A NEW APPROACH TO THE IDENTIFICATION OF INDUSTRIAL COMPLEXES USING INPUT-OUTPUT DATA

Howard Roepke, David Adams, and Robert Wiseman*

1. PURPOSE

While a number of industrial location models for single plants are available, models of entire industries, of industrial concentrations, and of the industrial sector of a national economy have been few and weak. One detailed model, the input-output matrix, does describe the relationship among these more aggregated units of industrial structure, but it has been little used for locational analysis because of its complexity and its lack of direct spatial reference.

An input-output matrix which records the transactions between large numbers of industries is of great value for tracing in detail the impact of specific changes upon the economy as a whole; it is too complex to be used to make more general statements on spatial patterns of activities. Therefore, what is needed is a way of simplifying the model for further analysis.

The research reported here applies factor analysis to an input-output table as a means of identifying functionally related groups of industries. This results in the identification of groups of industries which may be called industrial complexes. An industrial complex contains a base group of industries that have similar patterns of transactions, and it also includes other industries which are major suppliers or markets for those within the group.

Factor analysis makes possible the identification of three types of industrial complexes. The first is a grouping of industries on the basis of their total interaction. The interaction will consist of flows of both materials and products and thus is nondirectional. In addition, it is possible to identify two directionally-specific types of industrial complexes. One type is a grouping of industries on the basis of their common sources of materials. The other is a grouping on the basis of common destination industries for their products. Neither of these latter two necessarily implies any direct relationship between the industries within a complex, but this approach may help to explain the apparent heterogeneity of industries found in certain areas.

It should be noted that this research is concerned with the identification and

* Professor of Geography, University of Illinois-Urbana, Ph.D. candidate in Geography, University of Illinois-Urbana, and Assistant Professor of Geography, University of Kansas, respectively.

Date received: November, 1972; revised, July, 1973.

analysis of functionally related groups of industries. There is no concern with the relationship between sector activities and final demand. Therefore, the assumption of fixed input proportions, necessary for predictive input-output analysis (and also used in classical general equilibrium analysis), is not a serious problem. There is no attempt here to predict the impact of change on the system but rather a recognition that the coefficients reflect an existing functional relationship for one point in time.

This paper is basically concerned with testing the feasibility of the approach. The concern is more with alerting researchers to the possibilities and problems of this approach than with reporting specific results from the initial analysis.

2. OTHER METHODS OF INDUSTRIAL COMPLEX ANALYSIS

Other investigators have used two different operational definitions of the industrial complex concept. One definition considers the industrial complex as a group of highly interrelated industries contained within some major economic unit. The other definition describes industrial complexes as interdependent industries located at a single center or within a common region. The former, an aspatial type of complex, can be examined by a straightforward analysis of interindustry flow data. This kind of industrial complex can be identified by grouping the industry sectors of an input-output table into major components which show a high degree of interaction among the constituent industries.

The most common technique for implementation of this analysis has been the triangulation of the interindustry flows represented in an input-output matrix.[1] Although this technique has been demonstrated to have some utility for recognizing industrial complexes on a gross scale, it does have certain deficiencies and considerable information loss. It is particularly limited in its ability to identify smaller complexes, complexes of smaller industries, and secondary groupings in which the included industries have stronger ties to a more important complex. Furthermore, it must be noted that in practice the blocs are difficult to recognize.[2] Another problem arises from the fact that these groupings of sectors result from a single ordering, optimized according to a certain criterion. There may be latent groupings which would result from other orderings of the sectors.

Other aspatial approaches to the recognition of industrial complexes include that of Czamanski [3]. In attempting to identify industrial clusters which constitute a substantial portion of the multiplier effects within an economy, he utilizes eigenvectors and eigenvalues derived from analysis of input-output coefficients. However, the groupings do not appear to be readily definable, and the nature of the relationships among the composite industries is difficult to understand. Another methodology is found in Campbell's analysis of interindustry flows as directed graphs [1]. The approach is conceptually pleasing, but it does not

---

[1] For a more extensive description of triangulation, see Yan [18, Chap. 6]. Operational details are discussed in Helmstädter [6] and in Simpson and Tsukui [15, pp. 444–446].

[2] For an illustration of this point, see the triangulated input-output table for Israel presented by Leontief [13, pp. 152–153].

provide a very sensitive tool for the recognition of industrial groupings, since much information is lost in translating the input-output table into graphic form.

