NEEDLESS COMPLEXITY IN THE IDENTIFICATION OF INDUSTRIAL COMPLEXES: A COMMENT

Stan Czamanski*

In a recent article in this Journal, Latham [4] argues for greater simplicity in the study of industrial complexes. Simplicity, especially when contrasted with "needless complexity," has an obvious appeal, but the meaning which Latham attaches to these terms is rather naive. To the scholar, the notion of simplicity is rooted either in the ability of a particular theory to explain observed facts with the help of fewer assumptions than alternative ones, or in its greater generality which permits it to cover a wider range of phenomena. An often quoted example is a comparison of the Copernican heliocentric with the Ptolemaic geocentric system. To the unsophisticated, simplicity resides in agreement with common sense and long familiar notions, and hence depends on his education and cultural background. Thus to most people acquainted with ordinary calculus, but not with ideas about invariants, Newtonian physics is simple, while Einstein's relativity is of immense complexity, despite its greater generality.

Latham's preference for Streit's [6] method over approaches based on graph theory or factor analysis does not depend upon the former criterion. Streit identified industrial groupings on the basis of positive links between pairs of industries examined seriatim. In some studies using graph theory or triangularization methods, an industry was considered to be a member of the grouping if its links with all industries in the set exceeded a certain threshold level, thus classifying it on the basis of several rather than on a single link. In studies relying on multivariate analysis (principal components or factor analysis), the criterion was even more general, since industries were classified as members of a grouping on the basis of similarity of their patterns of flows with all other industries in the economy, and not only with those in the set. (See Bergsman, Greenston, and Healy [1], Czamanski [3], and Roepke, Adams, and Wiseman [5].) The choice between the three approaches is therefore not a matter of simplicity vs. complexity, but depends upon how one defines terms and how useful they are analytically.

Nor do I see how mathematical complexity could be invoked to discriminate between the three approaches. After all, correlation analysis and principal components analysis, at the level used in the studies discussed, are both ordinarily taught in the same introductory statistics course, and practically all social scientists are familiar with both. Latham's assertion that his method is "intuitively more appealing than others" is singularly unconvincing. Intuition is not the most powerful criterion for choosing between alternative methods of analysis.

* Professor of City and Regional Planning, Cornell University.

Date received: November, 1976; revised, March, 1977.

In his pursuit of "simplicity," Latham objects to the analysis of national input-output tables, claiming that such methods "are not useful for policy purposes." In his discussion, however, he not only fails to substantiate this, but is above all inconsistent. For while he opens his paper by promising to review critically "the aspatial method of identifying complexes . . . so that others will not be tempted to salvage these methods by simply coupling them with a supplemental spatial analysis" [4, p. 45], a little later he states that "one should examine an input-output table for an entire country to discover through less limited interrelationships what industry might be most complementary to those already present" [4, p. 49], and ends up by using purely aspatial methods in his own modest study. The lack of constraints caused by friction of space in his research is due not to design but, as will be shown later, to his use of inadequate statistical techniques.

Latham fails to realize that in all studies dealing with industrial groupings, three different notions are involved. In a recent review, a consistent terminology has been applied in order to avoid confusion. (See Czamanski and Czamanski [2].) Thus, "clusters" were defined as subsets of industries of the economy connected by flows of goods and services stronger than those linking them to other sectors of the national economy. The concept is thus devoid of any spatial connotation. "Complexes," on the other hand, have been defined as groups of industries connected by important flows of goods and services, and showing in addition a significant similarity in their locational pattern. In some studies, the term "complex" has acquired a much narrower meaning, referring to one particular agglomeration in a specific location. All studies dealing with identification of industrial complexes are concerned with both aspects, i.e., intersectoral flows of goods and services, and locational patterns. Some of them begin by analyzing intersectoral flows, usually on the basis of national input-output tables, and proceed to an analysis of locational patterns, while others start with a narrow regional or urban economy and study existing intersectoral flows.

Streit, followed by Latham, classified as members of a complex, pairs of industries for which a statistically significant coefficient of spatial correlation exists, and for which the value of flows exceed a certain threshold. But while Streit worked at a high level of industrial aggregation and, more importantly, used as his units of observation German and French regions, Latham based his study on 199 industries and some 377 U.S. regions, presumably SMSA's or cities.[1] Streit's use of correlation analysis might have been statistically correct, but Latham's was almost certainly not. For the distribution of practically all phenomena among U.S. cities or metropolitan areas is highly skewed or J-shaped, and the resulting straight line coefficients of correlation are, as a rule, very high but meaningless. Transformations of variables are ordinarily insufficient to produce approximately normal bivariate populations, and recourse to partitioned arrays and pooled coefficients is needed, although even that is not always successful. No mention was made by Latham of any examination of the properties of his data, but he stated that all

[1] His explanation of the procedures used is imprecise and the reader is referred to unpublished material, a practice to be deplored. The claim that his work used the most detailed industrial classification is exaggerated since Bergsman, Greenston and Healy [1] used 186 sectors and Czamanski *et al.* [3] used 172.

his correlations were very high, and that the lack of spatial association was not a binding constraint in identifying complexes [4, p. 54]. It must be assumed that the absence of spatial constraints was spurious due to his application of correlation analysis to nonnormal populations.

Finally, the form of the paper deserves some comments. About two-thirds of it is devoted to a review of the work of others, and only one-third to the presentation of Latham's own contribution. As a review it is, however, incomplete for some of the most important contributions are not even mentioned, and beyond that it lacks precision to the point of distorting some of the studies described.

REFERENCES

[1] Bergsman, J., P. Greenston, and R. Healy. "The Agglomeration Process in Urban Growth," *Urban Studies*, 9 (1972), 263–288.

[2] Czamanski, D. Z. and S. Czamanski "Industrial Complexes: Their Typology, Structure and Relation to Regional Development," *Papers, Regional Science Association*, 38 (1977), forthcoming.

[3] Czamanski, S. *et al. Study of Clustering of Industries.* Halifax: Dalhousie University, 1974.

[4] Latham, W. R. "Needless Complexity in the Identification of Industrial Complexes," *Journal of Regional Science*, 16 (1976), 45–55.

[5] Roepke, H., D. Adams, and R. Wiseman. "A New Approach to the Identification of Industrial Complexes Using Input-Output Data," *Journal of Regional Science*, 14 (1976), 15–29.

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