DRAFT

A Supplement to the LISREL IV Computer Program

October 1979

Eugene School District 4J
Research, Development
and Evaluation Division
200 North Monroe Street
Eugene, Oregon 97402
(503) 687-3311

Co-authors:
Les Hendrickson
Jim Bavry

LISREL

This paper is intended to be a modest supplement to the LISREL User's Guide. It contains information which will be useful to the inexperienced program user. This supplement is limited to a discussion of simple models and does not discuss simultaneous group comparisons, longitudinal models, multitrait-multimethod matrices or testing the equality of covariance and correlation matrices. LISREL is an extraordinarily versatile and complex tool and no claim is made that this supplement covers even most analyses that LISREL can do. This supplement is meant to be read in conjunction with the User's Guide and is intended to introduce inexperienced users to basic program features. Both the supplement and the User's Guide presuppose a knowledge of regression analysis, calculus, matrix algebra and some mathematical statistics.

This paper briefly discusses methodological trends in the last decade. Then it presents a hypothetical LISREL model and describes its parts. The use of Jöreskog and Sörbom's exploratory factor analysis program (EFAP) is referred to in this discussion. The eight basic matrices in a LISREL program are discussed next. LISREL and ordinary least squares are compared. Then more LISREL concepts are introduced, followed by discussions of identification, specification, and what is a good fitting model. The supplement ends by discussing how model fit can be improved.

What is LISREL and Why Should I Know About It?

LISREL stands for linear structural relations. It is a tool; a set of mathematical statistical routines which are available in a computer program. These routines currently are the most sophisticated procedures known for analyzing the interrelationships between sets of variables that are thought

to be causally linked. This program supercedes the cruder procedure of multiple-regression based path analysis which had been used to establish weights for estimating the degree to which one variable influences another.

In the early and middle 1960's, in the United States, in methodological circles, there was an increasing concern with linear equations. Econometric techniques were becoming more sophisticated and sociologists such as Blalock were graphically, systematically examining measurement error and zero order and partial correlations in sets of variables.

The systematic sorting out of how partial correlations and beta weights worked in small variable sets given different casual assumptions took place throughout the middle 1960's. The technique of path analysis was introduced in sociology in the late 1960's by Duncan and popularized rapidly through Sociological Methodology and other means. Originally used in genetic biology by Sewall Wright in the 1920's and 1930's, path analysis used standardized beta weights to estimate the weights of relations within a variable set. The technique disciplines users to graphically portray which variables influence which other variables and to know the temporal ordering of them. The technique uses conventional multiple regression least squares procedures.

LISREL has at least six advantages over path analysis: 1) It always produces consistent estimates; 2) the variables in the equation system may be either directly observed variables or unmeasured hypothetical variables which are not observed but are related to observed variables; 3) the program allows for both structural errors in equations (also called residuals, or disturbances) and measurement errors in the observed variables,

(also called observational errors); 4) the program calculates a residual covariance matrix and measurement error covariance matrix; 5) based on the original data matrix and the estimated matrix the program produces a maximum likelihood chi-square value which provides an indication of how well the hypothesized relations of the user's model fit the original data, and; 6) finally, LISREL is computerized. Users with an interactive terminal can rapidly alter program specifications and examine resulting changes in output before printing results. It supercedes older computer programs such as ACOVS and COFAMM.

A Hypothetical LISREL Model

Part I. The Measurement Model and the Use of EFAP II to Determine It

Figure 1 shows a hypothetical LISREL model taken from page 43 of the User's Guide. The model contains both observed and unobserved independent and dependent variables, measurement and structural parameters, and error terms.

There are seven independent observed variables denoted by x's. Each observed independent variable has a measurement error associated with it. These error terms are denoted by δ, delta. Long (1976) briefly discusses the history of theoretical approaches for dealing with error. There are three unobserved independent variables denoted by ξ, ksi. These unobserved independent variables are also called factors, latent variables or hypothetical constructs.

There are four dependent observed variables denoted by y's. Each observed dependent variable has an error term associated with it. These error terms are denoted by ε, epsilon. There are two unobserved dependent variables or factors in the model shown. These unobserved dependent factors are denoted by η, eta. The lines or arrows depict which unobserved factors are measured by which observed variables. It is a LISREL convention

