

**BUILDING BETTER BLOCKMODELS:
A NON-HIERARCHICAL EXTENSION OF CONCOR
WITH APPLICATIONS TO REGRESSION ANALYSIS**

Steven R. Steiber
American Medical Association

Mid-American Review of Sociology, 1981, Vol. VI, No. 1:17-40

Cluster analysis has been used by many as a tool for identifying relatively small numbers of categories which represent the salient contrasts within bodies of data. One method which has received widespread use among sociologists in general, and sociometrists in particular, is CONCOR. This paper presents a non-hierarchical extension of the CONCOR algorithm and demonstrates the advantages of this extension with three exemplary bodies of data. One of the bodies of data is then used as a vehicle for expanding on the information which is customarily derived from cluster, or blockmodel, analyses. Specifically, a means by which the results of blockmodeling may be translated into a regression model is described and demonstrated using these network data.

One of the fundamental concerns of researchers in sociology specifically, and all sciences in general, is how to distill into the most parsimonious package possible the maximum information available in a body of data. To this end, blockmodeling and assorted clustering techniques have been used in condensing the contrasts or similarities among members of larger collectives into their smaller salient components (clusters). Applied to social network data, blockmodeling efforts such as those of Breiger et al. (1975), White et al. (1976), Boorman and White (1976), Arabie et al. (1978), and Snyder and Kick (1979) represent useful empirical examples which identify relatively homogenous subgroups, cliques, or global structural positions.

The algorithm on which the above clustering efforts, as well as others, have been based is referred to by the acronym CONCOR. CONCOR provides a means by which the similarity among cases

within a larger collective may be used to successively divide that collective into smaller elements which reflect some presumed group structure. As a divisive hierarchical method, however, CONCOR is unable to subsequently reassign members of clusters without creating new clusters. This is not a major shortcoming when applied to smaller data sets and it may be managed, in part, by creating progressively more clusters which provide finer distinctions. It more severely misrepresents the categoric distinctions, however, among larger bodies of data. Paradoxically, it is among these larger data bases that a clustering algorithm can be most usefully applied.

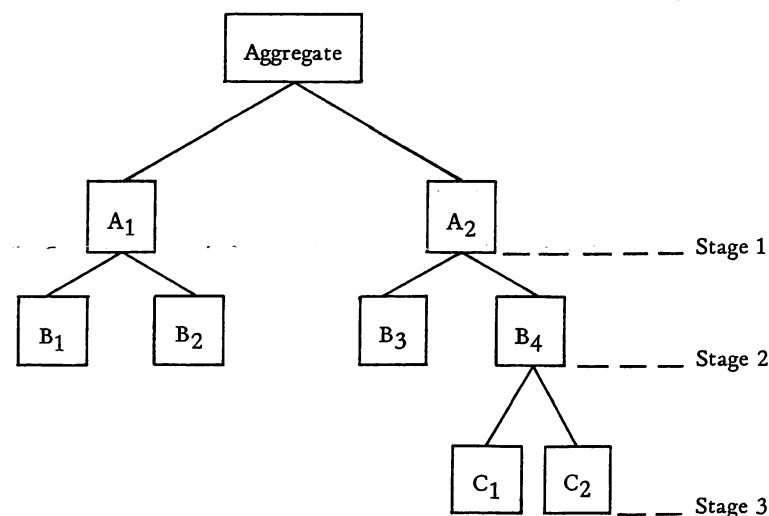
This paper provides an extension and refinement of CONCOR's basic criterion for defining blocks of like units within larger aggregations. The method (referred to here as RECOR) non-hierarchically elevates the empirical similarity of units within clusters. With a prescribed starting point, RECOR successively relocates units with others with which they are more similar. It ultimately, then, offers a solution where all data points are grouped with those with whom they are more similar. Given the constraint imposed by the predetermined starting number of clusters, RECOR offers, of course, locally maximal similarity within clusters.

This reclustering method is first statistically compared with CONCOR and then applied to three bodies of data: Sampson's (1969) Monastery data; Homan's (1950) Bank-Wiring Room data; and Snyder and Kick's (1979) World Structure Data. In each case RECOR provides an increment in the average similarity of units within clusters (measured by a Pearson product-moment correlation) and the extent to which a "lean fit" to the data is achieved (see Breiger et al. [1975:333] for a discussion of "lean fit"). Additionally, following Snyder and Kick's (1979) application of CONCOR results to an analysis of covariance framework, this paper presents an extension of clustering results to regression models which account for both positions and linkages. These regression models have been referred to as models of "structural relations" (Steiber, 1980). Implications for block-modeling in particular and social network analysis in general are also offered.