The second definition of an industrial complex described above implies a spatial association among the industries that comprise it. For instance, according to Isard and Smolensky [9, p. 110], an industrial complex is "a set of activities occurring at a given location and belonging to a group... of activities which reap important external economies because of their close production, marketing or other linkages." The idea is similar to the territorial-production complex described by Kolosovsky [12].

These spatially defined complexes have been sought in a variety of ways. Isard, Schooler, and Vietorisz [8] studied in a very detailed fashion specific technical bonds between types of manufacturing plants and the resultant interdependencies. Richter [14] and Streit [16] have used correlation analysis of the number employed by presumably functionally linked industries in order to suggest the existence of both interdependency and spatial association for certain groups of industries. Karaska [10] utilized input-output data for the Philadelphia region to identify industries for which local supply and demand linkages were of relatively great importance within this particular study area.

Although such searches for agglomerative forces within an industrial economy clearly have major value, there are certain difficulties with these attempts to recognize spatially-defined industrial complexes. As Richter [14] noted, high intercorrelations in the data are likely to be in part due to forces other than the tendency for the location of one type of industry to influence the location of another. A more difficult problem is presented by the question of distances that may occur between spatially associated firms. With a modern intercity transport system, significant agglomeration economies could accrue for firms within, say, 150 miles of each other. Thus, in the correlation approach to measurement of spatial association, the size of the observational unit is of critical importance.

This brief review of previous work suggests that recognition of industrial complexes has so far been hindered by a lack of sufficiently powerful techniques. In this study, principal axis factor analysis with varimax rotation is proposed as an approach which may avoid some of the problems noted here. It is suggested that factor analysis can be used as a highly sensitive replacement for the triangulation procedure in the recognition of industrial complexes defined solely upon the basis of observed interdependency within an economic unit. The identification of aspatial complexes can serve as an input to locational studies searching for spatial relationships within functioning complexes in specific regions.

3. DATA AND ANALYSIS

The investigation utilized an input-output table prepared for the Canadian Province of Ontario [4]. The original table of 1965 direct transactions was reduced from its 51 rows and columns to a 44 by 44 matrix containing only interindustry product linkages. The seven omissions include transportation and service industries as well as more general economic sectors.

Factor analysis of an input-output matrix is done in the same fashion as with any other data matrix. Column vectors in the matrix, representing direct transactions that are *inputs* to particular industrial sectors, are treated as sets of observed values for particular variables. Row vectors in the matrix, containing transactions that are *outputs* of particular sectors, are treated as values for variables at a particular observational unit. Thus, a_{ij}, which represents the dollar value of goods and services sold by industry i (outputs, rows) to industry j (inputs, columns), becomes x_{ij}, the value at the ith location of variable j.

Three separate factor analyses were conducted on the Ontario data. The first was done on an aggregated transactions matrix where flows between industries are combined so that the total interchange between industries is shown regardless of the input or output roles of either. This symmetric matrix B is derived from A, the direct transactions matrix, so that $b_{ij} = a_{ij} + a_{ji}$. The other two analyses were applied to directionally specified input-output data. One was an R-mode factoring which produced groups of industries receiving common patterns of inputs—their materials came from the same industries. The other was a Q-mode factoring which produced groups of industries having common destinations for their outputs—they market to the same industries.

In each of the three analyses, an interindustry correlation matrix was computed and then subjected to principal components analysis. The derived components with eigenvalues of one or greater were rotated to a varimax solution. The highest loading industries on each factor thus represent a group with similar linkage patterns. To further characterize the groups, factor scores were examined which express the relationship between a factor and the original data. High loadings and scores were used to operationally define industrial complexes.

Aggregated Matrix Results

Thirteen dimensions are derived from the aggregated transactions matrix about which industries are grouped by similarities of transaction flows. Together they account for 85 percent of the variance found in the data. Table 1 shows that Factor I, explaining 12 percent of the total variance, is composed primarily of metal working industries—Metal Stamping, Fabricated Metals, Electrical Appliances, etc. Examination of factor scores reveals a strong tie between the factor and Iron and Steel Mills (6.15 standard deviations). To a lesser extent, the factor is related to Other Metal Fabricating Industries. Thus the salient characteristic of this complex is a heavy dependence on Iron and Steel Mills, an industry external to the group.

Factor II represents a grouping made up entirely of textile industries; it accounts for 12 percent of the total variance. Synthetic Textiles, Other Textiles, Cotton Cloth and Yarn, and Clothing Industries have both the highest factor loadings and the significant factor scores. Hence, the ties that relate the components of this textile complex appear to be predominantly internal to the industries grouped on the factor. An examination of the transactions shown in the input-output data confirms that the most important flows for each industry link it with another in the group.

