DOCUMENT RESUME

ED 271 504 TM 860 413

AUTHOR	Ethington, Corinna A.
TITLE	The Robustness of LISREL Estimates in Structural Equation Models with Categorical Variables.
PUB DATE	Apr 86
NOTE	32p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Computer Software; *Correlation; Estimation (Mathematics); Goodness of Fit; Hypothesis Testing; *Least Squares Statistics; Mathematical Models; *Matrices; *Maximum Likelihood Statistics; *Path Analysis; Statistical Bias; Statistical Studies
IDENTIFIERS	Dichotomous Variables; *LISREL Computer Program; Pearson Product Moment Correlation; *Robustness

ABSTRACT

This study examined the effect of type of correlation matrix on the robustness of LISREL maximum likelihood and unweighted least squares structural parameter estimates for models with categorical manifest variables. Two types of correlation matrices were analyzed; one containing Pearson product-moment correlations and one containing tetrachoric, polyserial, and product-moment correlations as appropriate. Using continuous variables generated according to the equations defining the population model, three cases were considered by dichotomizing some of the variables with varying degrees of skewness. When Pearson product-moment correlations were used to estimate associations involving dichotomous variables, the structural parameter estimates were biased when skewness was present in the dichotomous variables. Moreover, the degree of bias was consistent for both the maximum likelihood and unweighted least squares estimates. The analysis of mixed matrices produced average estimates that more closely approximated the model parameters except in the case where the dichotomous variables were skewed in opposite directions. (Author/BS)

THE ROBUSTNESS OF LISREL ESTIMATES IN STRUCTURAL EQUATION MODELS WITH CATEGORICAL VARIABLES

Corinna A. Ethington

University of Illinois at Chicago

This paper was prepared for delivery at the annual meeting of the American Educational Research Association, San Francisco, April 16-20, 1986.

THE ROBUSTNESS OF LISREL ESTIMATES IN STRUCTURAL EQUATION MODELS WITH CATEGORICAL VARIABLES

ABSTRACT

This study was an examination of the effect of type of correlation matrix on the robustness of LISREL maximum likelihood and unweighted least squares structural parameter estimates for models with categorical manifest variables. Two types of correlation matrices were analyzed; one containing Pearson product-moment correlations and one containing tetrachoric, polyserial, and product-moment correlations as appropriate. Using continuous variables generated according to the equations defining the population model, three cases were considered by dichotomizing some of the variables with varying degrees of skewness. When Pearson product-moment correlations were used to estimate associations involving dichotomous variables, the structural parameter estimates were biased when skewness was present in the dichotomous variables. Moreover, the degree of bias was consistent for both the maximum likelihood and unweighted least squares estimates. The analysis of mixed matrices produced average estimates that more closely approximated the model parameters except in the case where the dichotomous variables were skewed in opposite directions.

THE ROBUSTNESS OF LISREL ESTIMATES IN STRUCTURAL EQUATION MODELS WITH CATEGORICAL VARIABLES

The use of models to represent and explain phenomena is pervasive in every field of study and, in the social sciences, a class of models that has become widely applied is the structural equation model. The seminal works of Joreskog (1970) and Goldberger (1971) led to the first general structural equation model, the ACOVS model (Joreskog, 1970, 1973b; Joreskog, Gruvaeus, and van Thillo, 1970), followed by a more general model that has become known as the Lisrel model (Joreskog, 1973a). Several models incorporating different sets of assumptions defining a specific covariance or moment structure have subsequently appeared in the literature (e.g., Bentler, 1982, 1983; Bentler and Weeks, 1980; Browne, 1974, 1982; Lohmoller, 1981; McDonald, 1978; Muthen, 1979, 1983b, 1984; Wold, 1980, 1982). Nevertheless, the public availability of the computer program, LISREL (Joreskog and van Thillo, 1972), now in its sixth edition (Joreskog and Sorbom, 1983), has resulted in the Lisrel model becoming the most widely applied. Substantive examples of the Lisrel model can be found in the psychology, sociology, economics, and education literature, and reviews of the applied literature by Bentler (1980) and Bielby and Hauser (1977) contain hundreds of references.

The Lisrel model consists of two parts, the measurement model defined by:

$$x = \Lambda_x \xi + \delta$$

$$y = \Lambda_y \eta + \epsilon$$

and the structural model defined by:

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

The measurement model specifies how the observed variables, x and y, are determined through Λ_x and Λ_y by the latent variables, ξ and η, respectively; the δ and ϵ terms represent residuals in x and y unexplained by ξ and η. The structural model specifies the causal relationships among the latent endogenous variables in β, between the exogenous and endogenous variables in Γ, and describes unexplained residuals of the latent factors in ζ (Joreskog and Sorbom, 1983). The elements of β and Γ are regression coefficients resulting from the regression of the endogenous latent factors on their respective antecedent causal factors.

