

Chapter 14

Introduction to Structural Equations with Latent Variables

Chapter Table of Contents

| | |
|---|-----|
| OVERVIEW | 189 |
| COMPARISON OF THE CALIS AND SYSLIN PROCEDURES | 189 |
| MODEL SPECIFICATION | 190 |
| ESTIMATION METHODS | 191 |
| STATISTICAL INFERENCE | 191 |
| GOODNESS-OF-FIT STATISTICS | 193 |
| OPTIMIZATION METHODS | 193 |
| SPECIFYING STRUCTURAL EQUATION MODELS | 194 |
| Identification of Models | 199 |
| PATH DIAGRAMS AND THE RAM MODEL | 203 |
| SOME MEASUREMENT MODELS | 209 |
| A COMBINED MEASUREMENT-STRUCTURAL MODEL WITH RE- CIPROCAL INFLUENCE AND CORRELATED RESIDUALS | 222 |
| REFERENCES | 250 |

Chapter 14

Introduction to Structural Equations with Latent Variables

Overview

You can use the CALIS procedure for analysis of covariance structures, fitting systems of linear structural equations, and path analysis. These terms are more or less interchangeable, but they emphasize different aspects of the analysis. The analysis of covariance structures refers to the formulation of a model for the variances and covariances among a set of variables and the fitting of the model to an observed covariance matrix. In linear structural equations, the model is formulated as a system of equations relating several random variables with assumptions about the variances and covariances of the random variables. In path analysis, the model is formulated as a path diagram, in which arrows connecting variables represent (co)variances and regression coefficients. Path models and linear structural equation models can be converted to models of the covariance matrix and can, therefore, be fitted by the methods of covariance structure analysis. All of these methods allow the use of hypothetical latent variables or measurement errors in the models.

Loehlin (1987) provides an excellent introduction to latent variable models using path diagrams and structural equations. A more advanced treatment of structural equation models with latent variables is given by Bollen (1989). Fuller (1987) provides a highly technical statistical treatment of measurement-error models.

Comparison of the CALIS and SYSLIN Procedures

The SYSLIN procedure in the SAS/ETS product can also fit certain kinds of path models and linear structural equation models. PROC CALIS differs from PROC SYSLIN in that PROC CALIS allows more generality in the use of latent variables in the models. Latent variables are unobserved, hypothetical variables, as distinct from manifest variables, which are the observed data. PROC SYSLIN allows at most one latent variable, the error term, in each equation. PROC CALIS allows several latent variables to appear in an equation—in fact, all the variables in an equation can be latent as long as there are other equations that relate the latent variables to manifest variables.

Both the CALIS and SYSLIN procedures enable you to specify a model as a system of linear equations. When there are several equations, a given variable may be a dependent variable in one equation and an independent variable in other equations. Therefore, additional terminology is needed to describe unambiguously the roles of variables in the system. Variables with values that are determined jointly and simultaneously by the system of equations are called *endogenous variables*. Variables

with values that are determined outside the system, that is, in a manner separate from the process described by the system of equations, are called exogenous variables. The purpose of the system of equations is to explain the variation of each endogenous variable in terms of exogenous variables or other endogenous variables or both. Refer to Loehlin (1987, p. 4) for further discussion of endogenous and exogenous variables. In the econometric literature, error and disturbance terms are usually distinguished from exogenous variables, but in systems with more than one latent variable in an equation, the distinction is not always clear.

In PROC SYSLIN, endogenous variables are identified by the ENDOGENOUS statement. When you specify structural equations in PROC CALIS, endogenous variables are assumed to be those that appear on the left-hand sides of the equations; a given variable may appear on the left-hand side of at most one equation.

PROC SYSLIN provides many methods of estimation, some of which are applicable only in special cases. For example, ordinary least-squares estimates are suitable in certain kinds of systems but may be statistically biased and inconsistent in other kinds. PROC CALIS provides three methods of estimation that can be used with most models. Both the CALIS and SYSLIN procedures can do maximum likelihood estimation, which PROC CALIS calls ML and PROC SYSLIN calls FIML. PROC SYSLIN can be much faster than PROC CALIS in those special cases for which it provides computationally efficient estimation methods. However, PROC CALIS has a variety of sophisticated algorithms for maximum likelihood estimation that may be much faster than FIML in PROC SYSLIN.

PROC CALIS can impose a wider variety of constraints on the parameters, including nonlinear constraints, than can PROC SYSLIN. For example, PROC CALIS can constrain error variances or covariances to equal specified constants, or it can constrain two error variances to have a specified ratio.