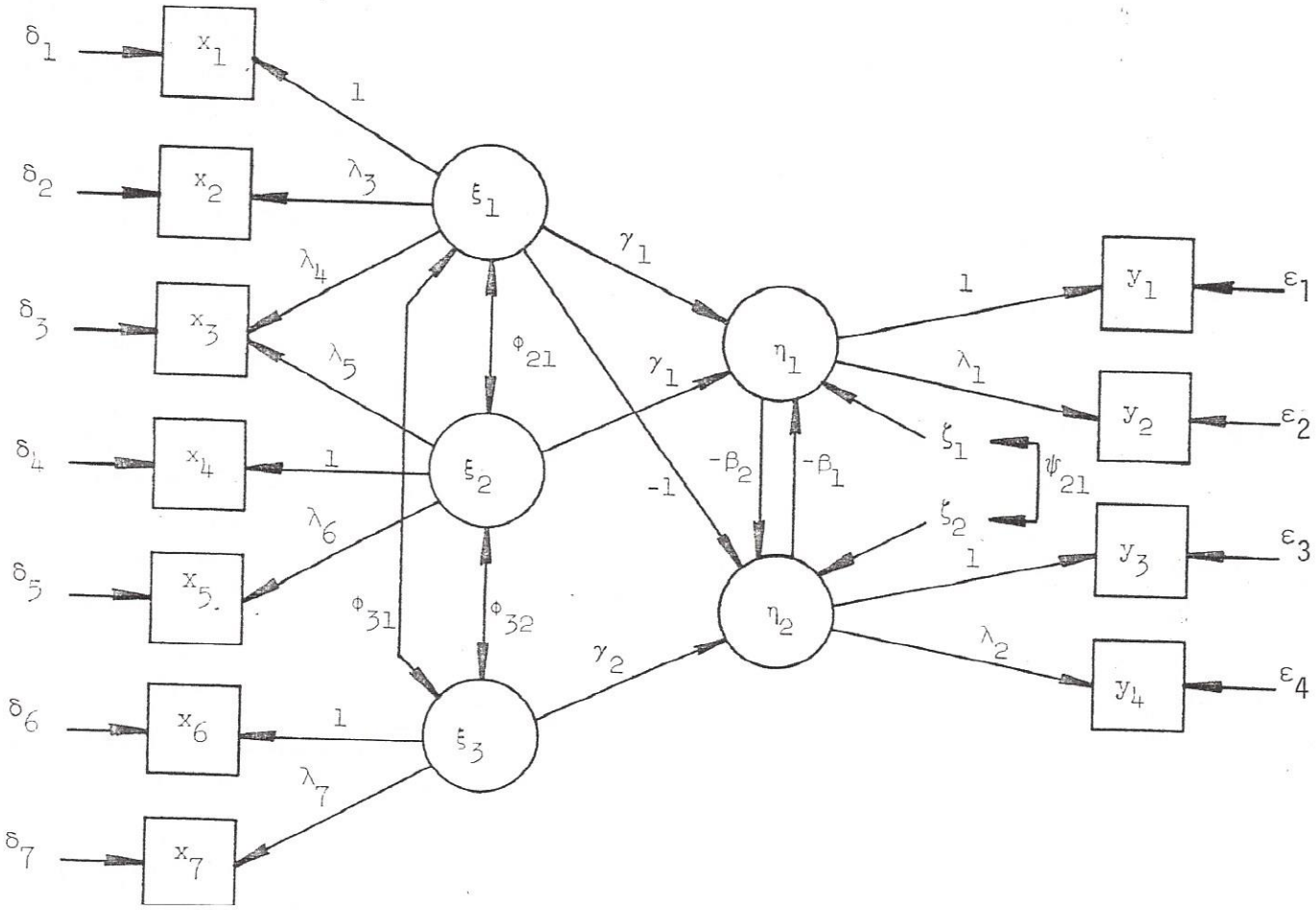

Figure 4: A Hypothetical Model Taken from p. 43 of the LISREL User's Guide.

that the directions of the arrows run from the unobserved to the observed variables. For example, x_1 and x_2 are a measure of ξ_1, x_3 is a measure of ξ_1 and ξ_2, and so on.

In Figure 1, the measurement coefficients linking the observed to the unobserved factors are denoted by λ, lambda. One of the arrow's (coefficients) leading from each unobserved factor is set equal to one in order to establish a metric or scale for unobserved factors. The number of arrows from the unobserved factors minus the number of unobserved factors equals the number of measurement coefficients that need to be estimated in this model. The general procedure for setting a measurement coefficient is to pick a variable you know something about. If you fix x_1, as in Figure 1, for example, it means a change of one unit in ξ_1 equals a one unit change in x_1.

The relations between the observed and unobserved variables are called the "measurement model". There are two possible measurement models: One for independent and one for dependent variables. These measurement models can be factor analytic models. The LISREL program presupposes that you know the relations between your observed and unobserved variables. Before using LISREL, it is often necessary to use a factor analysis program on your data if you have reason to believe that your observed variables may be tapping into different underlying factors. Jöreskog and Sörbom have available the Exploratory Factor Analysis Program (EFAP), now in its second version, EFAP II. This program is easy to use and is formatted similarly to LISREL.

Jöreskog recommends that you have at least three observed measures of each underlying factor and that the variables be measured on an interval scale. EFAP II provides four factor analytic methods to choose

understanding of how they apply to different models the user wishes to estimate is essential in order to properly use this complex program.

Formulas of the Structural and Measurement Models

The relationship between the unobserved independent and dependent variables (the structural model) is given by

$$\underset{\sim}{B}\underset{\sim}{\eta} = \underset{\sim}{\Gamma}\underset{\sim}{\xi} + \underset{\sim}{\zeta} \qquad (1)$$

In other words, five matrices are used to calculate this relation. In matrix form, (1) as applied to Figure 1, would be written:

$$\begin{bmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 & \gamma_1 & 0 \\ -1 & 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

In this example, the first row of the gamma matrix contains two γ coefficients of the same value.

There are two formulas for calculating the measurement models. One for the y variables and one for the x variables. These formulas are shown in (2) and (3).

$$\underset{\sim}{y} = \underset{\sim}{\Lambda}_{y}\underset{\sim}{\eta} + \underset{\sim}{\varepsilon} \qquad (2)$$

$$\underset{\sim}{x} = \underset{\sim}{\Lambda}_{x}\underset{\sim}{\xi} + \underset{\sim}{\delta} \qquad (3)$$

In (2) and (3) each observed variable is expressed as a function of the measurement coefficient times the unobserved factor plus an error of measurement term.

In matrix form, (2), as applied to Figure 1, would be written:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_1 & 0 \\ 0 & 1 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

As applied to Figure 1, (3) would be written:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \lambda_3 & 0 & 0 \\ \lambda_4 & \lambda_5 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_6 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_7 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \end{bmatrix}$$

Zeros are placed in $\underset{\sim}{\Lambda}_y$ and $\underset{\sim}{\Lambda}_x$ where there are no arrows linking the observed to unobserved variables. For example, there is no arrow linking y_1 to η_2, thus the first row, second column of $\underset{\sim}{\Lambda}_y$ has a zero in it. We also see that in the $\underset{\sim}{\Lambda}_x$ and $\underset{\sim}{\Lambda}_y$ matrices there is a one in each column and no row has all zeros.

Figure 2 presents the basic formulas for the measurement and structural formulas and then describes each of the matrices comprising the formulas. In order to use the LISREL program, users have to specify characteristics of each of these matrices. Are they symmetric, diagonal, zero, identity, full or fixed? Generally, one or two of these characteristics are necessary to describe each matrix to the LISREL program. The LISREL manual defines these terms.

Figure 3 provides a description of each of the eight matrices showing their Greek letters, the dimensions, the LISREL computer label for them and the possible characteristic forms these matrices can take.

Matrices $\underset{\sim}{\Lambda}_x$ and $\underset{\sim}{\Lambda}_y$ contain the measurement cefficients of the x and y measurement models. These matrices are used to identify the relations between the observed and unobserved variables and to set the scale of the unobserved variables.

When the dependent variables are thought to influence each other, the matrix $\underset{\sim}{B}$ is used to estimate the betas in Figure 1. This matrix indicates to LISREL what the analyst wishes in terms of estimating parameters indicating interdependencies among dependent variables.

The $\underset{\sim}{\Gamma}$, gamma, matrix contains the regression coefficients stating which unobserved factors influence which observed factors.

The $\underset{\sim}{\Phi}$, phi, matrix contains the variance and covariances of (un)observed independent variables. The diagonal elements are the variances of the factors and the off-diagonal elements are the covariances. The matrix is always symmetric. If your model contains no unobserved independent factors, then the $\underset{\sim}{\Phi}$ matrix is equal to the variance-covariance matrix of the observed independent variables. In LISREL terminology, this is called the fixed x case. If there are no unobserved independent factors, the parameters of phi do not have to be estimated iteratively because the variance-covariance matrix is the unbiased maximum likelihood estimator of phi.