The only estimation procedure available in the earlier versions of LISREL was the maximum likelihood procedure. With the assumption that the observed variables have a multivariate normal distribution, this procedure uses a modification of an iterative minimization procedure described by Fletcher and Powell (1963) to fit the estimated covariance matrix implied by the model, Σ, to the sample covariance matrix, S. The estimates of the parameters of the hypothesized model are those values minimizing the maximum likelihood fitting function:

$$F = \log|\Sigma| + tr(S\Sigma^{-1}) - \log|S| - (p + q).$$

A likelihood-ratio chi-square statistic comparing Σ and S is computed, and gives an indication of the goodness of fit of the whole model.

The assumption of a multivariate normal distribution for the data is the weakest part of the LISREL program (de Leeuw, et al., 1983), for in many instances the distributions of variables are unknown or suspected to be far from normal. Consequently, the fifth edition of LISREL (Joreskog and Sorbom, 1981) included generalized least squares and unweighted least squares estimation procedures in addition to the maximum likelihood. These limited information procedures minimize the general fitting function:

$$F = (S - \sigma(\theta))' W^{-1} (S - \sigma(\theta))$$

where the weight matrix, W, is a consistent estimator of the asymptotic covariance matrix of S. In the unweighted least squares procedure, W = I, and the procedure reduces to an iterative ordinary least squares method. Unfortunately, the chi-square statistic and standard errors of estimates are not available with these procedures.

While the robustness of more traditional statistical procedures against violations of assumptions has often been tested, methods applied in the analysis of covariance structures have only recently begun to be examined for robustness (Boomsma, 1982, 1983; Browne, 1982; Huba and Bentler, 1983; Huba and Harlow, 1984; Joreskog and Goldberger, 1972; Muthen, 1978, 1983a; Muthen and Kaplan, 1984; Olsson, 1979; Tanaka, 1984). These robustness studies have primarily been concerned with the effects of non-normality as evidenced by the inclusion of dichotomous or ordered polychotomous variables with assumed underlying continuities as indicator variables in Lisrel-type models. These studies have examined the impact of choice of correlation type in factor analysis models and differences among various estimation procedures for both factor analysis models and structural equation models.

The studies of factor analysis models containing categorical variables have implications concerning the measurement portion of Lisrel-type models. Using their own computer programs, Muthen (1983a) and Olsson (1979) examined the impact of correlation type and found that the analysis of Pearson product-moment matrices resulted in downwardly biased estimates of the factor loadings of the categorical variables and inflated values for the chi-square goodness-of-fit statistics. Muthen (1983a) found that the use of tetrachoric and polychoric correlations to measure associations among the variables produced more robust parameter estimates and better fitting models. As for the estimation of Lisrel models, these results suggest that the reliabilities of categorical variables would be underestimated with product-moment correlations and the models too often rejected for lack of fit.

In studies of the robustness of various estimation procedures (including some not necessarily found in LISREL) for factor analysis models containing categorical variables (Boomsma, 1982, 1983; Muthen, 1978; Muthen and Kaplan, 1984; Olsson, 1979; Tanaka, 1984), the maximum likelihood (ML), generalized least squares (GLS), categorical variable methodology (CVM), and asymptotically distribution free (ADF) procedures were all found to perform well for data that did not deviate too drastically from normality. In cases of extreme skewness, distortions in the ML and GLS chi-squares and standard errors were found; the chi-squares were inflated and the standard errors were biased downwards. However, the parameter estimates were generally unbiased for large sample sizes (N > 400). An exception was found by Tanaka (1984) wherein both the ML and ADF were found to underestimate not only the standard errors but the factor loadings as well. When the parameters of the model were seen to hold for the underlying continuous variables, only the CVM estimation produced robust results.

The case studies by Browne (1982), Huba and Bentler (1983), Huba and Harlow (1984), and Joreskog and Goldberger (1972) provided comparisons of the ML, GLS, and ADF estimators in factor analyses of categorical variables. No large differences were found among estimated factor loadings for the procedures, although Browne (1982) and Joreskog and Goldberger (1972) reported that in general the GLS and ADF parameter estimates were lower than those obtained with ML. Moreover, the ADF chi-square was consistently lower than either the GLS or ML, and the estimated standard errors were generally smaller for GLS and ML.

Results from the above studies suggest in terms of the measurement portion of the Lisrel model, that the choice of estimation procedure may not be as important as the choice of correlation type when the concern is with the robustness of parameter estimates. However, for valid hypothesis testing and assessment of fit, ADF or CVM seem to be the preferred estimation procedures.

The extension of these kinds of studies to the structural portion of Lisrel-type models has just begun. To date there has been only one study of the robustness of the structural parameter estimates against non-normality (Boomsma, 1983). Using an adaptation of LISREL-III (Joreskog and Sorbom, 1976) which he called LISREP, Boomsma examined the robustness of the maximum likelihood estimates for models in which all indicator variables for both the exogenous and endogenous variables were categorical.

In the models estimated, Boomsma found no bias in the parameter estimates or estimates of standard errors. He did find both a categorization effect and a skewness effect on the standard deviations of the estimates. They were found to be generally too small for model variations with zero or small skewnesses, and generally too large for variations with moderate and large skewnesses. Thus, with increased skewness, on the average, the estimates are not too far from the population values, but for a single sample, the parameter estimates may deviate substantially. Boomsma also reported only a minor effect of categorization on the chi-square goodness-of-fit statistic, but with increased skewness the model was rejected too often. Thus, the number of categories of variables had less effect than the skewness of the variables.