In a similar manner, the remaining industrial complexes may be defined and described from the results of the analysis presented in Table 1. Factors III (Chemicals), VI (Paper and Printing), VII (Wood), VIII (Extractive), and X

TABLE 1: Factors Derived from Aggregated Matrix with Factor Loadings Greater than 0.5 and Factor Scores Greater than 1.0

Factor I: Metal Using Industries
- Loadings
  - –.96 Metal Stamping, Pressing, Coating
  - –.95 Fabricated and Structural Metals
  - –.90 Electrical Appliances
  - –.75 Other Metal Fabricating
  - –.59 Other Transport Equipment
  - –.52 Electrical Industrial Equipment
- Scores
  - –6.15 Iron and Steel Mills
  - –1.22 Other Metal Fabricating

Factor II: Textiles
- Loadings
  - –.96 Knitting Mills
  - –.93 Other Textiles
  - –.90 Clothing Industries
  - –.90 Cotton Yarn and Textiles
  - –.74 Synthetic Textiles
- Scores
  - –4.11 Other Textiles
  - –3.21 Synthetic Textiles
  - –2.29 Cotton Yarn and Cloth
  - –1.89 Clothing Industries

Factor III: Chemical Industries
- Loadings
  - –.87 Plastics and Synthetic Resins
  - –.82 Pharmaceuticals and Medicines
  - –.74 Paint and Varnish
  - –.72 Other Chemical Industries
- Scores
  - –6.06 Other Chemical Industries

Factor IV: Farm Products
- Loadings
  - –.96 Grain Mills
  - –.94 Dairy Products
  - –.92 Meat and Poultry
- Scores
  - –6.35 Agriculture, Forestry, and Fishing

Factor V: Food Products Industries
- Loadings
  - –.86 Biscuits and Bakeries
  - –.84 Soft Drinks
  - –.82 Other Food Industries
  - –.79 Distilleries, Breweries, and Wineries
  - –.63 Sugar and Confectionaries
- Scores
  - –4.78 Other Food Industries
  - –2.36 Grain Mills
  - –1.96 Sugar and Confectionaries
  - –1.89 Paper Products

Factor VI: Paper and Printing
Loadings
−.98 Pulp and Paper Mills
−.91 Printing and Publishing
−.84 Paper Products
Scores
−5.50 Pulp and Paper Mills
−2.64 Paper Products
−1.38 Printing and Publishing

Factor VII: Wood Industries
Loadings
−.90 Sawmills
−.89 Other Wood Industries
−.69 Furniture and Fixtures
Scores
−4.03 Sawmills
−3.82 Other Wood Industries
−1.53 Furniture and Fixtures
−1.22 Other Metal Fabricating

Factor VIII: Extractive Industries
Loadings
−.83 Petroleum Refineries and Coal Products
−.76 Clay, Lime, and Cement
−.48 Mining
Scores
−5.13 Mining
−2.22 Petroleum Refineries and Coal Products
−2.18 Clay, Lime, and Cement

Factor IX: Motor Vehicles-Aircraft and Non-Metallic Minerals
Loadings
−.88 Motor Vehicles and Aircraft Industries
−.85 Other Non-Metallic Mineral Products
Scores
−5.76 Motor Vehicles and Aircraft Industries
−2.10 Other Non-Metallic Mineral Products

Factor X: Communications and Electrical Products
Loadings
−.94 Communications Equipment
−.75 Other Electrical Products
−.70 Electrical Industrial Equipment
Scores
−5.50 Other Primary Metals
−2.31 Communications Equipment
−1.20 Electrical Industrial Equipment

Factor XI: Agriculture and Leather Industries
Loadings
−.75 Agriculture, Forestry, and Fishing
−.75 Leather and Leather Products
Scores
−4.66 Meat and Poultry
−2.50 Grain Mills
−2.40 Leather and Leather Products
−2.14 Sugar and Confectionaries

TABLE 1—*Continued*

Factor XII: Heavy Metals Industries
Loadings
- −.68 Iron and Steel Mills
- −.61 Other Primary Metals Industries
- −.61 Other Transport Equipment

Scores
- −4.88 Other Metal Fabricating
- −1.61 Motor Vehicle and Aircraft
- −1.18 Metal Stamping, Coating, and Pressing
- −1.14 Other Transport Equipment
- −1.03 Mining
- 1.01 Iron and Steel Mills
- 1.08 Miscellaneous Manufacturing
- 1.12 Clay, Lime, and Cement