The nonmeasurement or structural error part of the system is represented in LISREL notation through the matrix $\underset{\sim}{\Psi}$, psi. The diagonal elements of this matrix measure the variances in the errors of prediction in the structural model. Psi is generally assumed to be diagonal, implying the errors of prediction are uncorrelated and that off-diagonal elements $\underset{\sim}{\Psi}$ are zero. Because psi is a variance-covariance matrix, it is symmetric.

The two theta matrices $\underset{\sim}{\theta}_{\varepsilon}$ and $\underset{\sim}{\theta}_{\delta}$ contain the variance-covariances of the errors in the measurement models. Matrix $\underset{\sim}{\theta}_{\varepsilon}$ contains the errors in the dependent variable measurement model; $\underset{\sim}{\theta}_{\delta}$ the independent variable measurement errors. From LISREL's standpoint, it is not the errors per se that are estimated. It is their variance-covariance matrix that is important.

Thus, there are eight matrices whose form must be characterized and whose parameters must be specified. Four of the matrices, $\underset{\sim}{\Lambda}_x$ $\underset{\sim}{\Lambda}_y$, $\underset{\sim}{B}$ and $\underset{\sim}{\Gamma}$ contain coefficients; and four, $\underset{\sim}{\Phi}$, $\underset{\sim}{\Psi}$, $\underset{\sim}{\theta}_{\varepsilon}$, and $\underset{\sim}{\theta}_{\delta}$, are variance-covariance matrices.

DEPENDENT VARIABLE MEASUREMENT MODEL

$\underset{\sim}{y} = \underset{\sim}{\Lambda}_y \underset{\sim}{\eta} + \underset{\sim}{\varepsilon}$

$\underset{\sim}{y}$ is the (px1) vector of observed dependent variables

$\underset{\sim}{\Lambda}_y$ is a (pxm) matrix of dependent variable measurement coefficients

$\underset{\sim}{\eta}$ is a (mx1) vector of unobserved dependent factors

$\underset{\sim}{\varepsilon}$ is a (px1) vector of measurement errors in the observed dependent variables

$\underset{\sim}{\Theta}_\varepsilon$ is a (pxp) variance-covariance matrix of the errors in ε.

$\underset{\sim}{C}$ is a (mxm) variance-covariance matrix of the unobserved variables in $\underset{\sim}{\eta}$

INDEPENDENT VARIABLE MEASUREMENT MODEL

$\underset{\sim}{x} = \underset{\sim}{\Lambda}_x \underset{\sim}{\xi} + \underset{\sim}{\delta}$

$\underset{\sim}{x}$ is the (qx1) vector of observed independent variables

$\underset{\sim}{\Lambda}_x$ is a (qxn) matrix of independent variable measurement coefficients.

$\underset{\sim}{\xi}$ is a (nx1) vector of unobserved independent factors

$\underset{\sim}{\delta}$ is a (qx1) vector of measurement errors in the observed independent variables

$\underset{\sim}{\Theta}_\delta$ is a (qxq) variance-covariance matrix of the errors in $\underset{\sim}{\zeta}$

$\underset{\sim}{\Phi}$ is a (nxn) variance-covariance matrix of the latent (or original) variables in $\underset{\sim}{\xi}$

STRUCTURAL EQUATION MODEL

$\underset{\sim}{B}\,\underset{\sim}{\eta} = \underset{\sim}{\Gamma}\,\underset{\sim}{\xi} + \underset{\sim}{\zeta}$

$\underset{\sim}{B}$ is a (mxm) full rank matrix reflecting the dependencies among the dependent variables

$\underset{\sim}{\eta}$ is a (mx1) vector of unobserved dependent variables

$\underset{\sim}{\Gamma}$ is a (mxn) matrix of regression coefficients allowing for relationships specified or estimated using $\underset{\sim}{B}$

$\underset{\sim}{\xi}$ is a (mx1) vector of latent independent variables

$\underset{\sim}{\zeta}$ is a (mx1) vector of structural errors

$\underset{\sim}{\Psi}$ is a (mxm) variance-covariance matrix of the errors in $\underset{\sim}{\zeta}$

Where m = number of latent dependent variables
n = number of latent independent variables
p = number of observed dependent variables
q = number of observed independent variables

FIGURE 2 Summary of Sub-models and Selected Matrices Which May be Estimated in Whole or in Part in a LISREL run.

Matrix	Symbol	Dimension	Description	LISREL Label	Possible Form
Lambda X	$\underset{\sim}{\Lambda}_x$	NX by NKSI	factor loadings of x variables on exogenous KSI factors	LX	identity, diagonal, or full
Lambda Y	$\underset{\sim}{\Lambda}_y$	NY by NETA	factor loadings of y variables on endogenous ETA factors	LY	identity, diagonal or full
Beta	$\underset{\sim}{B}$	NETA by NETA	structural coefficients among latent endogenous ETA factors	BE	identity, diagonal, or full
Gamma	$\underset{\sim}{\Gamma}$	NETA by NKSI	structural coefficients from exogenous KSI factors to endogenous ETA factors	GA	identity, diagonal, or full
Phi	$\underset{\sim}{\Phi}$	NKSI by NKSI	covariances among the exogenous KSI factors	PH	identity, diagonal, or symmetric
Psi	$\underset{\sim}{\Psi}$	NETA by NETA	covariances among residuals in the structural model	PS	zero, diagonal, or symmetric
Theta Epsilon	$\underset{\sim}{\Theta}_\varepsilon$	NY by NY	covariances among the measurement errors of the y variables	TE	zero, diagonal, or symmetric
Theta Delta	$\underset{\sim}{\Theta}_\delta$	NX by NX	covariances among the measurement errors of the x variables	TD	zero, diagonal, or symmetric

KEY TO DIMENSIONS:
NKSI - number of latent exogenous factors (n)
NETA - number of latent endogenous factors (m) (see structural model)
NX - number of X indicators (observed variables) (q)
NY - number of Y indicators (observed variables) (p) (see measurement model)

STRUCTURAL MODEL

$\underset{\sim}{\beta}\underset{\sim}{\eta} = \underset{\sim}{\Gamma}\underset{\sim}{\xi} + \underset{\sim}{\zeta}$