From the studies reviewed above, it appears that the LISREL maximum likelihood measurement parameter estimates are probably non-robust when skewness is present in the observed variables and product-moment correlations are used to estimate the model. What is not known at the present is the influence of the choice of correlation type on LISREL latent variable structural parameter estimates, or the robustness of the LISREL limited information procedures. The purpose of this study is to examine the effect of correlation type on the robustness of LISREL maximum likelihood and unweighted least squares estimates of the structural parameters of models containing categorical variables as indicators of latent factors.

METHODOLOGY

Ideally, this study would examine the robustness of the three estimation procedures in LISREL (maximum likelihood, generalized least squares, and unweighted least squares) by the analysis of covariance matrices. However, there are several restrictions placed on the methodology used in this study by the limitations of the LISREL-VI program. LISREL does not provide estimates of variances and covariances corresponding to the tetrachoric, polychoric, and polyserial correlations; thus, only standardized estimates are reported since this study was of necessity restricted to the analysis of correlation matrices. Furthermore, the weight matrix used in the generalized least squares procedure assumes the input of a Pearson product-moment matrix and thus precludes the use of the LISREL generalized least squares procedure to analyze a matrix including tetrachoric or polyserial correlations. As a result, analyses reported here employed only the ML and ULS procedures. Finally, the technical and computational problems of computing the chi-square statistic and the standard errors of parameter estimates have not been solved for the ULS procedures in LISREL. Consequently, comparisons between procedures and correlation type can only be made for the beta and gamma parameter estimates and not their standard errors or the chi-square goodness-of-fit statistics.

This study was conducted using data generated to fit the equations defining the Lisrel model shown in Figure 1. Fifty sets of data were generated using a SAS PROC MATRIX (Sas Institute, 1982) program, each set containing 500 observations on ten standard normal variables. These variables were subsequently dichotomized and represent the continuum underlying the dichotomized variables. The average correlations among these continuous variables are given in Table 1.

Insert Figure 1 About here

Insert Table 1 About Here

The robustness of the estimates of the model parameters was then examined for three cases of non-normality determined by the dichotomization of the manifest indicators of η_1 and η_2 with varying degrees of skewness. In the first case, the skewness of each variable was zero (50% of the data were in each category); in the second case, each variable had a skewness of 1.5 (80% vs. 20%); and in the third case, the indicators of η_1 had a skewness of 1.5 (80%

vs. 20%) and the indicators of η_2 had a skewness of -1.5 (20% vs. 80%). Two types of correlation matrices were computed for each case: one consisting only of Pearson product-moment correlations, and one consisting of tetrachoric, polyserial, and product-moment correlations as appropriate. Table 2 gives the averages of each type of correlation matrix for each of the three cases. It is evident, by comparing these matrices to the matrix of correlations among the continuous variables given in Table 1, that in every case the mixed matrix more closely approximated the correlations of Table 1 while the product-moment matrix underestimated the associations involving the dichotomous variables. The attenuation in the product-moment correlations was expected to be manifested by an underestimation of the structural parameter estimates, and the mixed matrix was expected to produce estimates more closely approximating the model parameters.

Insert Table 2 About Here

The quality of the data generation procedure can be checked by examining the results of the analysis of the continuous data, since any deviations of these estimates from the model parameters are due to the randomness involved in the data generation process. Statistics for the structural parameter estimates using these data are shown in Table 3. The maximum likelihood chi-square statistic gives an indication of how well the generated data fit the model. For each of the 50 estimations of the model, this statistic was compared with the critical value of $\chi^2_{.05;25} = 37.6526$. With an alpha level of .05, probability theory indicates that for the 50 data sets used to estimate the model, 2.5 rejections would occur. Since there were only two rejections, overall the data fit the model very well. There is a slight inflation in the mean square errors of the estimates for β_{31} and β_{32} indicating some bias in these estimates. The average estimate for β_{31} overestimates the parameter while β_{32} is underestimated. The bias present in these estimates is probably a function of the data generation procedure and the use of only 50 replications, but should be kept in mind in the interpretation of subsequent results.

Insert Table 3 About Here

RESULTS

Tables 4, 5, and 6 give the results of the analyses for each of the three cases of dichotomization. The likelihood ratio chi-square statistic shows little effect of skewness or categorization in terms of number of rejections for lack of fit when product-moment matrices are analyzed, although the rejection of 4 of the 50 estimations of the model in Case 1 is slightly more than the expected 2.5. This is in contrast to previous studies (e.g., Boomsma, 1983; Muthen, 1983a; Olsson, 1979) where inflation was found in the maximum likelihood chi-square for increasing skewness. However, the empirical distribution of the chi-square statistic does appear to be affected. For the model used in this study, this statistic is distributed as χ^2 with 25 degrees of freedom. In Cases 1 and 2, the variance is slightly larger than expected, and in Case 3, both the mean and variance are considerably smaller than expected.