THE ALGORITHMS: CONCOR AND RECOR

The means by which CONCOR arrives at clusters which contain elements similar to one another is relatively straightforward. Only a very general description is offered here, as there are extensive technical discussions elsewhere in the literature (e.g., Breiger et al., 1975; Arabie et al., 1978; among others). Computing the Pearson product-moment correlations among all combinations of column vectors of stacked data matrices yields an $n \times n$ similarity matrix of each case with the remaining $n-1$ cases which provides a starting point for CONCOR. Through successive iterations, the division which distinguishes two elements of a larger whole above a prescribed level of correlation is derived. The underlying assumption in generating blockmodels in this manner is that if one identifies clusters of cases which are most consistent (correlated) with one another, then one has a very close approximation of the group structure in the network. This consistency criterion is a surrogate for the "connectivity" which has been the focus of early sociometrists (Arabie et al., 1978:37).

Figure 1. Representation of Hypothetical Clustering of CONCOR



completed. Clusters are subsequently reidentified to reflect additions to, or deletions from, their ranks. The matrix is then recomputed, and another iteration is begun. When there are no further membership changes, convergence is attained. In instances where they are few cases, convergence can be reached in a few iterations or less of RECOR (n 20); where one encounters appreciably larger samples (n 100), substantially more iterations are necessary.

Sampson's (1969) Monastery Data

Beginning with a social network of 18 individuals, data on positively and negatively affective relations among a cloister of monks described by Sampson (1969) are analyzed. These data reflect the sociometric evaluations of 18 monks in a cloister undergoing a great deal of turmoil during a twelve month period during which Sampson maintained detailed records of four types of social ties among members. These data have since been statistically clustered using the CONCOR algorithm by White et al. (1976) and also by Breiger et al. (1975) using a reduced body of data. These latter data are reanalyzed here.

Applying CONCOR to the Sampson data, Breiger et al. (1975) identified, through repeated application of the algorithm, a three-block model which seemed to offer the substantively-salient contrasts in the social network. Using Sampson's own labels of cliques within the monastery, the first block isolated by Breiger et al., was designated the Loyal Opposition joined by three individuals earlier labeled as "waverers." Those of the Loyal Opposition were seen as maintaining the status quo of the monastic order in spite of some differences of opinion. The second major cluster was termed by Sampson, the Young Turks, and these are wholly and exclusively found in the second cluster of Breiger's blockmodel. The Young Turks were those who were in constant conflict over monastery policy with both the Loyal Opposition and the senior monks (not represented in the data). Finally, the last class of individuals is called the Outcasts. These individuals Sampson characterized as not only withdrawn from

the mainstream of the monastery, but also receiving negative affective evaluations from the others of the cloister.

Table 1. Block Membership, Reclustered Sampson Data

Block	Monks
I	Monk 4, Monk 5, Monk 6, Monk 8, Monk 9, Monk 10, Monk 11 (N = 7)
II	Monk 1, Monk 2, Monk 12, Monk 14, Monk 15 (N = 5)
III	Monk 3, Monk 7, Monk 13, Monk 16, Monk 17, Monk 18 (N = 6)
(Total N = 18)	

Reclustering the monastery data offers convergence after a single iteration of RECOR. This has the effect of removing one individual from the Loyal Opposition and two from the Young Turks and placing all three in the Outcasts groups (the reclustered memberships may be found in Table 1). The former member of the Loyal Opposition was one of the "interstitial members" of the cloister first identified by Sampson, thereby relegating him to a conceivably ambiguous position. The change in this individual's membership is in keeping with Sampson's original analysis as well as the other clustering efforts using the full data set (White et al. 1976:753). The change in the cluster assignment of monks 7 and 16 is also in keeping with actual events as they were recorded by Sampson. After the expulsion of four monks from the monastery for their "problems" with the senior staff, these two left of their own volition within a week (Sampson, 1969:373ff).

Figure 2. Block Densities and Binary Images, Reclustered Sampson Monastery Data

Positive Affect:			Negative Affect:				
	I	II	III		I	II	III
I	.43	.11	.02	I	.00	.14	.40
II	.06	.60	.07	II	.23	.10	.30
III	.00	.17	.27	III	.38	.07	.20
		1 0 0			0 0 1		
		0 1 0			1 0 1		
		0 0 1			1 0 1		
(Mean Density = .19)			(Mean Density = .20)				