$\underset{\sim}{\eta}$ - latent endogenous factors (ETA)
$\underset{\sim}{\xi}$ - latent exogenous factors (KSI)
$\underset{\sim}{\zeta}$ - latent residuals (ZETA)

MEASUREMENT MODELS

$\underset{\sim}{y} = \underset{\sim}{\Lambda}_y\underset{\sim}{\eta} + \underset{\sim}{\varepsilon}$
$\underset{\sim}{x} = \underset{\sim}{\Lambda}_x\underset{\sim}{\xi} + \underset{\sim}{\delta}$

$\underset{\sim}{x}, \underset{\sim}{y}$ - observed variables
$\underset{\sim}{\varepsilon}, \underset{\sim}{\delta}$ - errors of measurement

FIGURE 3 Description of Eight Basic LISREL Matrices. (Figure courtesy of Abt Associates, Boston, MA, 1979)

DISCUSSION OF SPECIFIC ISSUES

The purpose of this brief manual is to help inexperienced users learn how to use the LISREL program. Two major uncertainties that users encounter are: a) specifying uniquely estimable and meaningful parameters and b) setting starting values for the iterative model testing. These uncertainties arise because users have many options in the selection of which parameters to estimate and which starting values to set for each parameter.

The preceding discussion has introduced the eight basic matrices and an understanding of how they apply to the user's model will help the user with the first uncertainty cited above. While discussing how starting values can be assigned to the parameters of the users model, it is desirable to discuss specific issues. Hopefully, the discussion will provide information describing LISREL concepts and procedures that will make it easier for users to use the program. Specific issues to be covered are contrasts between LISREL and ordinary least squares, further introduction to LISREL concepts, the nature of sequential model fitting the identification and specification problems encountered by users, the kind of data used by the program, and when a model has a good fit.

LISREL and Ordinary Least Squares (OLS)

Many model fitting procedures can be used in exploring the fit of competing models to a data set. The Statistical Package for the Social Sciences (SPSS) Manual (Kin and Kohourt, 1975, page 345) lists four sequential methods for regression analysis and presents an example where blocks of variables are entered in a predetermined order. These procedures apply to LISREL IV models except that parameters rather than variables

are added. A major difference between LISREL and least-squares assumptions is that the LISREL analysis reproduces a covariance or correlation matrix rather than minimizing a sum of squares as in regression. Least squares regressions can be run on LISREL with the appropriate specification of paramaters.

The generality of the models fit by LISREL is exemplified in Jöreskog (1978). Jöreskog and his colleague, Sörbom, have created a tool capable of analyzing various models formed from given parameter configurations. Thus, the 1978 article shows the correspondence between given parameter configurations and the confirmatory factor analysis model, multitrait-multimethod models, regression models, measurement models and variations and combinations of these models.

Econometric and path analysis procedures can be seen as extensions of ordinary least squares regression models. In the ordinary least squares approach a dependent variable, say y, is expressed as a function of independent variables, say $\underset{\sim}{x}$, where $\underset{\sim}{x}$ is a vector of q predictor variables. The linear structural model is then given as $y = \underset{\sim}{\pi}\,\underset{\sim}{x} + \varepsilon$, where $\underset{\sim}{\pi}$ serves as a set of weights to be used in combining the variables in $\underset{\sim}{x}$ to predict a given observation, y.

In this model, the equation in terms of the variance of y in a population of observations is given by:

$$\sigma^2_y = \underset{\sim}{\pi}\,\underset{\sim\sim}{\Sigma xx}\,\pi' + \sigma^2_e$$

where σ^2_y = the variance of variable y in the population

$\underset{\sim}{\pi}$ = the regression parameters of y on the predictor variables

π' = the transpose of $\underset{\sim}{\pi}$

$\underset{\sim\sim}{\Sigma xx}$ = the covariance matrix of the predictor variables

σ^2_e = the variance of the errors in predicting y from $\underset{\sim}{\beta}\,\underset{\sim}{x}$

Suppose there are two independent variables, y_1 and y_2, say. Then the analyst might wish to examine a model wherein not only the individual variances of the dependent variables are reproduced, but also where the interrelationship between y_1 and y_2 is taken into account. However, in ordinary least squares the relation between y_1 and y_2 makes no difference. The ordinary least squares regressions will be the same no matter what the covariance between y_1 and y_2. A linear structural model as formulated for LISREL permits the analyst to allow for relationships between y_1 and y_2. If the analyst makes the standard assumption that errors in prediction are uncorrelated, then the interrelationship of y_1 and y_2 can be allowed for in the $\underset{\sim}{B}$ (BE) matrix of the LISREL model. In such a situation the analyst could interpret the off diagonal coefficient in $\underset{\sim}{B}$ as indicating causal relationships between y_1 and y_2.

Another way of comparing LISREL and ordinary least squares is to ask what is the difference between structural and ordinary regression parameters? Why is the claim made that LISREL generates causal or structural parameters? The general answer to these questions is that the LISREL procedures for generating effect parameters can more accurately describe the mechanisms generating the data. Structural parameters are not the same estimates as ordinary regression weights. There are at least three situations where LISREL parameters are a needed alternative to ordinary regression, where you have: measurement error, reciprocal causation, and omitted variables.

For example, consider the case of measurement error. Figure 4 describes two variables each influenced by a third and by measurement error.

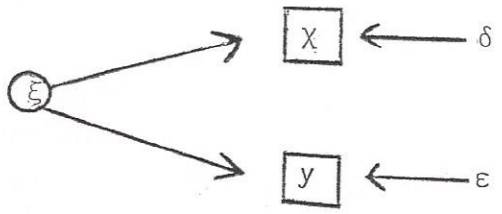

Figure 4 Three Variable Case with Measurement Error

Now $x = \xi + \delta$, a true score plus a measurement error

$y = \beta \xi + \varepsilon$, the true regression (4)

Assume ξ, δ, and ε are uncorrelated, then the covariance of x and y equals,

$$\underset{\sim}{\Sigma} = \begin{pmatrix} \delta^2_{\xi} + \delta^2_{\delta} & \\ \beta\delta^2_{\xi} & \beta^2\delta^2_{\xi} + \delta^2_{\varepsilon} \end{pmatrix} \qquad (5)$$

and $$\beta_{y.x} = \beta \frac{\delta^2_{\xi}}{\delta^2_{\xi} + \delta^2_{\delta}} \neq \beta \qquad (6)$$

In the case where measurement error is present, as shown in (6), the least squares regression weight is not equal to the structural parameter because of the existence of measurement error. Models of reciprocal causation and omitted variables could likewise be developed to demonstrate the difference between beta and the structural coefficient.