Insert Tables 4, 5, and 6 About Here

The highly inflated chi-squares occurring in the analyses of the mixed matrices cannot be compared with previous studies that calculated goodness-of-fit measures without using LISREL, for in LISREL this statistic is computed on the assumption that the correlations used in the analysis are product-moment correlations. Based on these chi-squares, the population model was rejected 45, 49, and 48 times for Cases 1, 2, and 3, respectively. This should be an important consideration for researchers estimating substantive Lisrel models for it indicates that the rejection of an estimated model using tetrachoric or polyserial correlations would probably not be a function of the data not fitting the model, but of the type of correlation used in the analysis. In any event, the user would not know the cause of rejection.

Case 1

The first case of non-normality examined the effect of categorization only. The manifest indicators for the two independent endogenous variables, η_1 and η_2, were dichotomized with fifty percent of the cases in each category, giving a symmetrical distribution with no skewness. When the matrix used in the analysis is the mixed matrix of tetrachoric, polyserial, and product-moment correlations, both the average maximum likelihood (ML) and unweighted least squares (ULS) estimates are robust against categorization (see Table 4). The analysis of product-moment matrices generally underestimated the model parameters with no substantial differences between the ML and ULS estimates. However, the bias in these estimates does not appear to be too serious since there is only a slight inflation of the mean square errors. The exceptions are the average estimates of β_{31} and β_{32}, and the direction of bias is opposite of that found in the analysis of the continuous data.

Case 2

The second case estimated the model with each of the manifest indicators of η_1 and η_2 dichotomized with a skewness of 1.5 (80% in the first category and 20% in the second). Results here show the effects of correlation type and estimation procedure on the parameter estimates when each of the dichotomized variables are skewed in the same direction.

The underestimation of the associations involving these variables by the product-moment correlations was greater here than in Case 1 (see Table 2). The impact of this underestimation is evident in the bias that appears in the estimates resulting from the analyses of product-moment matrices (see Table 5). Again there are no substantial differences between the ML and ULS estimates. With both the ML and ULS procedures, each of the parameters were underestimated with the exception of γ_{32}, leading to a general underestimation of the magnitude of the influence of one latent variable on another.

The analyses of matrices containing tetrachoric, polyserial, and product-moment matrices result in more robust parameter estimates, regardless of estimation procedure. While three of the estimates did exhibit a slight inflation in their mean square errors (β_{31}, β_{32}, and γ_{21}), the bias in the estimates of the two beta parameters was consistent with that appearing when the continuous data was analyzed.

Case 3

The final case of non-normality considered in this study had the manifest indicators of η_1 (y_1 and y_2) and of η_2 (y_3 and y_4) skewed in opposite directions. The skewness for the indicators of η_1 was set at 1.5 (80% vs. 20%) and for η_2 was set at -1.5 (20% vs. 80%). The attenuation in the product-moment correlations involving these variables is quite large in this case (see Table 2). For example, the correlation between y_2 and y_3 measured as continuous variables is .553, but the product-moment estimate of this correlation after the variables are dichotomized and skewed in opposite directions is only .207 while the tetrachoric estimate is .546. The large attenuation in the correlations was expected to substantially affect the parameter estimates, and the analyses of the mixed matrices were expected to produce more robust estimates. In the event, however, the results were not as expected.

When product-moment matrices were analyzed, as expected, the beta parameter estimates were severely biased, and the degree of bias was consistent for both the ML and ULS estimation procedures (see Table 6). The estimates for the gammas, however, did not exhibit the same increase in bias. In fact, they were no more biased than the estimates in Case 1 where the dichotomized variables each had no skew and the zero-order associations between the variables were much better estimated.

Turning to the estimates resulting from the analyses of the mixed matrices, substantial bias was also found in both the ML and ULS estimates of many of the parameters. Only four of the nine parameter estimates (β_{21}, γ_{11}, γ_{22}, and γ_{31}) did not show substantial bias, and, surprisingly, the estimates of γ_{21} and γ_{32} were extremely biased while the corresponding product-moment estimates were not. This was not expected, for the closer approximation of the correlations between the underlying variables was expected to result in more robust estimates from the analyses of the mixed matrices than from the product-moment analyses. Evidently the extreme skewness in opposite directions that was present among some of the variables was not entirely compensated for by the closer approximations of the correlations involving these variables by the tetrachoric and polyserial correlations. In addition, the estimates from the analyses of the mixed matrices exhibited a much greater variability. The increasing variability of the estimates for increasing skewness has also been reported in the studies of Boomsma (1983) and Muthen and Kaplan (1984).

CONCLUSIONS

While it is noted that the results of studies of the effect of non-normality in structural equation modeling are not completely independent of the models used in the studies, they do pose potentially serious problems for the applied researcher. From the results reported here, a general qualitative conclusion can be drawn: the analysis of product-moment correlation matrices resulted in biased estimates of the structural parameters of the model used in this study when skewness was present in the dichotomous variables. Moreover, the degree of bias was consistent for both the maximum likelihood and unweighted least squares procedures. With the exception of Case 3, the analysis of mixed matrices produced average estimates that more closely approximated the model parameters.