Factor XIII: Miscellaneous Manufacturing and Rubber Products
Loadings
- −.87 Miscellaneous Manufacturing
- −.54 Rubber Products

Scores
- −5.09 Plastics and Synthetic Resins
- −2.21 Miscellaneous Manufacturing
- −1.14 Rubber Products
- −1.03 Leather and Leather Products
- 1.30 Grain Mills
- 1.41 Other Non-Metallic Mineral Products

(Communications Equipment and Electrical Products) showed primarily internal product flows. Factors IV (Farm Products), XI (Agriculture and Leather Industries), XII (Heavy Metals),[3] and XIII (Miscellaneous Manufacturing and Rubber Products) were all strongly tied to industries external to the group represented by the factor. Factor V (Food Products) has both internal and external ties. Factor IX contains two unrelated industries; each has ties largely with itself.

R-Mode and Q-Mode Analyses

The results of *R*-mode factoring of input flows and *Q*-mode analysis of output links produced striking similarities to the results of the aggregate matrix analysis in: (1) the number of factors derived, (2) the amount of total variance explained by all factors, and (3) the industries found to be related by factor loadings and/or scores. The presence of similarly grouped industries within all three of the factor analyses suggests that the Ontario economy contains several clearly defined industrial complexes of functionally related industries. Those which stand out most clearly are shown in Table 2. Differences in the grouping of industries between the different factor analyses are generally confined to the last

[3] This seems to be functionally similar to Factor I, as many industries which load on one show high factor scores on the other.

TABLE 2: Percentage of Total Variance Explained in Three Factor Analyses

	Aggregate Factors	Origin Factors	Destination Factors
Similarly Defined Factors			
Metal Using Industries	12.3	11.5	14.0
Textiles	11.6	5.8	8.6
Farm Products	8.8	4.7	10.2
Food Products	8.2	3.2	5.4
Chemicals	7.2	6.5	9.7
Paper and Printing	6.1	3.1	8.1
Communications and Electrical Products	6.0	3.3	4.3
Wood Industries	5.1	5.2	5.0
Extractive Industries	4.1	3.1	5.2
Total	67.4	46.4	70.5
Balance of Rotated Factors	18.1	34.1	15.7
Percentage of Total Variance Explained by All Rotated Factors	85.5	80.5	86.2
Number of Rotated Factors	13	16	13

few groups derived in each factoring. (These groups are not included in Table 2.) It should be noted that, while the complexes have been identified with simple descriptive words, many of the component industries in the groups correspond to the 3-digit level of disaggregation in the United States S.I.C.

There are, however, small differences in the makeup of similar factors as they occur in each analysis. Rather than enumerate all such cases, three examples have been selected to demonstrate some of the similarities and differences found among the three analyses.

Figure 1 depicts the relationships among common factors for each of three complexes. A wood industries complex clearly emerges whether one considers industries on the basis of input patterns, distribution ties, or aggregate flows. Examination of factor scores and actual input-output data demonstrates that these industries are strongly related to each other by product flows. Hence salient linkages are primarily internal to the complex.

The textile complex is characterized by slightly varying factor composition among the three analyses. Industries comprising both origin and destination groups are contained within the aggregate factor, but the origin group excludes Clothing Industries. Similarly the destination factor omits Synthetic Textiles. This relationship may be explained by actual product flows, since all origin industries input heavily to the Clothing Industry. Conversely the destination group, Knitting, Cotton Cloth, and Yarn, as well as Other Textiles, receives significant input from Synthetic Textiles. One can conclude that, although product links among origin and destination groups are both internal and strongly external, they are all contained within the aggregate factor.

The Chemicals complex represents a more diffused set of relationships. The input-output data reveal that the industries grouped by origin input strongly to Other Chemicals and conversely the destination group receives significant input

