished. Similarly, the saliences associated with Connecticut were reduced from 11 to 4. (20,884 out-of-state residents were enrolled in Connecticut's colleges, while 42,965 Connecticut residents attended college outside the state.) Interestingly, Virginia, also a substantial net exporter of students, has a salience from Connecticut under the hypothesis of quasi-symmetry, but not symmetry.

On the other hand, the number of saliences associated with Pennsylvania increased from 19 to 22. (Pennsylvania was a relatively weak net exporter of students, receiving 62,302 and sending 76,420.) New York's saliences decreased in number from 29 to 27. (New York received 69,424 students and sent 135,981.) The number of saliences from Ohio increased from 6 to 11. (Ohio was a net importer, receiving 57,428 and sending 50,215.) Florida is shown in Figure 3 to be a strong attracter of students from Northeastern states, possibly for climatic reasons. (Climatic differences may also explain Minnesota's salience to California.) California and Hawaii have exceptional appeal for Florida residents. The number of saliences to Utah decreased, while those to Tennessee and Texas increased.

Perhaps the most surprising salience revealed in Figure 3 is that of Oregon to Pennsylvania. Its presence results in the digraph associated with Figure 3 being cyclic, in contrast to the acyclic digraph corresponding to Figure 2. (For instance, Pennsylvania → California, California → Oregon, Oregon → Pennsylvania. This is not, however, the maximal cycle containing the Oregon → Pennsylvania link.) In the Fall of 1968, Pennsylvania sent 82 students to Oregon, while Oregon sent 225 to Pennsylvania. Oregon received roughly a quarter more students than it sent, while Pennsylvania sent approximately a quarter more than it accepted from out-of-state.

5. Conclusions

Substantial departures from quasi-independence, symmetry, and quasi-symmetry models of the Fall 1968 state-to-state table of college students occur. By concentrating on the largest residuals from these models, with the use of graph theoretic procedures, dominant patterns of deviations can be perceived. The procedures employed are of general value in analyzing the rich body of transaction flow data.

ACKNOWLEDGEMENT

The data analyzed were brought to the author's attention by William H. Miernyk.

REFERENCES

Auguston, J. G. and Minker, J., 1970, "An Analysis of Some Graph Theoretical Cluster Techniques," Journal of the Association for Computing Machinery, 17, 571-588.

Bradford, B. C., 1973, "On the Structure of Soviet Urban Migration," Geography Department, University of Wisconsin, Madison, Wisconsin, U.S.A.

Brams, S. J., 1966a, "Transaction Flows in the International System," American Political Science Review, 60,880-898.

Brams, S. J., 1966b, "Trade in the North Atlantic Area: An Approach to the Analysis of Transformations in a System," Peace Research Society Papers, 6, 143-164.

Brams, S. J., 1968a, "Measuring the Concentration of Power in Political Systems," American Political Science Review, 62, 461-475.

Brams, S. J., 1968b, "DECOMP: A Computer Program for the Condensation of a Directed Graph and the Hierarchical Ordering of its Strong Components," Behavioral Science, 13, 344-345.

Britton, J. N. H., 1971, "Methodology in Flow Analysis," East Lakes Geographer, 7, 22-36.

Campbell, J., 1972, "Growth Pole Theory, Digraph Analysis and Interindustry Relationships," Tijdschrift voor Economische en Sociale Geographie, 63, 79-87.

Campbell, J., 1974, "Selected Aspects of the Interindustry Structure of the State of Washington, 1967," Economic Geography, 50, 35-46.

Chadwick, R. W. and Deutsch, K. W., 1973, "International Trade and Economic Integration: Further Developments in Trade Matrix Analysis," Comparative Political Studies, 6, 84-110.

Chambers, J. M., 1971, "Algorithm 410 Partial Sorting (ML)," Communications of the Association for Computing Machinery, 14, 357-358.

Clark, C., 1973, "The Impact of Size on the Savage-Deutsch RA Statistic: A Research Note," Comparative Political Studies, 6, 109-122.

Doreian, P., 1969, "A Note on the Detection of Cliques in Valued Graphs," Sociometry, 32, 237-242.

Filani, M. O., 1972, "Changing Patterns of Central Places and Functional Regions: Temporal and Spatial Dynamics of Air Traffic Flows in the United States," Ph. D. Thesis, Pennsylvania State University, State College, Pennsylvania.

Fisher, W. D., 1969, Clustering and Aggregation in Economics (Baltimore: Johns Hopkins Press).

Ghosh, A. and Sakar, H., 1970, "An Input-Output Matrix as a Spatial Configuration," Economics of Planning, 10, 133-142.

Goodman, L. A., 1963, "Statistical Methods for the Preliminary Analysis of Transaction Flows," Econometrica, 31, 197-208.

Goodman, L. A., 1964, "A Short Computer Program for the Analysis of Transaction Flows," Behavioral Science, 9, 176-186.