In cases where the categorical variables have approximately symmetric distributions, there are choices to be made. A conservative approach would be to analyze product-moment matrices with the maximum likelihood procedure. The parameter estimates exhibit little bias and the maximum likelihood chi-square statistic appears generally reliable. Additionally, both standardized and unscaled estimates may be obtained along with standard errors. The use of tetrachoric or polyserial correlations produce unbiased parameter estimates on the average, but the lack of standard errors and goodness-of-fit measures detract from their use.

At present, researchers do not have alternative estimation procedures available for the analysis of latent variable structural equation models. Programs such as EQS by Bentler (1982) and LISCOMP (described as LACCI in Muthen, 1984) that employ distribution-free estimation procedures that consider the skewness and kurtosis of the variables hold promise for the future and should be available for public distribution shortly. Until that time, researchers are ill-advised to employ LISREL in the estimation of models containing skewed categorical manifest variables. Perhaps the best recommendation for researchers who wish to use LISREL to estimate latent variable structural equation models is to avoid the use of categorical variables, particularly dichotomies. Careful consideration of the nature of the data to be analyzed is of utmost importance. Consideration given to how the variables are to be measured while the research is in the conceptual stage would avoid the problems that have been identified in this study. Great care should be taken to use interval scales of measurement. However, this does not assure that the data gathered from a sample will have a multivariate normal distribution. The impact of skewness in continuous distributions on LISREL estimates is thus of great importance, and future research should address this question.

REFERENCES

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. Annual Review of Psychology, 31, 419-456.

Bentler, P. M. (1982). Linear simultaneous equation systems with multiple levels and types of latent variables. In K. G. Joreskog and H. Wold (Eds.), Systems under indirect observation: Causality, structure, prediction. Amsterdam: North-Holland.

Bentler, P. M. (1983). Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. Psychometrika, 48, 493-517.

Bentler, P. M., and Weeks, D. G. (1980). Linear structural equations with latent variables. Psychometrika, 45, 289-308.

Bielby, W. T., and Hauser, R. M. (1977). Structural equation models. Annual Review of Sociology, 3, 137-161.

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. S. Joreskog and H. Wold (Eds.), Systems under indirect observation: Causality, structure, prediction. Amsterdam: North-Holland.

Boomsma, A. (1983). On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality. Unpublished doctoral dissertation, University of Groningen, The Netherlands.

Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariance structures. South African Statistical Journal, 8, 1-24.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), Topics in applied multivariate analysis. Cambridge: Cambridge University Press.

de Leeuw, J., Keller, W. J., and Wansbeek, T. (1983). Editor's introduction. Journal of Econometrics, 22, 1-12.

Fletcher, R., and Powell, M. J. D. (1963). A rapidly convergent descent method for minimization. The Computer Journal, 6, 163-168.

Goldberger, A. S. (1971). Econometrics and psychometrics: A survey of communalities. Psychometrika, 36, 83-107.

Huba, G. J., and Bentler, P. M. (1983). Test of a drug use causal model using asymptotically distribution free methods. Journal of Drug Education, 13, 3-17.

Huba, G. J., and Harlow, L. L. (1984, April). Comparison of maximum likelihood, generalized least squares, and asymptotically distribution free parameter estimates in educational latent variable causal models. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Joreskog, K. G. (1970). A general method for analysis of covariance structures. Biometrika, 57, 239-251.

Joreskog, K. G. (1973a). A general method for estimating a linear structural equation system. In A. S. Goldberger and O. D. Duncan (Eds.), Structural Equation Models in the Social Sciences. New York: Seminar Press.

Joreskog, K. G. (1973b). Analysis of covariance structures. In P. R. Krishnaiah (Ed.), Multivariate Analysis - III. New York: Academic Press.

Joreskog, K. G., and Goldberger, A. S. (1972). Factor analysis by generalized least squares. Psychometrika, 37, 243-260.

Joreskog, K. G., Gruvaeus, G. T., and van Thillo, M. (1970). ACOVS - A general computer program for analysis of covariance structures. Princeton: Educational Testing Service Research Bulletin 70-15.

Joreskog, K. G., and Sorbom, D. (1976). Statistical models and methods for test-retest situations. In D. N. M. de Gruijter and L. J. Th. van der Kamp (Eds.), Advances in psychological and educational measurement. London: Wiley.

Joreskog, K. G., and Sorbom, D. (1981). LISREL: Analysis of linear structural relationships by maximum likelihood and least squares methods, user's guide, version V. Chicago: International Educational Services.

Joreskog, K. G., and Sorbom, D. (1983). LISREL VI: analysis of linear structural relationships by maximum likelihood and least squares methods. Chicago: National Educational Resources.

Joreskog, K. G., and van Thillo, M. (1972). LISREL: A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. Princeton: Educational Testing Service Research Bulletin 72-56.

Lohmoller, J. B. (1981). LVPLS 1.6 program manual: Latent variables path analysis with partial least-squares estimation. Forschungsbericht 81.04. Hochschule der Bundeswehr Munchen, Fachbereich Padagogik, Munich.

McDonald, R. P. (1978). A simple comprehensive model of the analysis of covariance structures. British Journal of Mathematical and Statistical Psychology, 31, 59-72.

Muthen, B. (1978). Contributions to factor analysis of dichotomous variables. Psychometrika, 43, 551-560.