More LISREL Concepts

The following terms have been used in this manual before. However, they will be restated here to establish their usage explicitly. The term "measurement parameter" will be used to refer to: a) the measurement coefficients in $\underset{\sim}{\Lambda}_x$ and $\underset{\sim}{\Lambda}_y$, b) the elements in the covariance matrices of the measurement errors, $\underset{\sim}{\theta}_{\varepsilon}$ and $\underset{\sim}{\theta}_{\delta}$, and c) the elements in the covariance matrix of the latent variables in $\underset{\sim}{\Phi}$. The term "structural coefficient"

will be used to refer to parameters in $\underset{\sim}{B}$ and $\underset{\sim}{\Gamma}$. The term "structural error variance-covariance matrix" shall refer to $\underset{\sim}{\Psi}$. This nomenclature is consistent with the terms in Figure 2. Readers are cautioned that statistical nomenclature is not standardized and the variety of terms, subscript notation and formulas referring to the same concepts can be confusing. For example, Bielby & Hauser's notation (1977) use of gamma and lambda is different than that of Jöreskog and Sörbom.

The eight basic matrices are sufficiently general to subsume the specific model of the user. In each of the eight matrices, the parameters contained in the matrix may be set at predetermined values and not allowed to vary, and thus not be estimated, in a LISREL run. The parameters held at a constant value, often zero or 1, in a program run are referred to as "fixed" parameters. Other parameters, which are also elements of one of the eight matrices, will be allowed to vary in the program run and these are the parameters whose values will be estimated by the program. Such parameters are referred to as "free". "Constrained" parameters are those parameters which are equal to each other; as are two of the γ parameters in Figure 1. Within any of the eight matrices, the parameters may be of three different types--fixed, free, or constrained.

LISREL calculation routines differ from the usual routines used to calculate familiar statistics. For example, estimates of the mean, standard deviation, covariances, regression coefficients, and the range can be calculated using known formulas. In contrast, estimation routines in both LISREL and log linear approaches are not based on closed formulas. Rather, the user creates a starting point called a "starting value" and the estimation procedure goes through a sequence of successive approximations. All parameters must be assigned starting values. The program

iterates using the starting values until it finds the set of parameter values that best reproduces the original data matrix according to a criterion of fit.

The Fletcher-Powell approximation procedure as used in LISREL is based on the theory of normally distributed, continuous variables. If dichotomous or noncontinuous variables are used, the results may be difficult to interpret. It is most inadvisable to use noncontinuous dependent variables since that fixes the variance you are trying to explain. Henry (1973) has studied the use of discrete variables in measurement models and states ". . . parameter specifications that seem natural for continuous variables may be unnecessary or contradictory when analyzing dichotomous variables", (Henry, 1973, page 66).

The authors of this supplement worked with a dichotomous independent variable which yielded a negative measurement error variance. Fixing the parameter at zero and estimating again gained a degree of freedom and did not significantly affect other results. Dichotomies violate normality assumptions. Negative variance estimates can also be obtained with normal distributions, but setting negative estimates to zero complicates the distribution theory for the parameters, (Scheffé, 1959, pages 228-229).

The LISREL estimation is based on the maximum likelihood principle. Given: a) an assumption about the distribution form in a population, and b) a sample of observations from the population, then an estimate of a population parameter is chosen if the estimate maximizes the likelihood that the given sample results would be observed. The likelihood of a sample such as that obtained, given that a population parameter is a certain value, can be calculated. If the unknown parameter is estimated by a given statistic, and the probability of the sample is highest when the given statistic is used to estimate the unknown parameter, then the estimate giving such a maximum probability is called the "maximum-likelihood"

estimate of the unknown population parameter. For a simple example and definition when more than one parameter is estimated, see Mood and Graybill (1963, pages 178-182).

The Data Entered

The original data matrix should not contain linear dependencies. The sum or other linear functions of variables should not be analyzed along with the original measures.

A matrix of starting values for a covariance matrix should not contain linear dependencies. It would contain dependencies if all starting values were set at the same value such as .5 .5 .5 .5 .5 .5.

Either a correlation, covariance, the raw data or a moments matrix can be used to enter data. In the author's experience when convergence was difficult with a correlation matrix, it was significantly more difficult using the variance-covariance matrix. Furthermore, Jöreskog recommends that variables should not be standardized.

Is it better to start with a complex model first and simplify or start simple and add complexities? In regression there is evidence that forward selection procedures are more efficient than "backward deletions", (Kennedy and Bancroft, 1971). Jöreskog recommends that if you have large models, LISREL will be minimizing complicated functions and, therefore, running the model in pieces will help economize what you do.

The Identification Problem

The term "identification problem" is related to the necessity of using observed values to calculate the value of unknown quantities. In the LISREL program the known values are the nonredundant (diagonal and below-diagonal) values in the matrix to be analyzed. The number of

known values equals the number of values on the main diagonal, plus the number below the diagonal. For example, if you have q = 7 independent variables and p = 4 dependent variables, the 11 x 11 correlation matrix will have 11 values in the main diagonal and 55 below. This matrix is said to have 66 degrees of freedom. In general, the number of degrees of freedom in a matrix is equal to (1/2) (p + q) (p + q + 1).

One degree of freedom is used for each distinct parameter estimated. Let t be the number of unknown values to be identified. Then a necessary condition for the identification of all parameters is:

$$t < (1/2)\ (p + q)\ (p + q + 1)$$

As the LISREL User's Guide indicates, page 10, "It is sometimes difficult, in specific cases, to determine whether or not a parameter is identified and whether or not the whole model is identified." For example, linear dependencies among starting values will reduce the degrees of freedom and affect calculation routines which presupose matrices that contain no linear dependencies.

Measurement models have a scale indeterminacy not present in the structural model. Unless the scale of measurement is specified for a given unobserved factor, there will exist an infinity of solutions since many different multiplicative combinations of factor loadings and true scores can produce the same results. Fixing the scale of the unobserved factor removes the indeterminacy involved. This fixing can be done by setting the diagonal elements of Φ equal to 1 or by fixing a 1 in each column of lambda. Using both methods is too much.

Generally, identification of scale in the measurement models is done by fixing the measurement coefficient of one of the observed variables

associated with an unobserved factor as specified above. In a confirmatory factor analysis model where there are no independent variables, the scale of measurement of the dependent unobserved factors can be fixed by setting the diagonal elements of $\underset{\sim}{\Psi}$ equal to 1. This is illustrated in example 5 of the LISREL User's Guide, pp. 40ff.