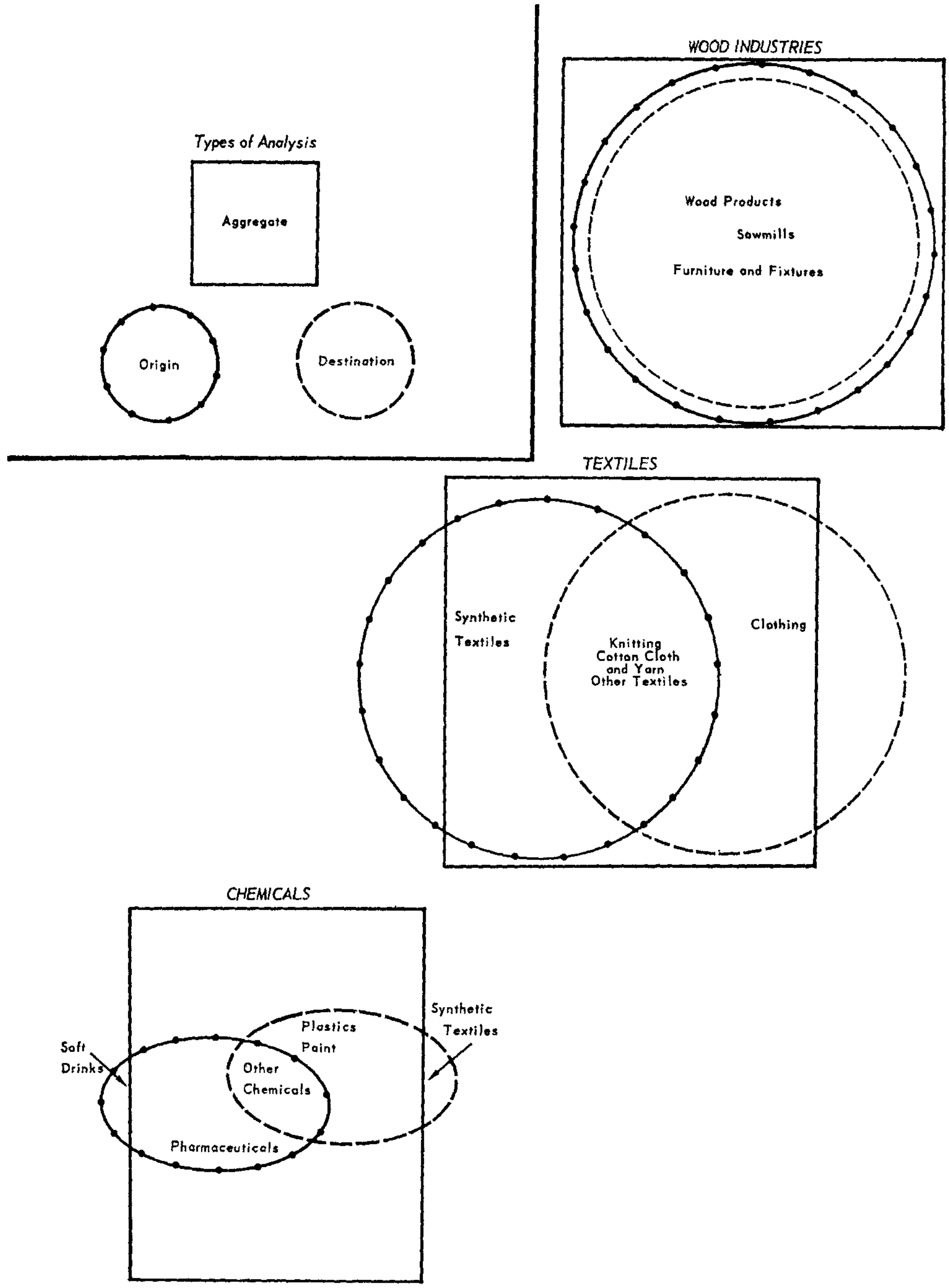

FIGURE 1: Comparative Factor Composition by Three Types of Analysis (Loadings >0.5).

from Other Chemicals. This is implied in Figure 1. However, there are many external ties and the complex displays markedly different composition as one considers input, output, and aggregate linkage patterns.

4. TEST OF COMPLEXES

In order to provide some supporting evidence of the interdependence within the industrial complexes derived by factor analysis, a measure of interindustry linkage was calculated for these groupings.

Streit [16] proposes the simple and useful index, L_{ij}, for determining the strength of an economic linkage between two industries i and j. The index is calculated from the matrix which presents annual interindustry transactions in money units. This matrix is represented as $A = a_{ij(n,n)}$, for n industrial sectors. The index is then calculated as follows

$$(1)\quad L_{ij} = L_{ji} = (1/4)[(a_{ij}/\textstyle\sum_{i=1}^{n} a_{ij}) + (a_{ij}/\sum_{j=1}^{n} a_{ij}) + (a_{ji}/\sum_{j=1}^{n} a_{ji}) + (a_{ji}/\sum_{j=1}^{n} a_{ji})]$$

with $0 \leq L_{ij} \leq 1.0$.