Muthen, B. (1979). A structural probit model with latent variables. Journal of the American Statistical Association, 74, 807-811.

Muthen, B. (1983a, April). Categorical vs. continuous variables in factor analysis and structural equation modeling. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Muthen, B. (1983b). Latent variable structural equation modeling with categorical data. Journal of Econometrics, 22, 43-65.

Muthen, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika, 49, 115-132.

Muthen, B., and Kaplan, D. (1984, April). A Comparison of some methodologies for the factor analysis of non-normal likert variables. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Olsson, U. (1979). On the robustness of factor analysis against crude classification of the observations. Multivariate Behavioral Research, 14, 485-500.

SAS Institute Inc. (1982). SAS user's guide: Statistics 1982 edition. Cary, North Carolina: SAS Institute Inc.

Tanaka, J. S. (1984). Some results on the estimation of covariance structure models. Unpublished doctoral dissertation, University of California, Los Angeles.

Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares. In J. Kmenta and J. Ramsay (Eds.), Evaluation of econometric models. New York: Seminar Press.

Wold, H. (1982). Soft modeling· The basic design and some extensions. In K. G. Joreskog and H. Wold (Eds.), Systems under indirect observation: Causality, structure, prediction. Amsterdam: North-Holland.

Table 1

Average Correlations for Continuous Variables

	X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
X_1	1.000									
X_2	.697	1.000								
X_3	.345	.344	1.000							
X_4	.337	.335	.802	1.000						
Y_1	.115	.117	.108	.110	1.000					
Y_2	.124	.128	.120	.120	.664	1.000				
Y_3	.258	.261	.323	.318	.519	.553	1.000			
Y_4	.234	.234	.288	.282	.472	.509	.753	1.000		
Y_5	.068	.069	.214	.213	.527	.566	.504	.458	1.000	
Y_6	.065	.067	.204	.209	.506	.547	.479	.433	.731	1.000

Table 2

Average Correlations for Three Cases of Dichotomization[a]

	X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
					Case 1					
X_1	---	.697	.345	.337	.095	.094	.203	.187	.068	.065
X_2	.697	---	.344	.335	.092	.098	.210	.186	.069	.067
X_3	.345	.344	---	.802	.087	.101	.255	.232	.214	.204
X_4	.337	.335	.802	---	.089	.100	.257	.227	.213	.209
Y_1	.119	.115	.109	.112	---	.462	.342	.305	.419	.405
Y_2	.117	.123	.127	.125	.663	---	.366	.337	.456	.434
Y_3	.254	.263	.320	.322	.512	.544	---	.538	.405	.382
Y_4	.234	.232	.291	.284	.460	.504	.746	---	.371	.346
Y_5	.068	.069	.214	.213	.542	.571	.507	.464	---	.731
Y_6	.065	.067	.204	.209	.507	.545	.480	.434	.731	---
					Case 2					
X_1	---	.697	.345	.337	.079	.083	.179	.161	.068	.065
X_2	.697	---	.344	.335	.085	.088	.181	.161	.069	.067
X_3	.345	.344	---	.802	.075	.079	.234	.201	.214	.204
X_4	.337	.335	.802	---	.077	.079	.227	.196	.213	.209
Y_1	.113	.121	.107	.110	---	.419	.305	.270	.364	.351
Y_2	.118	.125	.113	.113	.657	---	.333	.297	.393	.381
Y_3	.253	.255	.332	.322	.511	.550	---	.506	.361	.339
Y_4	.229	.227	.285	.278	.462	.500	.750	---	.317	.302
Y_5	.068	.069	.214	.213	.520	.559	.516	.454	---	.731
Y_6	.065	.067	.204	.209	.503	.544	.485	.433	.731	---
					Case 3					
X_1	---	.697	.345	.337	.079	.083	.182	.163	.068	.065
X_2	.697	---	.344	.335	.085	.088	.183	.160	.069	.067
X_3	.345	.344	---	.802	.075	.079	.223	.202	.214	.204
X_4	.337	.335	.802	---	.077	.079	.222	.199	.213	.209
Y_1	.113	.121	.107	.110	---	.419	.201	.190	.364	.351
Y_2	.118	.125	.113	.113	.657	---	.207	.199	.393	.381
Y_3	.263	.266	.322	.322	.521	.546	---	.512	.351	.334
Y_4	.235	.232	.291	.288	.481	.517	.757	---	.323	.304
Y_5	.068	.069	.214	.213	.520	.559	.503	.463	---	.731
Y_6	.065	.067	.204	.209	.503	.544	.477	.435	.731	---

[a] Product-moment matrix is above the diagonal and mixed matrix below.

Table 3

Structural Parameter Estimates and Sampling Variability Using Continuous Variables

Parameter	ML Estimate	ML Variability*	ULS Estimate	ULS Variability*
β_{21} = .663	.667	.0013 .0013	.668	.0013 .0013
β_{31} = .669	.689	.0057 .0061	.691	.0059 .0064
β_{32} = .127	.109	.0071 .0074	.106	.0074 .0079
γ_{11} = .132	.134	.0034 .0034	.134	.0031 .0031
γ_{12} = .105	.097	.0034 .0035	.097	.0033 .0034
γ_{21} = .126	.115	.0016 .0017	.115	.0015 .0016
γ_{22} = .238	.234	.0017 .0017	.232	.0017 .0017
γ_{31} = -.153	-.155	.0025 .0025	-.153	.0024 .0024
γ_{32} = .188	.193	.0023 .0023	.194	.0024 .0024
x^2		24.1654		
Var (x^2)		52.1977		
Rejections		2		

*The entries under variability are variance and mean square error.