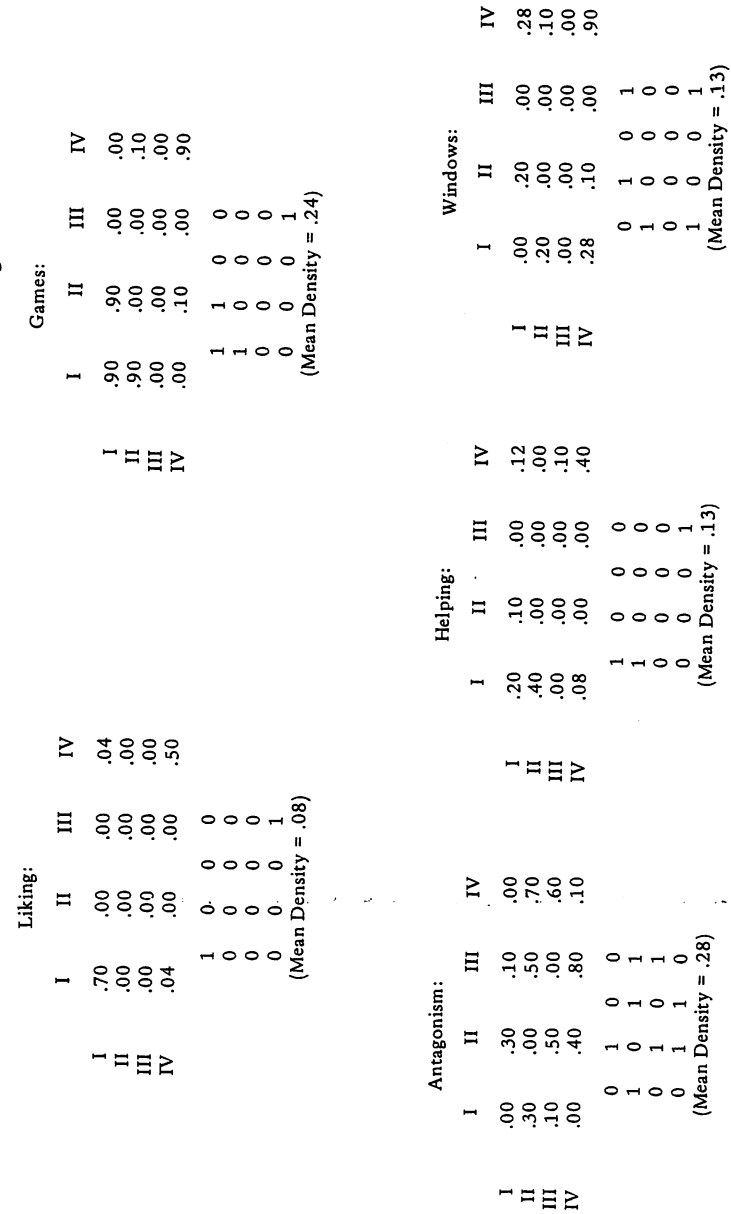
Empirically, the reclustering of the monastery data also offers an improvement in the "leanness" of the fit of the original blockmodel configuration (Breiger et al., 1975:355). Though the number of zeroblocks (a "lean fit" criterion) in any permutation of this data matrix will be admittedly small because there are so many ties among members of the monastery, Breiger finds only one. The application of the non-hierarchical RECOR, however, offers two (see Figure 2).³ Similarly, the average correlation of each individual vector with those of others of his own cluster (a social distance criterion) is increased from .150 to .162, a factor increase of 1.08. This is a relatively small increment in the case of the correlation coefficient and a small absolute increment in the number of zeroblocks, but they are both mutually consistent. They also are consistent with subsequent findings reported for other data sets in this analysis.⁴

Homans' (1950) Bank-Wiring Room Data

Among the fourteen individuals in the Bank-Wiring Room, Homans originally identified two cliques of five and four members. In addition, there were five workers whom Homans claimed were either outsiders of who had only slight orientations to the cliques (Homans, 1975:70ff). Subsequent quantitative analysis by Breiger et al. (1975) found two, four, and nine blocks of individuals within the social network of the Wiring Room. Though the nine-block model offers finer distinctions more sensitive to the subtle associational differences in the Wiring Room, the four-block model more closely preserves Homans' original interpretation. Indeed, the four-block model provides the basis for Breiger's discussion. The average correlation between the vector describing each individual and the actors with whom he is clustered in the four-block model is .377. A non-hierarchical reclustering of the data, however, shows that this is not an optimum empirical blocking of the data either in terms of correlations of individual vectors with members of blocks or the total zeroblocks afforded by the clustering of the data.

After two iteration of RECOR, the average correlation on convergence between each individual vector and the others of the

Figure 3. Block Densities and Binary Images, Reclustered Homan's Bank Wiring Room Data



respective clusters is increased ($r = .401$), representing a factor increment of 1.06. The number of zeroblocks in this reclustered model is also augmented by this application of RECOR. Whereas the original number of zeroblocks in the Breiger et al. (1975:346) model is 37, the number associated with this reclustered model is 45, a factor increase of 1.22 (see Figure 3).

Empirical considerations, though supportive of the efficiency of a method, do not provide the only basis for the preference of one algorithm over another. In a more substantive vein, the results using RECOR are in closer accord with the patterns originally generalized from the Bank-Wiring Room by Homans. The source of the difference between the earlier clustering effort of Breiger et al. (1975) and this present effort is in the assignment of "hangers-on." Specifically, Homans first identified men W5, S2, and I3 as non-participants with either of the two larger cliques (W1 through I1 or W7 through S4); rather, they were outliers ("hangers-on").