In order to interpret the values of this index, it should be realized that a uniform distribution of interindustry flow would yield equal matrix elements. The result for the present Ontario 44 by 44 matrix would be that L_{ij} would equal $1/44$ or .023 for every possible combination of i and j. Therefore, linkage index values ranging upward from .023 indicate increasingly significant concentration of flows and increasingly strong ties.

To derive a summary measure of interindustry linkage for an entire complex of k industries, one may set up a triangular matrix of interindustry linkages. This can be used to determine the mean linkage index, $\bar{L}$, for the elements above and to the right of the principal diagonal of the linkage index matrix.

$$\bar{L} = (L_{1,2} + L_{1,3} + \cdots + L_{1,k} + L_{2,3} + \cdots + L_{2,k} + \cdots + L_{k,k})/[(k^2 - k)/2]$$

The diagonal elements, $L_{1,1}$, $L_{2,2}$, etc., measure *intra*industry linkage. Their inclusion would exaggerate the strength of *inter*industry ties.

Table 3 summarizes the mean interindustry linkage values for the industrial complexes resulting from the three different analyses described above. The complexes are matched, as in Table 2, according to similarities among the constituent industries. For this test, only industries that have loadings with an absolute value of .65 or more and industries that have scores with an absolute value of 2.0 or more are included in the complexes.

These mean linkage values are, with one exception, more than twice the .023 value which would indicate uniformity in interindustry transactions. This provides an indication that the complexes identified by factor analysis are indeed functionally interrelated.

The varying sizes of the groups of industries make it difficult to compare the mean linkage indices shown in Table 3. The reason for this is that the maximum possible mean linkage index varies with the size of the group. The maximum mean

TABLE 3: Mean Interindustry Linkage Index Values ($\bar{L}$) for Derived Industrial Complexes

	Aggregate Factors		Origin Factors		Destination Factors	
	n[a]	$\bar{L}$	n	$\bar{L}$	n	$\bar{L}$
Metal Using	5	.102	7	.000	7	.049
Textiles	5	.123	3	.136	5	.123
Food Products	4	.049	2	.063	4	.081
Paper and Printing	3	.117	2	.183	3	.117
Chemicals	4	.107	3	.146	4	.075
Farm Products	5	.083	4	.109	6	.092
Wood Industries	3	.136	3	.136	2	.263
Extractive Industries	3	.156	3	.156	3	.156
Communications and Electrical Products	4	.103	3	.098	4	.103

[a] n = number of industries included in the complex.

TABLE 4: Relative Mean Linkage Indices ($\bar{L}_R$)

	Aggregate Factors	Origin Factors	Destination Factors
Metal Using	.41	.00	.29
Textiles	.49	.27	.49
Food Products	.15	.06	.24
Paper and Printing	.23	.18	.23
Chemicals	.32	.29	.23
Farm Products	.33	.33	.46
Wood Industries	.27	.27	.26
Extractive Industries	.31	.31	.31
Communications and Electrical Products	.31	.20	.31

index for a group of 2 is 1.0, which may be restated symbolically as $\bar{L}_{\max}(2) = 1.0$. If one more industry is added to the group, then $\bar{L}_{\max}(3) = 0.5$. For a group of arbitrary size k, $\bar{L}_{\max}(k) = 1/(k - 1)$. Because of this property, it seems appropriate to calculate a relative mean linkage index, $\bar{L}_R = \bar{L}/\bar{L}_{\max}(k)$. This has been done for each derived complex, and the results are presented in Table 4.

It is reasonable to hypothesize that the most strongly linked complexes would be those which can be most readily recognized by factor analysis. If this is true, then the order of the relative mean indices ($\bar{L}_R$) should be similar to the order of the percentages of variance explained by each respective factor (as shown in Table 2). A test of this similarity for the aggregate industry factor and its related complexes yields a Spearman rank correlation coefficient of .604 (significant at the .05 level). The same test applied to origin and destination industry factors and their complexes resulted in very low correlation coefficients.

This result aids in the recognition of the differing roles which may be filled by complexes derived from the three different factor analysis procedures. The factoring of aggregate data is most useful when a high level of interrelationship is the most desired property for internally related complexes to be derived. The *R*- and *Q*-mode factoring procedures are better suited to situations in which one is interested in flows of goods moving from an origin set of industries to a destination set of industries. The *R*-mode procedure, for instance, would allow identification of distinct externally related complexes of industries which would be markets for specified origin industries.