Model Specification

PROC CALIS provides several ways to specify a model. Structural equations can be transcribed directly in the LINEQS statement. A path diagram can be described in the RAM statement. You can specify a first-order factor model in the FACTOR and MATRIX statements. Higher-order factor models and other complicated models can be expressed in the COSAN and MATRIX statements. For most applications, the LINEQS and RAM statements are easiest to use; the choice between these two statements is a matter of personal preference.

You can save a model specification in an OUTRAM= data set, which can then be used with the INRAM= option to specify the model in a subsequent analysis.

Estimation Methods

The CALIS procedure provides three methods of estimation specified by the METHOD= option:

| | |
|-----|--|
| ULS | unweighted least squares |
| GLS | generalized least squares |
| ML | maximum likelihood for multivariate normal distributions |

Each estimation method is based on finding parameter estimates that minimize a badness-of-fit function that measures the difference between the observed sample covariance matrix and the predicted covariance matrix given the model and the parameter estimates. See the section “Estimation Methods” on page 462 in Chapter 19, “The CALIS Procedure,” for formulas, or refer to Loehlin (1987, pp. 54–62) and Bollen (1989, pp. 104–123) for further discussion.

The default is METHOD=ML, which is the preferred method for most applications with respect to statistical considerations. The option METHOD=GLS usually produces very similar results to METHOD=ML. Both methods are suitable regardless of the scaling of the covariance matrix. The ULS method is appropriate for a covariance matrix only if the variables are measured on comparable scales; otherwise, METHOD=ULS should be applied to the correlation matrix. PROC CALIS cannot compute standard errors or test statistics with the ULS method.

You should not specify METHOD=ML or METHOD=GLS if the observed or predicted covariance matrix is singular—you should either remove variables involved in the linear dependencies or specify METHOD=ULS.

PROC CALIS should not be used if your data are extremely nonnormal data, especially if they have high kurtosis. You should remove outliers and try to transform variables that are skewed or heavy-tailed. This applies to all three estimation methods, since all the estimation methods depend on the sample covariance matrix, and the sample covariance matrix is a poor estimator for distributions with high kurtosis (Bollen 1989, pp. 415–418; Huber 1981; Hampel et. al 1986). PROC CALIS displays estimates of univariate and multivariate kurtosis (Bollen 1989, pp. 418–425) if you specify the KURTOSIS option in the PROC CALIS statement.

Statistical Inference

When you specify the ML or GLS estimates with appropriate models, PROC CALIS can compute

- a chi-square goodness-of-fit test of the specified model versus the alternative that the data are from a multivariate normal distribution with unconstrained covariance matrix (Loehlin 1987, pp. 62–64; Bollen 1989, pp. 110, 115, 263–269)

- approximate standard errors of the parameter estimates (Bollen 1989, pp. 109, 114, 286), displayed with the STDERR option
- various modification indices, requested via the MODIFICATION or MOD option, that give the approximate change in the chi-square statistic that would result from removing constraints on the parameters or constraining additional parameters to zero (Bollen 1989, pp. 293–303)

If you have two models such that one model results from imposing constraints on the parameters of the other, you can test the constrained model against the more general model by fitting both models with PROC CALIS. If the constrained model is correct, the difference between the chi-square goodness-of-fit statistics for the two models has an approximate chi-square distribution with degrees of freedom equal to the difference between the degrees of freedom for the two models (Loehlin 1987, pp. 62–67; Bollen 1989, pp. 291–292).

All of the test statistics and standard errors computed by PROC CALIS depend on the assumption of multivariate normality. Normality is a much more important requirement for data with random independent variables than it is for fixed independent variables. If the independent variables are random, distributions with high kurtosis tend to give liberal tests and excessively small standard errors, while low kurtosis tends to produce the opposite effects (Bollen 1989, pp. 266–267, 415–432).

The test statistics and standard errors computed by PROC CALIS are also based on asymptotic theory and should not be trusted in small samples. There are no firm guidelines on how large a sample must be for the asymptotic theory to apply with reasonable accuracy. Some simulation studies have indicated that problems are likely to occur with sample sizes less than 100 (Loehlin 1987, pp. 60–61; Bollen 1989, pp. 267–268). Extrapolating from experience with multiple regression would suggest that the sample size should be at least five to twenty times the number of parameters to be estimated in order to get reliable and interpretable results.

The asymptotic theory requires that the parameter estimates be in the interior of the parameter space. If you do an analysis with inequality constraints and one or more constraints are active at the solution (for example, if you constrain a variance to be nonnegative and the estimate turns out to be zero), the chi-square test and standard errors may not provide good approximations to the actual sampling distributions.