Jöreskog has described (1978, pages 455ff) a concept called the "Best Fitting Simple Structure". It recommends that the same observed variable cannot be used to fix the scale of more than one unobserved factor. That is, in the $\underset{\sim}{\Lambda}_x$ or $\underset{\sim}{\Lambda}_y$ matrices, a row containing a 1 has zeros in all other row elements. Factors are orthogonal. Insignificant factor loadings are set at zero and all nonzero loadings are significant. An example is given on page 45 of the User's Guide.

The theory of identification of structural coefficients has been comprehensively described in Fisher (1966). Basic results are given in Johnston (1972, pages 356-360) and Land (1973, pages 27-32). Land shows that structural models without reciprocal causation are completely identified assuming the error disturbances are uncorrelated. In such a system, $\underset{\sim}{B}$ is lower triangular, all diagonal elements are equal to 1 and only $m\ (m-1)/2$ more parameters are introduced over the simple regression situation. Such a model is "just-identified": the number of values estimated equals the degrees of freedom in the variance-covariance matrix.

An equation-by-equation analysis of identification can be made using Goldberger's (1964, pages 316-317) rank condition. A sparsely explained example of its use is presented on page 16 of the User's Guide. Goldberger's order condition is: a necessary but not sufficient condition for the g:th equation to be identified is that $n_{**} > m_* - 1$, where * signifies

included in the equation and ** signifies excluded, n = the number of independent variables, and m = the number of dependent variables. In words, the number of independent variables excluded from the equation must be at least as great as the number of dependent variables included in the equation less one.

Tintner's model of the U.S. meat market can be described as:

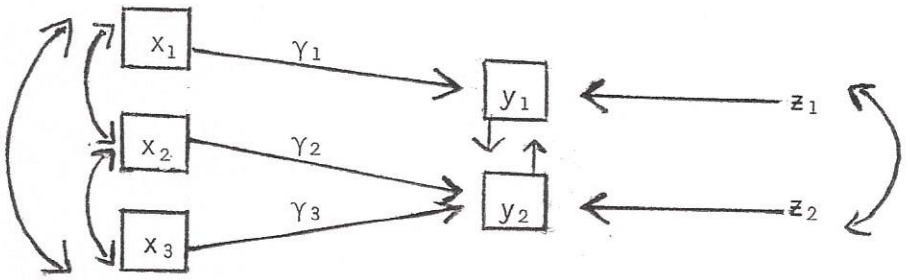

Figure 4 - Tintner's Model of the U.S. Meat Market - A Purely Structural Model.

The first equation is $y_1 = \beta_1 y_2 + \gamma_1 x_1 + \zeta_1$. Since two x variables are excluded $N_{**}=2$ and since 2 y variables are included $m_*=2$. Thus, $N_{**}-(m_*-1) = 1$ and the y_1 equation is said to be overidentified by 1 degree of freedom. The y_2 equation can be similarly analyzed. So it is possible to make an equation-by-equation analysis of a model to see if its parameters are identifiable.

The program will print out an error message, THE Nth FREE PARAMETER. MAY NOT BE IDENTIFIED, which specifies the last parameter in an indeterminate sequence. Usually this means that the particular parameter is not identified but it may mean more.

For further comments on identification, the user is referred to the LISREL User's Guide and the references cited above.

Specifying Models to LISREL

Chapter II of the User's Guide contains detailed instructions on how to set up the problem run. The following steps are used.

First, draw the model. Secondly, relate the eight matrices to the model. Which matrices are relevant and what are their characteristics. The user must tell the program, matrix-by-matrix, which parameters are to be estimated, which are equal to each other and which are fixed. Then, starting values are assigned to all values in all matrices used. Then the program is run.

The format cards for LISREL use two key letters to designate each matrix. These letters are provided in Figure 3 and are: $\underset{\sim}{\Lambda}_y$ = LY, $\underset{\sim}{B}$ = BE, $\underset{\sim}{\Phi}$ = PH, $\underset{\sim}{\Lambda}_X$ = LX, $\underset{\sim}{\Gamma}$ = GA, $\underset{\sim}{\Psi}$ = PS, $\underset{\sim}{\theta}_\varepsilon$ = TE and θ_δ = TD. These matrices are very versatile and can be used to express a variety of models. For example, the LISREL model for ordinary least squares is:

LY = IDENTITY, LX = IDENTITY BE = IDENTITY GA = FULL, FREE

PH = SYMMETRIC FIXED, PS = FREE, DIAGONAL TE = ZERO TD = ZERO

Since there are no unobserved factors LX and LY are set at the identity matrix. The measurement error variances are set at zero and there is no causal structure among the dependent variables since BE = IDENTITY. The only parameters that are free to vary are in GA and the residual variances in PS. The PH matrix is specified as symmetric and fixed. If LISREL were applied to these matrices the only estimates obtained would be regression coefficients and residual variances.

The above model specifications provide regression results but require assigning starting values for PH. A simpler specification exists.

LY = IDENTITY BE = IDENTITY GA = FULL, FREE

FIXED X PS = DIAGONAL, FREE TE = ZERO

The program then will automatically define values for LX, TD and PH. In the FIXED X case the program will use the variance-covariance matrix of the independent observed variables as starting values for PH.

This demonstrates how a single parameter specification can convey information about the matrix form, set starting values, or define the status of all parameters in a matrix. For example, the assumption of independent errors can be conveyed by specifying PS = DIAGONAL, FREE, TE = DIAGONAL, FREE or TD = DIAGONAL, FREE. These specifications tell the program to estimate error values in the matrices' diagonals.

The LISREL format procedures are flexible. Matrix specifications can be variously done. For example, if a user wishes to free only one parameter in a matrix, the entire matrix can be fixed and then another specification e.g., FREE BE (3,4), as illustrated on pages 72-73 of the User's Guide frees the desired parameter.

The constant 0.5 seems to give good results as a starting value for the structural coefficients in BE and GA, and for the estimated error variances. In the author's data, where a covariance matrix PH, PS, TE or TD, is entered, and it is not diagonal, 0.5 was used for diagonal variances and 0.25 for the off-diagonal covariances. For example, the starting values for a 3 x 3 PH matrix were entered as:

```
MA  PH
(6F 3.2)
.5 .25 .5 .25 .25 .5
```

A chi-square maximum-likelihood test of the goodness of fit of the model to the data is generated as part of the program's standard output. Use of this goodness-of-fit statistic in comparing successive models is an ad hoc precedure which should be used with judgment. If the user's model were correct and if all assumptions such as multi-normality

were present, then the χ^2 fit is in fact χ^2 distributed. However, assumptions are frequently not met and it makes more sense to view any given model as an approximation.

