PSYCHOMETRIKA—VOL. 50, NO. 3, 371–375.
SEPTEMBER 1985
REVIEWS

REVIEWS

J. Scott Long *Confirmatory Factor Analysis. A Preface to LISREL.* Sage University Paper series on Quantitative Application in the Social Sciences 07-033. Beverly Hills and London: Sage, 1983. 88 pp. $5.00.

J. Scott Long *Covariance Structure Models. An Introduction to LISREL.* Sage University Paper series on Quantitative Application in the Social Sciences 07-034. Beverly Hills and London: Sage, 1983. 95 pp. $5.00.

B. S. Everitt *An Introduction to Latent Variable Models.* Monographs on Statistics and Applied Probability. London and New York: Chapman and Hall, 1984. 107 pp. $20.00.

W. Saris and H. Stronkhorst *Causal Modeling in Nonexperimental Research. An Introduction to the LISREL Approach.* Amsterdam, The Netherlands, Sociometric Research Foundation, 1984. 335 pp. approximately $14.00.

Introduction

In recent years there have been important breakthroughs in the formulation, the fitting, and the statistical analysis of linear simultaneous equation models. These developments have taken place in econometrics, in sociometrics, and in psychometrics, with relatively little interaction between the disciplines. The psychometricians took factor analysis as their starting point, the sociometricians used multiple regression and path analysis, and the econometricians introduced errors-in-variables into their simultaneous equations models. As a consequence of these separate developments the field was in a rather confused state in the early seventies. Due to the efforts of econometricians like Goldberger and Wold, sociometricians like Blalock and Duncan, and psychometricians like Jöreskog, Bentler and Browne, this situation has improved tremendously. There is very little doubt, I think, that the efforts of Karl Jöreskog have been by far the most important unifying factor in this process. There were many people who knew that the separate developments that were going on were closely related, and in fact often duplicating the same results. There were not many people, however, who really tried to do something about this unfortunate state of affairs. Jöreskog and his co-workers wrote a series of successful computer programs, applied them to many examples from various disciplines, and traveled around the world giving lectures and courses to many different audiences. One of the consequences of all these activities is that there are, by now, many people who have difficulty in distinguishing linear simultaneous equation modeling from the application of the LISREL program. This is a new possible source of confusion, which must be kept in mind. Another source of confusion is the fact that sociometrists like to speak of "causal models" in this context. It is not at all clear if and in how far linear structural models have anything to do with causality. And finally the new sophisticated computer programs do not provide much guidance in the model selection problem; in fact, they tend to make it much more serious than it was, by providing a far wider range of models that can be fitted. I think that the present situation can be summarized by saying that the technical problems are mostly solved, but that as a consequence of this enormous technological progress the methodological problems have become more serious. And this, of course, is a very familiar state of affairs, not only in data analysis.

In this review I shall compare a number of recent books on "covariance structure

models", "latent variable models", "causal models", or "LISREL models". The comparison will be in terms of the range of models they cover, their statistical sophistication, their practical usefulness, but especially in terms of the methodological issues that have been mentioned above.

Causality

Let me first state my own position on the issue of "causality" in no uncertain terms. I think that the use of causal terminology in connection with linear structural models of the LISREL type means indulging in a modern, but nevertheless clearly recognizable, version of the "post hoc ergo propter hoc" fallacy. In this I agree with Guttman (1977), Ling (1982), and Cliff (1983). Of course one could maintain that causality (or "explained variance" or "direct effect" or "determination") is used merely in a metaphorical way. Such a cavalier attitude towards terminology becomes hard to defend if educational programs are based on your metaphors, such as the metaphor that "intelligence is largely genetically determined" or "allocation of resources to schools has only very minor impact on the careers of students". It seems to me that the use of cause-effect terminology cannot be defended, except in those rare cases (such as Mendelian genetics) in which information about the causal mechanism is available from other sources. If all the information we have is a set of correlations, then we can try to describe these correlations "parsimoniously" in terms of restrictive models. But that is about all we can do. Thus there simply is no LISREL method of theory construction.

The two books by Scott Long and the book by Everitt avoid methodological issues as much as possible. In the interpretations of the examples causal terminology is used, but sparingly. Saris and Stronkhorst use it freely, and give direct causal interpretations of their examples. This disparity reflects the quite different nature of the books. Scott Long and Everitt have written technical books, concentrating on identification and fitting. Saris and Stronkhorst are clearly in the sociological tradition, and stay quite close to such books as Kenny (1979) and James, Mulaik, and Brett (1982). The remarks made by Ling (1982) and Goldberger (1983) in their review of these books also apply here. The causal superstructure of model fitting consists of unclear and often unsound terminology and interpretation. It is unfortunate if this terminology is used freely, but it is even more unfortunate if introductory books do not contain warnings against such practices.