Standard errors may be inaccurate if you analyze a correlation matrix rather than a covariance matrix even for sample sizes as large as 400 (Boomsma 1983). The chi-square statistic is generally the same regardless of which matrix is analyzed provided that the model involves no scale-dependent constraints.

If you fit a model to a correlation matrix and the model constrains one or more elements of the predicted matrix to equal 1.0, the degrees of freedom of the chi-square statistic must be reduced by the number of such constraints. PROC CALIS attempts to determine which diagonal elements of the predicted correlation matrix are constrained to a constant, but it may fail to detect such constraints in complicated models, particularly when programming statements are used. If this happens, you should add parameters to the model to release the constraints on the diagonal elements.

Goodness-of-fit Statistics

In addition to the chi-square test, there are many other statistics for assessing the goodness of fit of the predicted correlation or covariance matrix to the observed matrix.

Akaike's (1987) information criterion (AIC) and Schwarz's (1978) Bayesian criterion (SBC) are useful for comparing models with different numbers of parameters—the model with the smallest value of AIC or SBC is considered best. Based on both theoretical considerations and various simulation studies, SBC seems to work better, since AIC tends to select models with too many parameters when the sample size is large.

There are many descriptive measures of goodness of fit that are scaled to range approximately from zero to one: the goodness of fit index (GFI) and GFI adjusted for degrees of freedom (AGFI) (Jöreskog and Sörbom 1988), centrality (McDonald 1989), and the parsimonious fit index (James, Mulaik, and Brett 1982). Bentler and Bonett (1980) and Bollen (1986) have proposed measures for comparing the goodness of fit of one model with another in a descriptive rather than inferential sense.

None of these measures of goodness of fit are related to the goodness of prediction of the structural equations. Goodness of fit is assessed by comparing the observed correlation or covariance matrix with the matrix computed from the model and parameter estimates. Goodness of prediction is assessed by comparing the actual values of the endogenous variables with their predicted values, usually in terms of root mean squared error or proportion of variance accounted for (R^2). For latent endogenous variables, root mean squared error and R^2 can be estimated from the fitted model.

Optimization Methods

PROC CALIS uses a variety of nonlinear optimization algorithms for computing parameter estimates. These algorithms are very complicated and do not always work. PROC CALIS will generally inform you when the computations fail, usually by displaying an error message about the iteration limit being exceeded. When this happens, you may be able to correct the problem simply by increasing the iteration limit (MAXITER= and MAXFUNC=). However, it is often more effective to change the optimization method (OMETHOD=) or initial values. For more details, see the section “Use of Optimization Techniques” on page 551 in Chapter 19, “The CALIS Procedure,” and refer to Bollen (1989, pp. 254–256).

PROC CALIS may sometimes converge to a local optimum rather than the global optimum. To gain some protection against local optima, you can run the analysis several times with different initial estimates. The RANDOM= option in the PROC CALIS statement is useful for generating a variety of initial estimates.

Specifying Structural Equation Models

Consider fitting a linear equation to two observed variables, Y and X . Simple linear regression uses the model of a particular form, labeled for purposes of discussion, as Model Form A.

Model Form A

$$Y = \alpha + \beta X + E_Y$$

where α and β are coefficients to be estimated and E_Y is an error term. If the values of X are fixed, the values of E_Y are assumed to be independent and identically distributed realizations of a normally distributed random variable with mean zero and variance $\text{Var}(E_Y)$. If X is a random variable, X and E_Y are assumed to have a bivariate normal distribution with zero correlation and variances $\text{Var}(X)$ and $\text{Var}(E_Y)$, respectively. Under either set of assumptions, the usual formulas hold for the estimates of the coefficients and their standard errors (see Chapter 3, “Introduction to Regression Procedures”).

In the REG or SYSLIN procedure, you would fit a simple linear regression model with a MODEL statement listing only the names of the manifest variables:

```
proc reg;
  model y=x;
run;
```

You can also fit this model with PROC CALIS, but you must explicitly specify the names of the parameters and the error terms (except for the intercept, which is assumed to be present in each equation). The linear equation is given in the LINEQS statement, and the error variance is specified in the STD statement.

```
proc calis cov;
  lineqs y=beta x + ex;
  std ex=vex;
run;
```

The parameters are the regression coefficient BETA and the variance VEX of the error term EX. You do not need to type an * between BETA and X to indicate the multiplication of the variable by the coefficient.

The LINEQS statement uses the convention that the names of error terms begin with the letter E, disturbances (error terms for latent variables) in equations begin with D, and other latent variables begin with F for “factor.” Names of variables in the input SAS data set can, of course, begin with any letter.