Goodman, L. A., 1968, "The Analysis of Cross-Classified Data: Independence, Quasi-Independence and Interactions in Contingency Tables With or Without Missing Values," Journal of the American Statistical Association, 63, 1091-1131.

Gotlieb, C. C. and Kumar, S., 1968, "Semantic Clustering of Index Terms," Journal of the Association for Computing Machinery, 15, 493-513.

Haberman, S. J., 1973, "The Analysis of Residuals in Cross-Classified Tables," Biometrics, 29, 205-220.

Harary, F., Norman, R. Z., and Cartwright, D., 1965, Structural Models: An Introduction to the Theory of Directed Graphs (John Wiley: New York).

Harary, F. and Miller, H., 1970, "A Graph-Theoretic Approach to the Analysis of International Relations," Journal of Conflict Resolution, 14, 57-63.

Hughes, B. B., 1972, "Transaction Data and Analysis: In Search of Concepts," International Organization, 26, 659-680.

IBM Corporation, 1968, System 360 Matrix Language (MATLAN), Manual H 20-0479-1.

Ireland, C. T., Ku, H. H., and Kullback, S., 1969, "Symmetry and Marginal Homogeneity of an r x r Contingency Table," Journal of the American Statistical Association, 64, 1323-1341.

Lijphart, A., 1964, "Tourist Traffic and Integration Potential," Journal of Common Market Studies, 2, 251-262.

Mann, C. A. and Chadwick, R. W., 1973, "Ranull-Transaction Flow Analysis," Behavioral Science, 18, 72-73.

Nystuen, J. D. and Dacey, M. F., 1961, "A Graph Theory Interpretation of Nodal Regions," Papers and Proceedings of the Regional Science Association, 7, 29-42.

Osteen, R. E., 1974, "Clique Detection Algorithms Based on Line Addition and Line Removal," SIAM Journal of Applied Mathematics, 26, 126-135.

Puchala, D., 1970, "International Transactions and Regional Integration," International Organization, 24, 732-763.

Rheinboldt, W., Basili, V., and Meszteniji, C., 1972, "On A Programming Language for Graph Algorithms," BIT, 12, 220-241.

Rouget, B., 1972, "Graph Theory and Hierarchisation Models," Regional and Urban Economics, 2, 263-296.

Roy, M. K., 1971, "A Note on the Computation of an Optimal Ordering for an Input-Output Matrix," Economics of Planning, 11, 95-97.

Savage, I. R. and Deutsch, K. W., 1960, "A Statistical Model of the Gross Analysis of Transaction Flows," Econometrica, 28, 551-572.

Seligson, M. A., 1973, "Transactions and Community Formation: Fifteen Years of Growth and Stagnation in Central America," Journal of Common Market Studies, 11, 173-190.

Slater, P. B., 1974, "Exploratory Analyses of Interindustrial Transaction Tables," Environment and Planning, 6A (1974), 207-214.

Smith, R. H. T., 1970a, "Concepts and Methods in Commodity Flow Analysis," Economic Geography, 46, 404-416.

Smith, R. H. T., 1970b, "Inter-regional Trade as a Component of National Unity," in Selected Papers, 21st International Geographical Congress, Vol. 2, Ed. S. P. Chattenjee (National Committee for Geography, Calcutta), pp. 285-293.

Soja, E. W., 1968, "Communications and Territorial Integration in East Africa: An Introduction to Transaction Flow Analysis," East Lakes Geographer, 4, 39-57.

Taliaferro, J. D. and Remmers, W. W., 1973, "Identifying Integrated Regions for Health Care Delivery," Health Services, 18, 337-343.

Wade, G. H., 1970, Residence and Migration of College Students, Fall 1968: Basic State-to-State Matrix Tables (National Center for Educational Statistics, Washington).

Wade, G. H., 1971, Residence and Migration of College Students (National Center for Educational Statistics, Washington).

Williams, A. V. and Zelinsky, W., 1970, "On Some Patterns in International Tourist Flows," Economic Geography, 46, 549-567.

Zadeh, L. A., 1965, "Fuzzy Sets," Information and Control, 8, 338-353.

TABLE 1

Hierarchical ordering of states based on 150 largest residuals from symmetry model

Level	States
6	New Jersey
5	Connecticut, New York
4	Pennsylvania
3	Florida, Maryland
2	Alaska, Hawaii, Illinois, Montana, Ohio, South Carolina
1	Arizona, California, Delaware, Georgia, Idaho, Iowa, Kentucky, Nevada, New Mexico, Virginia, Washington, Wyoming
0	Alabama, Colorado, District of Columbia, Indiana, Kansas, Louisiana, Maine, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Nebraska, New Hampshire, North Carolina, Oklahoma, Oregon, Rhode Island, South Dakota, Tennessee, Texas, Utah, Vermont, West Virginia, Wisconsin