Table 2. Block Memberships, Reclustered Homans Data

Block	Workers
I	Wireman 1, Wireman 3, Wireman 4, Solderman 1, Inspector 1 (N = 5)
II	Wireman 2, Wireman 5 (N = 2)
III	Solderman 2, Inspector 3 (N = 2)
IV	Wireman 6, Wireman 7, Wireman 8, Wireman 9, Solderman 4 (N = 5)
(Total N = 14)	

In neither of the earlier CONCOR assignment nor in this reclustered model were the "hangers-on" clustered with one another in a single unique clique. There were, however, assigned to two cliques of "hangers-on" by RECOR as displayed in Table 2; I3 and S2 were associated in one distinct clique and W5 was linked with individual W2 in another. This latter individual, W2, was earlier identified by Homans (1950:70) as entering "little into conversation with" the first clique and "had little to do with it"—an outlier. As outliers, both individuals I3 and S2 are joined

in a single unique block. Finally, the reassignment of W6 to the second large clique from his CONCOR assignment represents an empirical clarification of an ambiguous role. Indeed, Homans (1950:70) stated that W6 was "in many ways an outsider" to the second larger clique, but that his membership did appear to lie with that group more than any other. This is borne out here at least empirically.

Snyder and Kick's (1979) World Structure Data

As the above examples suggest, the degree to which a non-hierarchical clustering algorithm will offer different blockings than those of a hierarchical method is a function of both the number of variables available for each case and the total number of cases. For instance, only one iteration was necessary for convergence with 18 cases and two variables (Sampson's Monastery data), but two iterations were required before reaching convergence on a four variable, 14 case data base (Homans' Bank Wiring Room data). When either of these two parameters is substantially increased, one can expect that the number of iterations necessary before achieving convergence will be increased (though no precise function can be identified).

Using data reported by Snyder and Kick (1979) on the social, political, and economic interrelations among a sample of nations, RECOR is applied to a 118 case, four variable data set. In the original analysis, Snyder and Kick employed the CONCOR algorithm with their preferred blocking offering ten clusters of nations. These they identified as core nations (one block with such members as Canada, the United States, the United Kingdom, etc.), semiperipheral nations (three blocks with such nations as Ireland, the U.S.S.R., and Argentina as members), and peripheral nations (the remaining six blocks with most of Africa, Haiti, Mexico, and others). The four international ties used in the data were those identified by trading flows, military interventions, diplomats exchanged, and treaties binding nations.

The most general distinctions made across these tiers of a world-system were between the core and the non-core. In terms of the variables used in the analysis, for instance, the core was seen

to maintain "trade linkages to and from every other block in the system." The peripheral blocks, on the other hand, were "with minor exceptions integrated into the world economy only through their trade with the core block" (Snyder and Kick, 1978:1114). Additionally, the core block, though it "does not undertake military interventions into all other blocks, it policies much of the system." The core "clearly dominates the sending of information (diplomatic exchanges), but it does not receive diplomatic bonds from any other blocks" (Snyder and Kick, 1979:1115). Finally, though the core block maintains treaties with more nations than any other block, most treaties appear to bind blocks to their own membership or to other blocks geographically proximate to them.

In applying RECOR to the data, eleven iterations are required before there are no changes in block memberships of nations. Moreover, the total number of blocks is reduced from the original ten to nine.⁵ Direct comparison of the average correlations between individual vectors and clusters is denied by this reduction in the total clusters. Standardizing on cluster number, however, offers some basis for comparison. A simple formula for this comparison is:

$$F = \frac{(\bar{r}_i)/N_i}{(\bar{r}_j)/N_j}$$

where F represents the factor increase in average r's across block-models, \bar{r}_i is the average of the correlations between individual vectors and the members of the N_i clusters, and \bar{r}_j is the average Pearson correlation between individual vectors and the members of the N_j clusters. In the World Structure data, the \bar{r}_i (reclustered) is .552 and the \bar{r}_j (CONCOR) is .531 with a factor difference of 1.16. Using an analogous formula to compare zeroblocks for each blockmodel, the reclustered configuration offers 1.03 times the zeroblocks of the blockmodel derived by Snyder and Kick (1979).

Binary images for the nine-block model are presented in Figure 4. Identifying the first block (row 1) as Core, the next

Figure 4. Binary Images of Nine-Block Model, Reclustered, Snyder and Kick World Structure Data

		Interventions:									Treaties:								
		I	II	III	IV	V	VI	VII	VIII	IX	I	II	III	IV	V	VI	VII	VIII	IX
Trade:	I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	II	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	III	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	IV	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	V	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
	VI	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
	VII	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
	VIII	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
	IX	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
		(Mean Density = .34)									(Mean Density = .13)								
		Diplomats:									Treaties:								
		I	II	III	IV	V	VI	VII	VIII	IX	I	II	III	IV	V	VI	VII	VIII	IX
Diplomats:	I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	II	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	III	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	IV	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	V	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
	VI	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
	VII	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
	VIII	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
	IX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
		(Mean Density = .26)									(Mean Density = .13)								

three rows (2, 3, and 4) as semi-peripheral, and the remaining five rows as peripheral reveals much of the same relationships as Snyder and Kick's ten-block model. The relationships are simply more consistent, as witnessed by the "leaner fit" in the reclustered configurations.