Table 4

Structural Parameter Estimates and Sampling Variability for Case 1 (50/50).

	Pearson Matrix				Mixed Matrix			
	ML		ULS		ML		ULS	
Parameter	Estimate	Variability*	Estimate	Variability*	Estimate	Variability*	Estimate	Variability*
β_{21} = .663	.626	.0030 .0044	.625	.0031 .0045	.661	.0027 .0027	.660	.0028 .0028
β_{31} = .669	.636	.0075 .0086	.638	.0077 .0087	.678	.0107 .0108	.679	.0113 .0114
β_{32} = 127	.146	.0119 .0122	.142	.0127 .0130	.131	.0168 .0168	.129	.0175 .0175
γ_{11} = .132	.120	.0047 .0048	.121	.0044 .0045	.125	.0051 .0052	.127	.0047 .0047
γ_{12} = .105	.102	.0054 .0054	.101	.0052 .0052	.106	.0061 .0061	.106	.0056 .0056
γ_{21} = .126	.116	.0022 .0023	.117	.0022 .0023	.120	.0025 .0026	.120	.0025 .0025
γ_{22} = .238	.219	.0031 .0034	.219	.0030 .0033	.231	.0038 .0039	.230	.0035 .0035
γ_{31} = -.153	-.146	.0037 .0038	-.145	.0035 .0036	-.152	.0042 .0042	-.152	.0039 .0039
γ_{32} = .188	.187	.0039 .0039	.188	.0038 .0038	.180	.0046 .0047	.182	.0044 .0044
x^2		24.8112				57.5588		
Var (x^2)		63.9152				445.2100		
Rejections		4				45		

*The entries under variability are variance and mean square error.

Table 5

Structural Parameter Estimates and Sampling Variability for Case 2 (80/20).

Parameter	Pearson Matrix				Mixed Matrix			
	ML		ULS		ML		ULS	
	Estimate	Variability*	Estimate	Variability*	Estimate	Variability*	Estimate	Variability*
β_{21} = .663	.614	.0069 .0093	.612	.0067 .0093	.665	.0063 .0063	.665	.0058 .0058
β_{31} = .669	.599	.0143 .0192	.599	.0138 .0187	.693	.0238 .0244	.697	.0216 .0223
β_{32} = .127	.111	.0133 .0136	.111	.0127 .0130	.111	.0266 .0268	.107	.0235 .0239
γ_{11} = .132	.120	.0056 .0058	.120	.0056 .0057	.135	.0070 .0070	.136	.0069 .0069
γ_{12} = .105	.078	.0055 .0062	.080	.0054 .0061	.089	.0069 .0071	.092	.0069 .0071
γ_{21} = .126	.091	.0044 .0056	.093	.0040 .0051	.102	.0064 .0070	.103	.0056 .0062
γ_{22} = .238	.213	.0026 .0032	.212	.0026 .0033	.247	.0034 .0035	.244	.0035 .0035
γ_{31} = -.153	-.133	.0044 .0048	-.130	.0042 .0047	-.158	.0061 .0061	-.157	.0057 .0057
γ_{32} = .188	.217	.0034 .0042	.216	.0036 .0043	.196	.0053 .0054	.197	.0053 .0054
χ^2		25.9206				78.3846		
Var (χ^2)		59.3208				689.6611		
Rejections		2				49		

*The entries under variability are variance and mean square error.