It is the differences between the χ^2 for various models that provide indications as to which model better reproduces the original data matrix. The original data matrix is usually a correlation or variance-covariance matrix. Based on the LISREL estimates of the user's model an estimated variance-covariance matrix is constructed and compared to the original matrix. The chi-square statistic states how close the approximation is between the original and the estimated matrix. If the chi-square is large compared to the degrees of freedom in the model the model is said to fit poorly. If the chi-square value is close to the degrees of freedom the model is said to fit well. The comparison is a matter of judgment and users should be guided by their knowledge of the theoretical relations among the variables, (see Long, 1976, pp. 168-173).

It is possible to use an efficiency ratio defined as the χ^2 value for the model divided by the degrees of freedom in the model. This index seems appropriate given Marriott's (1974) rules of thumb for best fitting models. In a LISREL context, these characteristics could be restated as: a) all structural and measurement coefficients are significant, and b) no other combination of parameters gives a better fit.

A "probability level" is also printed out by the program. This is defined as the probability of getting a χ^2 value larger than the value actually attained, given that the hypothesized model is true. If you get a value of .10-.15 , be happy with the fit. When χ^2 = degrees of freedom, the probability level $\tilde{=}$ .50.

How Can The Fit Of A Model Be Improved

If your model doesn't fit, you need a theory to tell you which parameter to change or you can change the model on an empirical basis. You can look at the first derivatives. A matrix of derivatives is available as optional output and you should request that the first order derivatives be printed. The derivatives are partial derivatives of the function being minimized with respect to the particular parameter. Relax the largest first derivative, or two or three of the largest derivatives. "Relax" means to rerun the program again, but free the parameter(s) associated with each derivative and let the program estimate it. You can do this if you have enough degrees of freedom. This is an ad hoc procedure.

There is no set rule as to what constitutes a large derivative. The derivatives theoretically have standard errors. However, presently, they are not practically calculable. Present methods require the calculation of the inverse of a matrix whose order is equal to the number of free parameters in the model. Relaxing the derivative may not help, but it is currently the best procedure available.

References

The remaining section of this paper contains: a) references to authors cited in the paper and b) a bibliography of references pertinent to LISREL and structural equation theory. The bibliography was obtained from Jöreskog.

REFERENCES

BIELBY, W. T. and R. M. HAUSER. (1977) "Structural Equation Models." Annual Review of Sociology 3: 137-61.

BLALOCK, H. M. (1964) Casual Inferences in Nonexperimental Research, University of North Carolina Press, Chapel Hill.

FISHER, F. M. (1966) The Identification Problem in Econometrics, New York: McGraw-Hill.

GOLDBERGER, A. S. (1964) Econometric Theory, New York: Wiley.

HENRY, N. W. (1973) "Measurement models for continuous and discrete variables." In A. S. Goldberger and O. D. Duncan (Eds.), Structural Equation Models in the Social Sciences. New York: Seminar Press. Pp. 51-67.

JOHNSTON, J. (1972) Econometric Methods, New York: McGraw-Hill. Pp.356-360.

JÖRESKOG, K. G. (1978) "Structural analysis of covariance and correlation matrices," Psychometrika, 43: 443-477.

KENNEDY, W. J. & BANCROFT, T. A. (1972) "Model building for prediction in regression based upon repeated significance tests,"Annals of mathematical statistics, 42: 1273-1284.

KIM, J. and KOHOUT, F. J. (1975) "Multiple regression analysis: Subprogram regression." In H. N. Nie, et al, SPSS: statistical package for the social sciences. New York: McGraw-Hill. Pp.320-367.

LAND, K. C. (1973) "Identification, parameter estimation, and hypothesis testing in recursive sociological models." In A. S. Goldberger and O. D. Duncan (Eds.), Structural equation models in the social sciences. New York: Seminar Press. Pp. 19-49.

LONG, J. S. (1976) "Estimation and Hypothesis Testing in Linear Models Measurement Error." Sociological Methods and Research, 5: 157-207.

MARRIOTT, F. H. C. (1974) The interpretation of multiple observations. San Francisco: Academic Press.

MOOD, A. M. and GRAYBILL, F. A. (1963) Introduction to the theory of statistics. (Second ed.) New York: McGraw-Hill.

SCHEFFÉ, H. (1959) The analysis of variance. New York: John Wiley & Sons, Inc.

REFERENCES

Books

1. Harman, H.H. Modern factor analysis. Chicago: The University of Chicago Press, 1967.

2. Lawley, D.N., Maxwell, A.E. Factor analysis as a statistical method. (2nd ed.). New York: American Elsevier, 1971.

3. Mulaik, S.A. The foundation of factor analysis. New York: McGraw-Hill, 1972.

4. Gursuch, R.L. Factor analysis. Philadelphia: W.B. Saunders Company, 1974.

5. Lord, F.M. & Novick, M.E. Statistical theories of mental test scores. Reading: Addison-Wesley Publishing Co., 1968.

6. Blalock, H.M. Jr. Causal models in the social sciences. Chicago, Illinois: Aldine-Atherton, 1971.

7. Blalock, H.M. Jr. Measurement in the social sciences. Chicago: Aldine Publishing Company, 1974.

8. Heise, D.R. Causal analysis. New York: Wiley, 1975.

9. Duncan, O.D. Introduction to structural equation models. New York: Academic Press, 1975.

10. Goldberger, A.S. & Duncan, O.D. Structural equation models in the social sciences. New York: Seminar Press, 1973.

11. Aigner, D. J. & Goldberger, A.S. Latent variables in socio-economic models. Amsterdam: North-Holland Publishing Co., 1977.

12. Jöreskog, K.G. Statistical estimation in factor analysis. Stockholm: Almqvist & Wiksell, 1963.

13. Jöreskog, K.G. Some contributions to maximum likelihood factor analysis. Psychometrika, 1967, 32, 443-482.

14. Jöreskog, K.G. & Lawley, D.N. New methods in maximum likelihood factor analysis. British Journal of Mathematical and Statistical Psychology, 1968, 21, 85-96.

15. Jöreskog, K.G. & Goldberger, A.S. Factor analysis by generalized least squares. Psychometrika, 1972, 37, 243-260.

16. Jöreskog, K.G. Factor analysis by least squares and maximum likelihood methods. K. Enslein, A. Ralston & M.S. Wilf (Eds.), Statistical methods for digital computers. New York: Wiley, 1976, 125-165.