Table 3. Block Memberships Reclustered, Snyder and Kick

Block	Nations
I	United States, Canada, United Kingdom, The Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, Federal Republic of Germany, Austria, Italy, Yugoslavia, Greece, Sweden, Norway, Denmark, Turkey, Japan, India, Australia (N = 22)
II	Dominican Republic, Mexico, El Salvador, Nicaragua, Costa Rica, Panama, Venezuela, Ecuador, Peru, Brazil, Chile, Argentina, Uruguay (N = 13)
III	Cuba, Ireland, German Democratic Republic, Poland, Hungary, Czechoslovakia, Albania, Bulgaria, Romania, Soviet Union, Finland, Kenya, Israel, People's Republic of China, Pakistan (N = 15)
IV	Cyprus, South Africa, Iran, Taiwan, Republic of Korea, Burma, Sri Lanka, Republic of Vietnam, Malaysia, Philippines, Indonesia (N = 11)
V	Haiti, Guatemala, Honduras, Colombia, Bolivia, Paraguay (N = 6)
VI	Jamaica, Trinidad and Tobago, Malta, Iceland, Mali Republic, Mauritania, Somalia, Mongolia, People's Democratic Republic of Korea, Nepal, Thailand, Cambodia, Laos, New Zealand (N = 14)
VII	Afghanistan, Democratic Republic of Vietnam (N = 2)
VIII	Iraq, Syria, Lebanon, Jordan, Saudi Arabia, Yemen, Kuwait (N = 7)
IX	Senegal, Dahomey, Niger, Ivory Coast, Guinea, Upper Volta, Sierra Leone, Ghana, Togo, Cameroon, Nigeria, Gabon, Central African Republic, Chad, Democratic Republic of Congo, Zaire, Uganda, Burundi, Rwanda, Ethiopia, Malagasy, Morocco, Algeria, Tunisia, Libya, Sudan, United Arab Republic, Liberia (N = 28)
(Total N = 118)	

For the most part, the clusters identified by RECOR give not only empirically "leaner fits" but also conceptually "cleaner" clusters. That is, they represent here more homogenous geopolitical units within a world-system (see Table 3). Block number 2, for instance, represents most of Latin America, block number 8 is a Middle Eastern cluster, and block number 9 is exclusively African in its membership.

In an empirical sense (and close to the theoretical one), centrality is the underlying dimension which dictates cluster memberships. Though not all nations of block number 1 may be considered core power, per se, they all maximally represent centrality within a world-system and have close ties to other central nations. Where nations do not manifest such extensive ties, there understandably exist regional concentrations of ties, and hence the memberships of nearly all the peripheral clusters and most of the semi-peripheral clusters. This centrality versus non-centrality within a global network, while not precisely mirroring the theoretical core/non-core distinction, can nonetheless prove useful in modeling relations within a world-system.

In addition to the identification of binary images and clusters of the international relations data, Snyder and Kick included an extension of their analysis to multiple regression.⁶ They found, as theory had suggested, that over a period of time, core nations grew more rapidly in GNP per capita than did non-core nations. In short, there was a relative price to be paid for non-core status and a relative benefit to be gained from core membership. Explicit international flows were added to this analysis of variance model with nominal insight into the means by which the relations in the data contributed to differences in gains in GNP. Distinguishing between the receiving and the launching of armed interventions showed the one significant effect (launching) aside from block membership. Sending nations manifested less economic growth, on average, than non-sending nations.

FROM BLOCKMODELS TO REGRESSION MODELS

More than just a simple presentation of the relative benefits and costs attached to specific structural positions in the world-system has been offered by Snyder and Kick (1979). "They also demonstrated the effects of specific transnational flows which had served to define the structure of the system. Their models, then, may be aptly described as models of 'structure and relations'" (Steiber, 1980a:14).

Using dummy-coded variables in an analysis of variance framework, blockmodel structure was translated into a regression

model of nine variables (plus controls). In subsequent presentations, the relational components could be seen in the incorporation of selected flows as ratio variables (Snyder and Kick, 1979: 1119). That only one of the three flows used manifested a significant effect, however, suggests that (1) the flows selected were not those of most importance in identifying position in the world system, (2) the effects of relational ties were already largely identified by structural variables, (via blockmodeling), or (3) enough of the information from the original matrices of flows was lost in this translation that no systematic effects could be detected. Though the first two may be inferred from the following presentation, the last is more explicitly explored here.

Table 4. Regression Estimates for Changes in GNP/Capita (1960-1970) for Blockmodels

Independent Variables	Partial Regression Coefficients (Standard Errors)		
	Ten-Block Model	Nine-Block Model	Three-Block Model ^a
GNP/Capita (1960) ^b	-.065 (.049)	-.053 (.048)	-.060 (.048)
Secondary School Enrollment Ratio (1960) ^c	6.23 (1.67)**	6.40 (1.62)**	6.28 (1.64)**
Block I	-397.2 (127.6)**	-427.7 (115.5)**	-361.7 (95.1)**
Block II	-463.7 (128.8)**	-182.7 (108.0)	-425.3 (103.0)**
Block III	-509.4 (151.5)**	-471.1 (120.5)**	
Block IV	-340.7 (113.4)**	-434.8 (151.3)**	
Block V	-527.2 (127.2)**	-367.8 (117.6)**	
Block VI	-516.0 (141.6)**	-414.4 (305.0)	
Block VII	-459.7 (141.2)**	-671.8 (136.0)**	
Block VIII	-360.0 (114.6)**	-360.4 (114.3)**	
Block IX	-731.2 (147.7)**		
Regression Constant	482.9	416.0	434.4
R ²	.541	.535	.470

(N = 114 Nations)

** Significant beyond .01.

^a In the Three-Block Model Block I is a dummy-code for the contrast between semi-nations and all others, and Block II represents the distinction between peripheral nations and all others.