5. EVALUATION OF THIS APPROACH TO THE RECOGNITION OF INDUSTRIAL COMPLEXES

For the identification of industrial complexes the present methodology appears promising when compared to the previously discussed approaches. Factor analysis appears to differ from triangulation in that it is a simple way to identify minor as well as major industrial groupings. Although the results do not preserve the actual flow volumes upon which the linkages are based, they do make it possible to assess the nature of the relationships within each grouping. Campbell's approach [1], the use of directed graphs, has not yet been implemented for the purpose of identifying industrial complexes, but it appears that factor analysis would be more sensitive to the actual volumes of flow.

Czamanski [3] suggests a procedure designed to identify the sectors of an economy that generate substantial multiplier effects. This goal is complementary to the identification of functioning industrial complexes, and the methodology which he employs is somewhat similar to that suggested in this study. Czamanski starts with matrices of correlation coefficients relating industries on the basis of different types of product links. He then constructs a single matrix of r values such that

$$r_{ij} = r_{ji} = \max\left[r(a_{li}, a_{lj}), r(a_{il}, a_{jl}), r(a_{li}, a_{jl}), r(a_{il}, a_{lj})\right]$$

where $r(a_{li}, a_{lj})$ indicates the correlation of destination industries i and j on the basis of the relative importance to them of inputs from each origin industry l. His final step is the extraction of eigenvectors from this composite matrix. This has the advantage of identifying the greatest multiplier effects. However, as a result the ultimate groupings are not based upon readily identifiable sets of product links and coherent complexes become difficult to recognize.

The present study also extracts eigenvectors from a correlation matrix. However, actual correlation matrices of either input, output, or aggregate flows are employed, as opposed to Czamanski's composite matrix. The derived eigenvectors are interpreted as factors correlated with the original industries and are rotated to facilitate definition. This analysis of the Ontario data has produced readily identifiable and meaningful groupings of industries. Verification by use of the Streit Index tends to substantiate the validity of this approach in the recognition of industrial complexes. However, certain further comments may be useful in evaluating the suggested methodology.

It is important to note that the derived complexes are not based upon industry size or significance to the general economy. The principal grouping criterion is similarity of linkage patterns at all levels of activity; magnitude is significant only within the distribution of a particular industry's input or output ties. Thus one advantage of this approach is that it allows the consideration of small but possibly critical complexes within the industrial economy.

In the application of factor analysis to other areas of research, there is a tendency to attach significance to factors according to the amount of variance explained. This measure is applicable only to the pattern of multicollinearity indicated by the correlation matrix, and gives no information concerning the economic importance of any factor grouping. Measures of economic importance to use in conjunction with the technique suggested here could be derived by calculating simple indices from the interindustry transactions data aggregated or reordered on the basis of factor groupings.

The present research indicates that certain minor difficulties may be encountered in using the suggested methodology. Because factoring proceeds by deriving principal components, each of which successively removes or explains less variance within the matrix, the last few factors tend to be small and more difficult to interpret.

A second potential source of difficulty stems from the fact that it is possible for both factor scores and loadings to carry mixed algebraic signs. This is exemplified by Factors XII and XIII in Table 1. Factor XIII is most closely correlated with Miscellaneous Manufacturing and Rubber Products. It is characterized by strong links with Plastics and Synthetic Resins, weaker ties to Miscellaneous Manufacturing, Rubber Products, Leather and Leather Products, but also a tendency towards disassociation or lack of ties with Grain Mills and Other Non-Metallic Mineral Products. This may create problems of interpretation but, when applied spatially, may actually provide significant insights.

In the three analyses of the Ontario data, neither of these difficulties produced serious complications in interpreting results. In all cases, problem occurrence is confined to the last few factors derived. This suggests that the difficulties might be overcome by reducing the number of factors extracted through the choice of higher critical levels of factor acceptance (higher eigenvalues). Whenever other interpretation questions arise, a simple examination of the actual input-output data facilitates explanation. Hence we conclude that the suggested approach to recognition of industrial complexes has considerable utility.

It would be appropriate to conclude this section with a reminder that the concern here was with testing the feasibility of the approach. A 44 × 44 matrix is more aggregate than would be necessary for in-depth analysis. However, the technique has been shown to be operational and to produce meaningful results. Extensions and analytic applications are suggested in the final section.