Interpretation and Model Search

The excellent recent work of Wermuth, Hodapp, Speed, Kiiveri, Lauritzen, Darroch and others (cf. Kiiveri & Speed, 1982, for a review and references) has made it perfectly clear that causal models are equivalent to conditional independence statements, i.e., in the linear context to the vanishing of certain partial correlation coefficients. This is their mathematical content. Correlation, not causality. Thus the "LISREL-approach" searches among conditional independence models for a model with a good fit, high explanatory power, good parsimony. The search is guided, partly, by objective criteria such as chi-square, R^2, the AIC-statistic, but also partly by "prior knowledge" or "interpretability". It is difficult enough to combine the various objective criteria into a single model selection rule, but the subjective criteria which also play a prominent role make it debatable if two independent social scientists will ever come up with the same model. Thus the LISREL-approach is more subjective and exploratory than it is usually presented to be.

Both Scott Long and Everitt play little attention to model search (or specification search) as such. This means that they prefer to pretend that the LISREL-approach is confirmatory, and that the model is fixed before the data are collected or the analysis is done. It seems to me that this supposition does not correspond with actual practice. Saris

and Stronkhorst realize this, and pay a great deal of attention to techniques for modifying and polishing the model. None of the books pay much attention to the interpretation of the models in terms of conditional independence. Saris and Stronkhorst briefly discuss the 'Simon-Blalock' approach, but dismiss it as unsystematic and somewhat old-fashioned. The interpretation is in causal and manipulative terms: If we increase father's income by one standard deviation, then son's IQ will increase eight points. In introductory chapters it is usually mentioned that we need a "ceteris paribus" clause here: All other variables (including those not in the system that influence son's IQ) must be kept equal at their current values. It is not indicated how such an experiment could be performed, and the clause is rapidly forgotten on the level of concrete examples. Of course the qualitative interpretation of causal models is usually fixed from the beginning by the choice of the model, only the size of the path coefficients is left free. This often results in the kind of a-priorism that sociologists are familiar with, and in computerized forms of armchair theorizing.

Range of Models and the Naming Fallacy

The four books cover a different range of models. Scott Long's first volume discusses the LISREL measurement model, i.e., confirmatory factor analysis. His second book discusses the full LISREL model. The two books really should be considered as one, and the two volumes nicely correspond with the historical development of the LISREL-model. As the title already indicates the book by Everitt covers a much wider range of models. More than half of the book is on factor analysis and the LISREL model, but there is also some discussion of the general latent structure model, of the latent class model, and of the recent extensions of structural modelling to categorical variables by logit or probit approaches. This makes the range so wide, that the book is necessarily quite superficial. It is excellent for a quick introduction and overview, for instance for statisticians. The book is less suitable for research workers who want to apply the techniques in practice. The model range of the Saris and Stronkhorst book is quite surprising. It does not contain any models with latent variables, and it mentions their existence only briefly. Thus all models treated in the book are classical path models, in fact except for some examples in the section on identification they are all recursive. This is surprising, exactly because it implies that the program LISREL is not needed at all to fit the Saris and Stronkhorst models. A multiple regression program will do the job, and presumably much cheaper. Saris and Stronkhorst describe this as "the econometric approach" and reject it because it does not automatically generate a test of the model. Whether this is true or not depends on what you mean by "automatically". In any case this "disadvantage" seems rather far-fetched.

Of course avoiding latent variables avoids the "naming fallacy," discussed in the context of causal models by Cliff (1983). This fallacy has quite a long history in psychometrics, because factor analysts routinely "reify" their hypothetical constructs. In the LISREL-approach, in which each observed variable is interpreted as an indicator of a theoretical construct, presentation of the results is usually in terms of the constructs, which are named in suggestive ways, and treated as if they have independent existence and could be manipulated if necessary. There is a very brief warning in Everitt's book against this fallacy, Scott Long ignores it. Of course the long and troubled history of the most famous theoretical construct that was ever derived from a covariance matrix should be a warning, but the need for "interesting theory" seems stronger than methodological caution.