^b Information on Gross National Product per capita is derived from I.B.R.D. (1972).

^c Information on secondary school enrollment is derived from U.N.E.S.C.O. (1968).

Of general theoretical importance in Snyder and Kick's analysis was not the distinctions across all categories of a ten-block model (or here a nine-block model) or even a basic core versus non-core dichotomy. There was an explicit concern for the possible distinctions to be made across the three tiers of a global hierarchy: core, semi-periphery, and periphery. In Table 4, three regression models which highlight these distinctions are presented. In the first column of the table are the regression coefficients associated with the ten-block model originally identified by Snyder and Kick (1979). The general patterns found in the early report were obtained by using the data here as well: all non-core nations increase their GNP per capita, on average, less than core nations during the ten year interval. There are some minor departures from the precise regression estimates derived by Snyder and Kick, but these in all probability owe to the use of unadjusted secondary school enrollment ratios (versus population adjusted ratios) and the larger number of nations for whom data are available (114 here versus 105).

In the second column of coefficients are the regression estimates associated with the nine-block model identified by RECOR. With one degree of freedom difference between the two models, there is no significant difference between the respective variances explained between the ten- and nine-block models ($F = 1.35$, $d.f. = 1, 102$). Finally, moving to the farthest right column of the table, the distinctions among core, semi-periphery, and periphery do not capture significantly less of the explained variance than the Snyder and Kick ten-block model ($F = 2.25$, $d.f. = 7, 102$).

Each of the models in Table 4 describe only the *structure* of the world-system. Even Snyder and Kick's models included, at best, structure and relations. What is present in blockmodels, but heretofore lost in most regression models, is a means for representing structure, relations, and *structural relations*.

Structural relations may be identified by "categorical variables which effectively identify the origins of flows or process," and "... include also in specific variables the destinations and types of flows" as well as potential interactions among additive effects (Steiber, 1980a:15). In the present context, a model of

structural relations should identify the *nature of specific* ties (e.g., treaties, trade, diplomats exchanged) as well as the points of *origin* and *destination* of those ties. In addition, since aggregate wealth of nations (measured by GNP/capita) is of central concern here, the relative advantage that such wealth may differentially offer nations with different ties is included. Indeed, Gobalet and Diamond (1979) have demonstrated that there is an interaction between domestic wealth and the mechanisms of a world-system.

A range of models may be identified from this discussion which may be hierarchically ordered and from which testable models can be derived. The most basic model would be simply one of *structure*, followed hierarchically by models of *structure and relations*, and finally by models of *structural relations*. The first model used in this example, then, is the three-block model reported in Table 4 (here, termed Model A). Hierarchically, one may then include, as a block of variables, the additive effects associated with flows or ties of interest (here, termed Model B). Taking a simple case such as treaty agreements, this would be merely the number of treaties existing between each nation and those which may be identified as core, semi-peripheral, and peripheral (three variables). Where the ties are not symmetric as they are with treaties (e.g., military interventions), the number of variables necessary to identify the additive effects of the flows increases twofold. Next, hierarchically, there are two competing models representing the differential advantages owing to the interactions either between initial level of GNP and treaty agreements as a block of variables (Model C) or between structural positions and treaty agreements as a block of variables (Model D). Model D allows for non-additive distinctions across levels of the world hierarchy among treaty ties. Finally, both may be considered conjointly in a single equation along with the additive effects (Model E).⁷

That model (Model C) over which no other regression models significantly increase the value of the R^2 and which significantly increase the R^2 over any model hierarchically simpler (see lower panel of Table 5) is presented in Table 5. The additive effects associated with the structural and control variables are similar to those of the three-block model with the exception of the

Table 5. Regression Estimates for Changes in GNP/Capita (1960-1970) for Accepted Treaty Model^a

Independent Variables	Partial Regression Coefficients (Standard Errors)
GNP/Capita (1960)	.235 (.060)**
Secondary School Enrollment Ratio (1960)	3.710 (1.348)**
Semi-periphery Periphery	-201.28 (90.18)*
Core Treaties	-276.55 (94.59)**
Semi-periphery Treaties	-8.873 (12.297)
Periphery Treaties	.829 (8.052)
Core Treaties X GNP	10.281 (2.632)**
Semi-Periphery Treaties X GNP	.006 (.007)
Peripheral Treaties X GNP	.023 (.007)**
Regression Constant	-.044 (.005)**
R^2	183.71
	.715

(N = 114 Nations)

* Significant beyond .05.