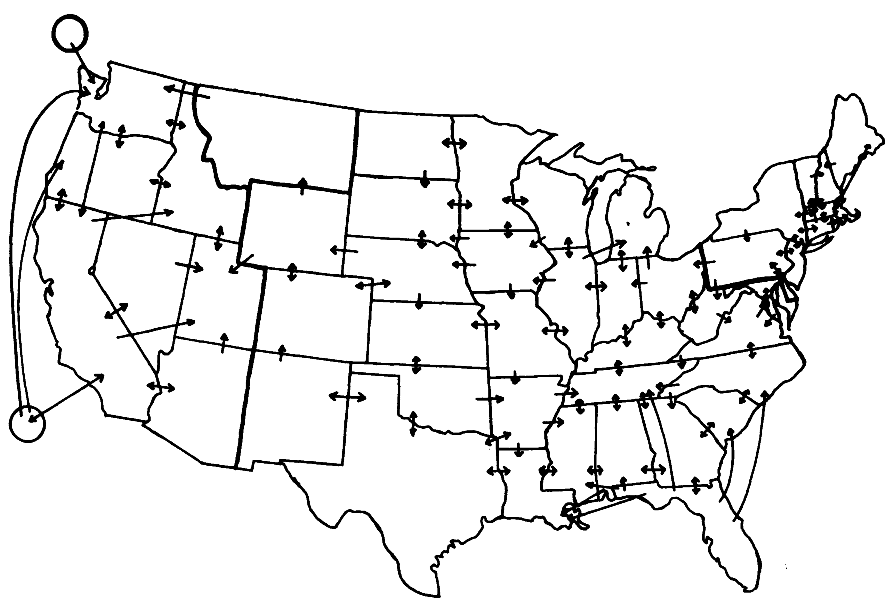

Figure 1. 150 largest residuals from quasi-independence model, and regionalization based upon strong components.

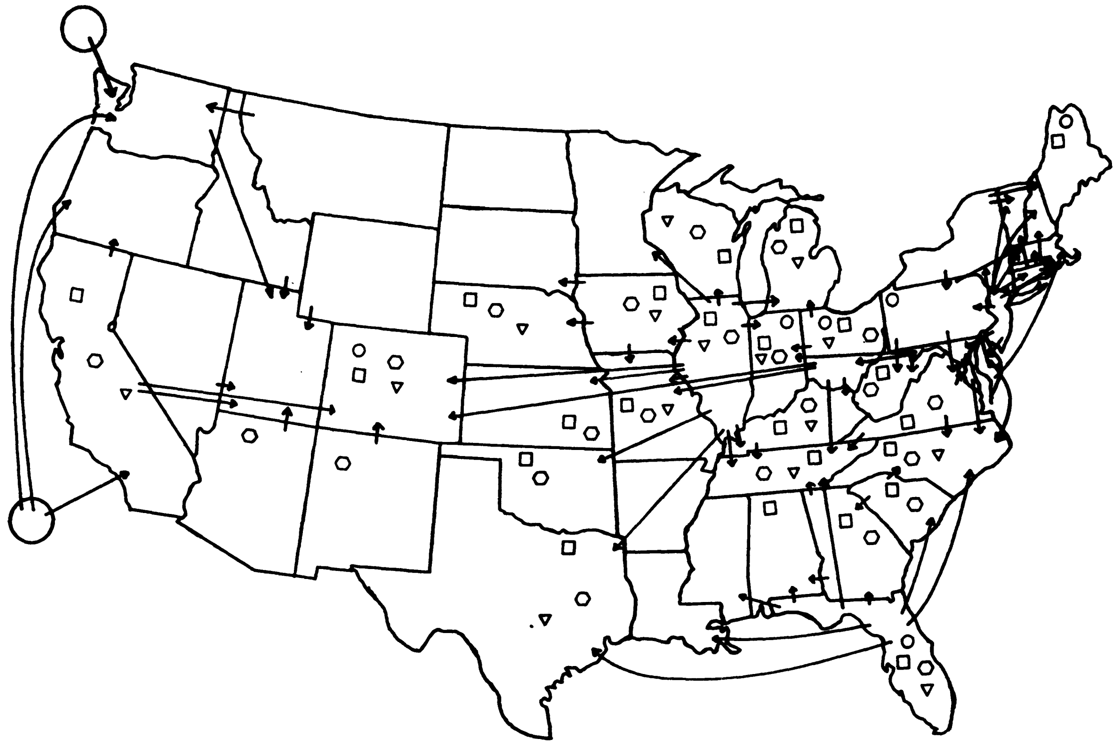

Figure 2. 150 largest residuals from symmetry model.

○ : ← Connecticut
□ : ← New Jersey
⬡ : ← New York
△ : ← Pennsylvania
District of Columbia ← {Connecticut, New Jersey, New York, Pennsylvania}

Figure 3. 150 largest residuals from quasi-symmetry model.

○ : ← Connecticut
□ : ← New Jersey
⬡ : ← New York
△ : ← Pennsylvania
▷ : ← Ohio
District of Columbia → {Illinois, North Carolina, Ohio}