Statistics

Saris and Stronkhorst, and to a lesser extent also Scott Long, argue that LISREL can be used to rigorously test assumed models, and that this feature is perhaps its major

advantage over other approaches (and certainly over armchair theorizing). This is a simplification of the real situation, for various reasons. In the first place, as Guttman has repeatedly pointed out, using restrictive models as null hypotheses makes little sense. In the second place the tests are based on assumptions which are hardly ever true. A detailed analysis of the famous Blau and Duncan model in this respect has been published recently by Freedman (1983), with quite devastating results. Moreover, tests only examine the overidentification restrictions *within* the framework of independent identically distributed multinormal observations, and this superstructure itself does not make much sense in most cases. And, finally, we have already seen that LISREL is often used in an exploratory way. Thus "data snooping" and "chance capitalization" make significance tests even more debatable as they already are. This point has also been emphasized by Cliff (1983). Although significance tests and confidence intervals are certainly useful statistics, which give interesting information about at least one form of stability of the technique, they should not be used for inference purposes, certainly not routinely.

Both Everitt and Scott Long have very little statistical theory in their books. They introduce maximum likelihood estimation and chi-square testing without much ado, and also with little motivation. Saris and Stronkhorst give a useful introduction to some of the statistical ideas that are needed. Their treatment of more advanced statistical topics such as maximum likelihood estimation, chi square testing, Pitman power, generalized least squares, is not entirely beyond reproach, but this is perhaps unavoidable in a book of this type. Their treatment of asymptotic statistics repeats most of the current prejudices, but at the same time their emphasis on power calculations is both original and useful, as are their remarks about robustness. Saris and Stronkhorst do not succeed very well in their comparison of LISREL with log-linear analysis, the Simon-Blalock approach, and the econometric approach (whatever that is). These comparisons could be made far more interesting by treating the models as statements about conditional independence, and by placing less emphasis on the "theoretical" and "inferential" aspects of the technique, which are mostly illusory anyway.

Summary

If you want to smooth your covariance or correlation matrix by fitting a restrictive path model, or if you want to investigate whether certain partial correlation coefficients are small, then you can do this by using LISREL. The program is very well documented, it is a useful omnibus-approach, and it has its own language and folkways which are currently quite popular and successful. If you want to use LISREL, you certainly need Jöreskog and Sörbom (1984), and it will help you a great deal to consult Saris and Stronkhorst if you are not too familiar with mathematical statistics and matrix algebra. If you only want to know about LISREL, you could borrow Everitt's book. The books by Scott Long are somewhere in between the two, and it is not clear for which audience they are really intended. For relatively unsophisticated users they are a bit compact and tangential, for statisticians they are inferior to Everitt's book. They are, of course, less expensive than the Saris and Stronkhorst book, and they do not have the peculiarity that they omit latent variables almost completely. Because they pay relatively little attention to path models with observed variables, the two books by Scott Long could be used as a sequel to Saris and Stronkhorst. The three books taken together give a 500 page extensive introduction to the LISREL approach, for less than $25.

But, much more important than these conditional recommendations, is the advice by Cliff (1983) not to suspend your normal critical facilities because of the apparent computational and statistical sophistication of LISREL and related approaches. In most social science situations the true problem is model selection (not necessarily from the LISREL class), and conventional statistical theory has preciously little to say about that problem.

It is a serious mistake to assume (or even to teach) that this problem disappears or becomes less serious when it is computerized. What we need, by now, is a book about the limitations of LISREL and the other correlational approaches. Books which sell techniques and programs to audiences who would be better off with methods for computing and handling relatively simple descriptive statistics cannot be recommended wholeheartedly.

UNIVERSITY OF LEIDEN *Jan de Leeuw*

References

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research, 18,* 115–128.

Freedman, D. A. (1983). *Structural-equation models: A case study* (Technical Report No. 22). Berkeley: University of California, Department of Statistics.

Goldberger, A. S. (1983). Review of *Causal analysis: Assumptions, models and data* by Lawrence J. James, Stanley A. Mulaik, and Jeanne M. Brett. *Contemporary Psychology, 28,* 858–859.

Guttman, L. (1977). What is not what in statistics. *The Statistician, 26,* 81–107.

James, L. J., Mulaik, S. A., & Brett, J. M. (1982). *Causal analysis: Assumptions, models, and data.* Beverly Hills: Sage.

Jöreskog, K. G., & Sörbom, D. (1984). *LISREL VI. Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods.* Mooresville, IN: Scientific Software.

Kenny, D. A. (1979). *Correlation and Causality.* New York: Wiley.

Kiiveri, H., & Speed, T. P. (1982). Structural analysis of multivariate data: A review. In S. Leinhardt (Ed.), *Sociological Methodology 1982.* San Francisco: Jossey-Bass.

Ling, R. F. (1982). Review of *Correlation and Causality* by David A. Kenny. *Journal of the American Statistical Association, 77,* 489–491.