** Significant beyond .01.

a.	F _(C-A) (d.f.)	=	14.17 (6, 103)
	F _(C-B) (d.f.)	=	29.03 (3, 103)
	F _(D-A) (d.f.)	=	1.19 (12, 97)
	F _(D-B) (d.f.)	=	1.49 (9, 97)
	F _(E-C) (d.f.)	=	0.85 (9, 94)

initial level of GNP per capita. In Model C, initial Gross National Product has a significant positive effect on change in GNP per capita during the period covered by the data. Interpreted, those nations which possess greater economic wherewithal at a given starting point fare better, on average, over a subsequent period.

One other main effect, that describing the number of treaties between a nation and all peripheral nations (Periphery Treaties), is significant and positive. This effect, however, is more accurately discussed with reference to the interaction associated with the joint effect of GNP level and peripheral treaties (Stolzenberg, 1979). Considering, then, the two effects conjointly, one may say

that some nations of the sample are not in a position to take the fullest economic advantage of their political and economic linkages (treaties). Specifically, nations entering into a treaty with peripheral nations tend to gain less economically if they already enjoy a relatively high level of Gross National Product. A nation with relatively greater economic assets will probably be more likely to engender financial obligations from a treaty with a peripheral nation than one without such assets. Additionally, a relatively more affluent nation may also be a likely candidate to supply military or economic aid to a less powerful nation with whom it is bound by treaty. The more such treaties that exist for these nations with relatively higher starting GNP per capita, the more will the nation's economic growth be deleteriously affected. Such would have been the position of the United States or other core powers in the late 1960s as well as more recent years.

Though a second interaction term manifests a significant coefficient, its interpretation is ill-advised given the absence of significant main effects associated with both its components. It is more likely that there is no "real" relationship described here, but rather a single outlier or a small number of outliers which have been captured by this interaction.

Though this single example is provided here, the generality of this mode of presenting the results of cluster analyses in a more conventional form (i.e., regression models) may be appreciated. The same exercise could be pursued with the Snyder and Kick (1979) data using other transnational flows incorporated by those analysts. The matrix of international trade relations, for instance, translated into regression models via blockmodels has been demonstrated elsewhere. There, different costs and benefits accruing to trade in even similar commodities, but between different levels of the global hierarchy, were demonstrated (Steiber, 1980b). Moreover, all transnational or other flows could be introduced into a model simultaneously, though it is likely that the relative payoffs in substantive insights gained would be outweighed by the costs inhering in such a cumbersome presentation. Indeed, in problems where one has very few degrees of freedom initially, discretion in the selection of ties or flows of interest is most certainly advised.

DISCUSSION AND IMPLICATIONS

The blockmodeling approach to presenting sociometric or other relational data has done much to distill information on large networks into more manageable and interpretable forms. Clustering algorithms such as CONCOR have greatly facilitated the initial blocking of such data in order that they might be presented in reduced matrices or blockmodels. It has been demonstrated here, however, that while CONCOR offers generally useful reductions in the number of categories necessary to capture the salient contrasts across members of a data set, it falls short of maximizing those contrasts. An alternative to CONCOR, termed RECOR, was shown to improve the "leanness" of fit among blocks within data sets and at the same time increase the average similarity of units within the clusters to which they have been assigned. These improvements in leanness and within-cluster similarity by RECOR were demonstrated in each of the three bodies of data presented.

In applications to more conventional "hypothesis testing," it was shown that blockmodeling results can easily be translated into models of *structure*. More important from a substantive perspective, however, are the means by which relations and structural relations can be presented in regression models based on blockmodels. Empirically, *structural relation* models for the sample data were able to explain significantly more of the error variance than simple models of structure and relations and offer means by which such processual relations might create structure. There is every indication that this would be the case with other data sets as well.

In the past, primary emphasis in blockmodeling has been on the underlying structure in a body of data. The means by which structure traditionally has been detected, however, has been through the processes or flows inherent in a relational network or body of data. It is, then, not too surprising that analysis of structure alone has offered substantial insight into the referents of relational data. Using regression methods (e.g., Snyder and Kick, 1979), structure has been able to explain a large share of the variance and identify the implications of structure

in conceptually-important dependent variables. The above example, however, demonstrates that even beyond that variance which processually defined structure has been able to explain, there remains a significant amount of variance uniquely described by the processes in the data themselves.

Aside from offering a "better mousetrap" for network analysts, the above is a suggestion for further, more rigorous analyses of networks. There are certainly more sophisticated means by which blockmodels may be translated into forms amenable to hypothesis testing. Space considerations, however, suggest that these efforts be reserved for subsequent research. The limitations to be imposed on the translation of social relations are not limited by the data at the sociologist's disposal, and the methodological means available for extending these efforts are virtually unlimited.