6. POSSIBLE APPLICATIONS AND FUTURE RESEARCH DIRECTIONS

An obvious research direction of interest is to investigate the extent to which industries identified in an industrial complex as functionally related are also spatially proximate. This problem is being pursued in connection with analysis of a U.S. input-output table. Determining the degree to which such industries are associated spatially may be an additional approach to measuring the efficiency of the spatial organization of an economy. Earlier attempts to accomplish this by linear programming have been made by both economists and geographers (Henderson [7] and King *et al.* [11]). It is hoped that the production complex approach may also provide the basis for the development of new models of the spatial arrangement of industry at a level more aggregated than that of the single plant or firm.

A further application of this technique may lie in the making of comparisons between different economic units. It will be interesting to determine whether similar production complexes may be identified in areas of different size and with varying resources. Differing industrial structures may be more clearly identified by this technique—particularly when the analysis is extended to encompass imports and exports. International comparisons have commonly dealt with overall measures of indirectness (multiplier effects), with sectoral interdependence, or with a comparison of individual input coefficients (Chenery and Clark [2, pp. 201–13]).

Further research using this approach, as well as that suggested by Czamanski, might furnish initial guidance in economic development efforts. The application of such research to several input-output tables may produce fairly consistent groupings of industries which in the absence of local input-output data might guide economic planners in their general investment decisions. A further use of such an approach could be in preliminary analysis in connection with the growth pole concept. One of the difficulties in implementing this concept is finding an easy and economical way to identify appropriate sets of industries. (See Tolosa and Reiner [17].)

In summary, it seems that the factor analysis of input-output data may prove to be a technique of value in both spatial and aspatial analysis. It is hoped that it will be applied by other investigators in order that both its potential and its problems may be better understood.

REFERENCES

[1] Campbell, J. "The Relevance of Input-Output Analysis and Digraph Concepts to Growth Pole Theory," Unpublished Ph.D. dissertation, University of Washington, 1970.

[2] Chenery, H. B. and P. G. Clark. *Interindustry Economics*. New York: Wiley, 1959.

[3] Czamanski, S. "Some Empirical Evidence of the Strengths of Linkages Between Groups of Related Industries in Urban-Regional Complexes," *Papers, Regional Science Association*, 27 (1971), 137–150.

[4] Department of Treasury and Economics, Province of Ontario. *Ontario Statistical Review, 1969*. Toronto: Government of Ontario, 1970.

[5] Executive Office of the President, Bureau of the Budget. *Standard Industrial Classification Manual*. Washington, D.C.: Government Printing Office, 1957.

[6] Helmstädter, E. "The Hierarchical Structure of Interindustrial Transactions," in *International Comparisons of Interindustry Data*. Industrial Planning and Programming Series, No. 2, United Nations Industrial Development Organization, Vienna. New York: United Nations, 1969.

[7] Henderson, J. *The Efficiency of the Coal Industry, An Application of Linear Programming*. Cambridge: Harvard University Press, 1958.

[8] Isard, W., E. Schooler, and T. Vietorisz. *Industrial Complex Analysis and Regional Development*. New York: Wiley, 1959.

[9] ——— and E. Smolensky. "Application of Input-Output Techniques to Regional Science," in T. Barna (ed.), *Structural Interdependence and Economic Development*. London: MacMillan, 1963.

[10] Karaska, G. J. "Manufacturing Linkages in the Philadelphia Economy: Some Evidence of External Agglomeration Forces," *Geographical Analysis*, 1 (1969), 354–369.

[11] King, L., E. Cassetti, J. Odland and K. Semple. "Optimal Transportation Patterns of Coal in the Great Lakes Region," *Economic Geography*, 47 (1971), 401–413.

[12] Kolosovsky, N. N. "The Territorial-Production Combination (Complex) in Soviet Economic Geography," *Journal of Regional Science*, 3 (1961), 1–25.

[13] Leontief, W. "The Structure of Development," *Scientific American*, 209 (1963), 148–166.

[14] Richter, C. E. "The Impact of Industrial Linkages on Geographic Association," *Journal of Regional Science*, 9 (1969), 19–27.

[15] Simpson, D. and J. Tsukui. "The Fundamental Structure of Input-Output Tables, An International Comparsion," *Review of Economics and Statistics*, 46 (1965), 434–446.

[16] Streit, M. E. "Spatial Associations and Economic Linkages Between Industries," *Journal of Regional Science*, 9 (1969), 177–188.

[17] Tolosa, H. and T. Reiner. "The Economic Programming of a System of Planned Poles," *Economic Geography*, 46 (1970), 449–458.

[18] Yan, C. *Introduction to Input-Output Economics*. New York: Holt, Rinehart and Winston, 1969.