17. Kaiser, H.F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

18. Hendrickson, A.E. & White, P.O. Promax: A quick method for rotation to oblique simple structure. British Journal of Statistical Psychology, 1964, 17, 65-70.

19. Jennrich, R.I. & Sampson, P. Rotation for simple loadings. Psychometrika, 1966, 31, 313-323.

20. Lawley, D.N. & Maxwell, A.E. Factor transformation methods. British Journal of Statistical Psychology, 1964, 17, 97-103.

21. Browne, M.W. On oblique Procrustes rotation. Psychometrika, 1967, 32, 125-154.

22. Browne, M.W. Oblique rotation to a partially specified target. British Journal of Mathematical and Statistical Psychology, 1972, 25, 207-212.

23. Archer, C.O. & Jennrich, R.I. Standard errors for rotated factor loadings. Psychometrika, 1973, 38, 581-592.

24. Jennrich, R.I. Standard errors for obliquely rotated factor loadings. Psychometrika, 1973, 38, 593-604.

Confirmatory Factor Analysis and Analysis of Covariance Structures

25. Jöreskog, K.G. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 1969, 34, 183-202.

26. Anderson, T.W. Statistical inference for covariance matrices with linear structure. In P.R. Krishnaiah (Ed.), Multivariate analysis - II. New York, Academic Press, 1969, 55 - 66.

27. Anderson, T.W. Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices. In R.C. Bose et al. (Eds.), Essays in probability and statistics. Chapel Hill: University of North Carolina Press, 1970, 1 - 24.

28. Anderson, T.W. Asymptotically efficient estimation of covariance matrices with linear structure. Annals of Statistics, 1973, 1, 135-141.

29. Bock, R.D. & Bargman, R.E. Analysis of covariance structures. Psychometrika, 1966, 31, 507-534.

30. Mukherjee, B.N. Likelihood ratio tests of statistical hypotheses associated with patterned covariance matrices in psychology. British Journal of Mathematical and Statistical Psychology, 1970, 23, 120.

31. Jöreskog, K.G. Statistical analysis of sets of congeneric tests. Psychometrika, 1971, 36, 109-133.

32. Jöreskog, K.G. A general method for analysis of covariance structures. Biometrika, 1970, 57, 239-251.

33. Jöreskog, K.G. Analyzing psychological data by structural analysis of covariance matrices. In R.C. Atkinson, D.H. Krantz, R.D. Luce & P. Suppes (Eds.), Contemporary developments in mathematical psychology - Volume II. San Francisco: W.H. Freeman & Co., 1974, 1-56.

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

35. Browne, M.W. Generalized least-squares estimators in the analysis of covariance structures. In D.J. Aigner and A.S. Goldberger (Eds.), Latent variables in socio-economic models. North Holland Publishing Co., 1970, 203-226.

36. Browne, M.W. The analysis of patterned correlation matrices by generalized least squares. British Journal of Mathematical and Statistical Psychology, 1977, 30, 113-124.

Simultaneous Factor Analysis in Several Populations

37. Meredith, W. Notes on factorial invariance. Psychomerika, 1964, 29, 177-185.

38. Meredith, W. Rotation to achieve factorial invariance. Psychometrika, 1964, 29, 187-206.

39. Jöreskog, K.G. Simultaneous factor analysis in several populations. Psychometrika, 1971, 36, 409-426.

40. Sörbom, D. A general method for studying differences in factor means and factor structure between groups. British Journal of Mathematical and Statistical Psychology, 1974, 27, 229-239.

Structural Equation Models

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

42. Goldberger, A.S. Structural equation methods in the social sciences. Econometrica, 1972, 40, 979-1001.

43. Geraci, V.J. Identification of simultaneous equation models with measurement error. Journal of Econometrics, 1976, 4, 263-283.

44. Griliches, Z. Errors in variables and other unobservables. Econometrica, 1974, 42, 971-998.

45. Duncan, O.D. Path analysis: sociological examples. American Journal of Sociology, 1966, 72, 1-16.

46. Werts, C.E. & Linn, R.L. Path analysis: Psychological examples. Psychological Bulletin, 1970, 74, 193-212.

47. Duncan, O.D., Haller, A.O. & Portes, A. Peer influences on aspirations, a reinterpretation. In H.M. Blalock, Jr. (Ed), Causal models in the social sciences. Aldine-Atherton, Inc., 1971, 219-2441

48. Hauser, R.M. & Goldberger, A.S. The treatment of unobservable variables in path analysis. In H.L. Costner (Ed.) Sociological methodology 1971, San Francisco: Jossey-Bass, 1971, 81-117.

49. Jöreskog, K.G. A general method for estimating a linear structural equation system. In A.S. Goldberger & O.D. Duncan (Eds.), Structural equation models in the social sciences. New York: Seminar Press, 1973, 85-112.

50. Jöreskog, K.G. Structural equation models in the social sciences: Specification, estimation and testing. In P.R. Krishnaiah (Ed), Proceedings of the Symposium on applications of statistics. Amsterdam: North Holland Publishing Co., 1977, 265-286.

51. Wiley, D.E. The identification problem for structural equation models with unmeasured variables. In A.S. Goldberger & O.D. Duncan (Eds.), Structural equation models in the social sciences. New York: Seminar Press, 1974, 69-83.

52. Jöreskog, K.G. & Goldberger, A.S. Estimation of a model with multiple indicators and multiple causes of a single latent variable. Journal of the American Statistical Association, 1975, 10, 631-639.

53. Jöreskog, K.G. Estimation and testing of simplex models. British Journal of Mathematical and Statistical Psychology, 1970, 23, 121-145.

54. Sörbom, D. Detection of correlated errors in longitudinal data. British Journal of Mathematical and Statistical Psychology, 1975, 28, 138-151.

55. Jöreskog, K.G. & Sörbom, D. Statistical models and methods for test-retest situations. In D.N.M. de Gruijter & L.J. Th. van der Kamp (Eds.), Advances in psychological and educational measurement. New York: Wiley, 1976, 135-157.

56. Jöreskog, K.G. & Sörbom, D. Statistical models and methods for analysis of longitudinal data. In D.J. Aigner & A.S. Goldberger (Eds.), Latent variables in socio-economic models. Amsterdam: North Holland Publishing Co., 1977, 285-235.

57. Jöreskog, K.G. An econometric model for multivariate panel data. Paper presented at the INSEE-CNRS Symposium on the Econometrics of Panel Data in Paris, August 22-24, 1977.