FOOTNOTES

1. It should be emphasized that the mindless application of a clustering algorithm to allow the data to "speak for themselves" is ill-advised. No algorithm offers definitive answers to questions, per se, nor should one be expected to provide such information. Conceptual and empirical considerations should guide the researcher in the use of any algorithm.
2. Through the computation of this similarity matrix, RECOR follows the same preliminary stages as Arabie's (1977) and Shepard and Arabie's (1979) ADCLUS. Their search for overlapping group membership, however, takes them in other directions beyond this point.
3. Zeroblocks referenced here are those whose block densities are actually 0.00. In binary images, however, zeroblocks are identified as those whose block densities simply fall below the mean block density for the permuted matrix.
4. Direct comparisons of the zeroblocks derived and the average correlation of individual vectors with the clusters to which they are assigned is possible only when the number of clusters a reclustered matrix and the original configuration are the same.

5. A decrease in the number of clusters can be achieved when the membership of any given cluster is successively diminished by the closer ties of its members with other clusters. Ultimately where there remain only weak ties among members within such a cluster, all members may, during the course of an iteration, manifest stronger similarities with members of other blocks than with their own. This can, then, remove all members from the cluster, thereby deleting the cluster.
6. Regression analysis and the mode of cluster analysis described in these pages are both based on correlation matrices. The results reported for the regression models of Snyder and Kick (1979) and reported here, however, introduce into the regression analysis a conceptually-advised dependent variable *not* included in the clustering effort.
7. Only two-way interactions are considered in this analysis. Higher order interactions not only are extremely complicated to interpret, but they also serve to create a virtually unmanageable number of variables and the potential for severe multicollinearity.

REFERENCES

- Arabie, P.
1977 "Clustering representations of group overlap." *The Journal of Mathematical Sociology* 5:113-128.
- Arabie, P., S. Boorman, and P. Levitt
1978 "Constructing blockmodels: how and why." *Journal of Mathematical Psychology* 17:21-63.
- Boorman, S. and H. White
1976 "Social structure and multiple networks II." *American Journal of Sociology* 81:1384-1446.
- Breiger, R., S. Boorman, and P. Arabie
1975 "An algorithm for clustering relational data with applications to social network analysis and comparison with multi-dimensional scaling." *Journal of Mathematical Psychology* 12: 328-383.
- Gobalet, J. and L. Diamond
1979 "Effects of investment dependence on economic growth: the role of internal structural characteristics and periods in the world economy." *International Studies Quarterly* 23:412-444.
- Homans, G.
1950 *The Human Group*. New York: Harcourt-Brace.
- I.B.R.D.
1972 *World Bank Atlas*. Washington, D.C.: International Bank for Reconstruction and Development.

Mid-American Review of Sociology

- Sampson, S.
1969 *Crisis in a Cloister*. Ann Arbor: University Microfilms. Number 69-5775.
- Shepard, R. and P. Arabia
1979 "Additive clustering: representation of similarities as combinations of discrete overlapping properties." *Psychological Review* 86:87-123.
- Snyder, D. and E. Kick
1979 "Structural position in the world system and economic growth, 1955-1970: a multiple network analysis of transnational interactions." *American Journal Sociology* 84:1096-1126.
- Steiber, S.
1980a "Quantitative models and methods for world-system analysis." *The Sociological Quarterly* 20:forthcoming.
1980b "International trade and domestic development: a cross-national assessment of dependent relations in the world-system." mimeo.
- Stolzenberg, R.
1979 "The measurement and decomposition of causal effects in nonlinear and nonadditive models." Pp. 459-488 in K. Schessler (ed.), *Sociological Methodology* 1980. San Francisco: Jossey-Bass.
- U.N.E.S.C.O
1968 *Statistical Yearbook 1966*. Paris: United Nations Educational, Scientific and Cultural Organization.
- White, H., S. Boorman, and R. Breiger
1976 "Social structure from multiple networks I." *American Journal of Sociology* 81:730-780.

IN RELATION TO SCIENCE AND SOCIAL
NONSCIENCE: A CRITIQUE OF
PEARSON AND FISHER

Robert J. Wazienski
University of Kansas

Mid-American Review of Sociology, 1981, Vol. VI, No. 1:41-49

The objective set before us in this work is to explore the direction which social science has taken in this century. We do not intend to focus our attention on the nature and occurrence of phenomena, for these empirical notations are important only if one wants them to be. Furthermore, they are not relevant to science. The goal here is to simply distinguish, using a few isolated examples, between science and empiricism.

There seems to be remarkable confusion as to the distinction between science and empiricism. We must note that this confusion is not new.

In ancient times there was no social science. What we call social science was then called social philosophy or political philosophy or ethics. Aristotle's political and ethical philosophy, for example, was the same as his social science. It was much latter, with the success of Newtonian science, that social thinkers felt that it was possible to create a modern social science similar to the modern natural sciences. This modern or positivistic social science would have one basic characteristic, it was thought: it would be scientific (Stern, p. 1).

It was thought, and believed, that this "scientific" approach along with the statements derived thereof could introduce the capability of interpersonal agreement. The worth of this notion is sound, in that it is consistent with the function of science and reflects an adequate goal of the social sciences. However, along the path from social inquiry to social science, a reversal of the