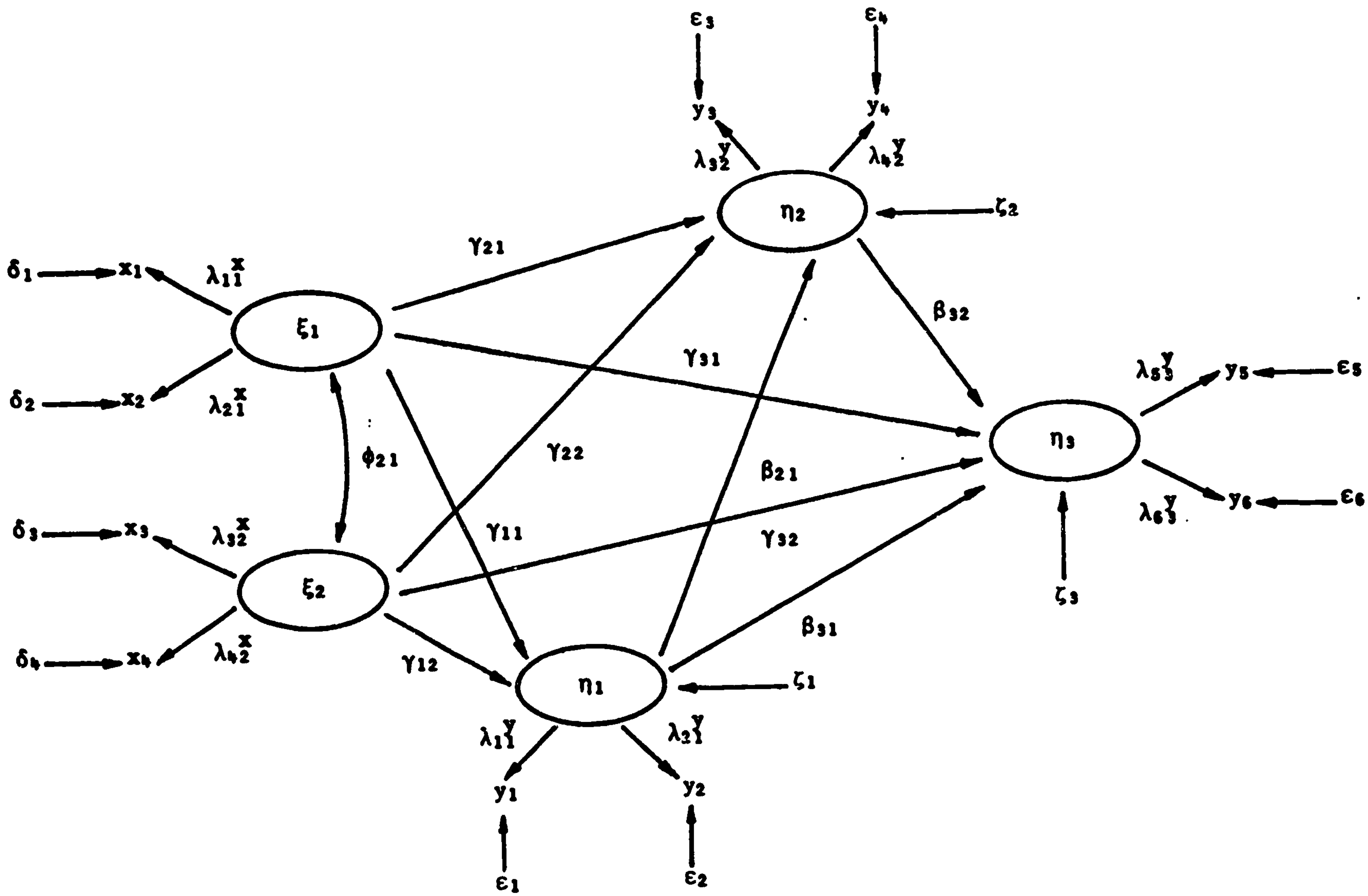
Table 6

Structural Parameter Estimates and Sampling Variability for Case 3 (80/20 vs. 20/80).

	Pearson Matrix				Mixed Matrix			
	ML		ULS		ML		ULS	
Parameter	Estimate	Variability*	Estimate	Variabilty*	Estimate	Variability*	Estimate	Variabilty*
β_{21} = .663	.381	.0030 .0825	.384	.0031 .0811	.690	.0122 .0129	.684	.0115 .0119
β_{31} = .669	.557	.0053 .0179	.556	.0052 .0180	.787	.0571 .0709	.774	.0519 .0629
β_{32} = .127	.283	.0042 .0285	.281	.0041 .0279	-.011	.0685 .0875	.010	.0648 .0785
γ_{11} = .132	.122	.0058 .0059	.121	.0056 .0057	.139	.0077 .0077	.137	.0069 .0069
γ_{12} = .105	.077	.0056 .0064	.079	.0054 .0061	.083	.0072 .0077	.091	.0069 .0071
γ_{21} = .126	.126	.0042 .0042	.127	.0042 .0042	.103	.0098 .0104	.111	.0082 .0084
γ_{22} = .238	.224	.0034 .0036	.222	.0035 .0038	.242	.0061 .0061	.239	.0065 .0065
γ_{31} = -.153	- 158	.0030 .0030	-.156	.0029 .0029	-.148	.0080 .0080	-.141	.0097 .0099
γ_{32} = .188	.179	.0032 .0033	.178	.0034 .0035	.228	.0088 .0104	.222	.0078 .0089
χ^2		22.7272				103.2198		
Var (χ^2)		36.2031				1983.9096		
Rejections		0				48		

*The entries under variability are variance and mean square error.

Figure 1. Structural Equation Model Used for Data Generation


ξ_1
ξ_2
η_1
η_2
η_3
δ_1
δ_2
δ_3
δ_4
x_1
x_2
x_3
x_4
λ_{11}^x
λ_{21}^x
λ_{32}^x
λ_{42}^x
ϕ_{21}
γ_{21}
γ_{31}
γ_{22}
γ_{11}
γ_{12}
γ_{32}
β_{32}
β_{21}
β_{31}
ε_3
ε_4
y_3
y_4
λ_{32}^y
λ_{42}^y
ζ_2
λ_{53}^y
y_5
ε_5
λ_{63}^y
y_6
ε_6
ζ_3
ζ_1
λ_{11}^y
λ_{21}^y
y_1
y_2
ε_1
ε_2