If you leave out the name of a coefficient, the value of the coefficient is assumed to be 1. If you leave out the name of a variance, the variance is assumed to be 0. So if you tried to write the model the same way you would in PROC REG, for example,

```
proc calis cov;
  lineqs y=x;
```

you would be fitting a model that says Y is equal to X plus an intercept, with no error.

The COV option is used because PROC CALIS, like PROC FACTOR, analyzes the correlation matrix by default, yielding standardized regression coefficients. The COV option causes the covariance matrix to be analyzed, producing raw regression coefficients. See Chapter 3, “Introduction to Regression Procedures,” for a discussion of the interpretation of raw and standardized regression coefficients.

Since the analysis of covariance structures is based on modeling the covariance matrix and the covariance matrix contains no information about means, PROC CALIS neglects the intercept parameter by default. To estimate the intercept, change the COV option to UCOV, which analyzes the uncorrected covariance matrix, and use the AUGMENT option, which adds a row and column for the intercept, called INTERCEP, to the matrix being analyzed. The model can then be specified as

```
proc calis ucov augment;
  lineqs y=alpha intercep + beta x + ex;
  std ex=vex;
run;
```

In the LINEQS statement, `intercep` represents a variable with a constant value of 1; hence, the coefficient `alpha` is the intercept parameter.

Other commonly used options in the PROC CALIS statement include

- MODIFICATION to display model modification indices
- RESIDUAL to display residual correlations or covariances
- STDERR to display approximate standard errors
- TOTEFF to display total effects

For ordinary unconstrained regression models, there is no reason to use PROC CALIS instead of PROC REG. But suppose that the observed variables Y and X are contaminated by error, and you want to estimate the linear relationship between their true, error-free scores. The model can be written in several forms. A model of Form B is as follows.

Model Form B

$$\begin{aligned}
 Y &= \alpha + \beta F_X + E_Y \\
 X &= F_X + E_X \\
 \text{Cov}(F_X, E_X) &= \text{Cov}(F_X, E_Y) = \text{Cov}(E_X, E_Y) = 0
 \end{aligned}$$

This model has two error terms, E_Y and E_X , as well as another latent variable F_X representing the true value corresponding to the manifest variable X . The true value corresponding to Y does not appear explicitly in this form of the model.

The assumption in Model Form B is that the error terms and the latent variable F_X are jointly uncorrelated is of critical importance. This assumption must be justified on substantive grounds such as the physical properties of the measurement process. If this assumption is violated, the estimators may be severely biased and inconsistent.

You can express Model Form B in PROC CALIS as follows:

```

proc calis cov;
  lineqs y=beta fx + ey,
         x=fx + ex;
  std fx=vfx,
      ey=vey,
      ex=vex;
run;

```

You must specify a variance for each of the latent variables in this model using the STD statement. You can specify either a name, in which case the variance is considered a parameter to be estimated, or a number, in which case the variance is constrained to equal that numeric value. In general, you must specify a variance for each latent exogenous variable in the model, including error and disturbance terms. The variance of a manifest exogenous variable is set equal to its sample variance by default. The variances of endogenous variables are predicted from the model and are not parameters. Covariances involving latent exogenous variables are assumed to be zero by default. Covariances between manifest exogenous variables are set equal to the sample covariances by default.

Fuller (1987, pp. 18–19) analyzes a data set from Voss (1969) involving corn yields (Y) and available soil nitrogen (X) for which there is a prior estimate of the measurement error for soil nitrogen $\text{Var}(E_X)$ of 57. You can fit Model Form B with this constraint using the following SAS statements.

```

data corn(type=cov);
  input _type_ $ _name_ $ y x;
  datalines;
n      . 11      11
mean   . 97.4545 70.6364
cov    y 87.6727 .
cov    x 104.8818 304.8545
;

proc calis data=corn cov stderr;
  lineqs y=beta fx + ey,
         x=fx + ex;
  std   ex=57,
        fx=vfx,
        ey=vey;
run;

```

In the STD statement, the variance of EX is given as the constant value 57. PROC CALIS produces the following estimates.

| The SAS System | | | | |
|--|-----------|-------------|----------------|-----------|
| The CALIS Procedure | | | | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | |
| y | = | 0.4232*fx | + | 1.0000 ey |
| Std Err | | 0.1658 beta | | |
| t Value | | 2.5520 | | |
| x | = | 1.0000 fx | + | 1.0000 ex |
| Variances of Exogenous Variables | | | | |
| Variable | Parameter | Estimate | Standard Error | t Value |
| fx | vfx | 247.85450 | 136.33508 | 1.82 |
| ey | vey | 43.29105 | 23.92488 | 1.81 |
| ex | | 57.00000 | | |

Figure 14.1. Measurement Error Model for Corn Data

PROC CALIS also displays information about the initial estimates that can be useful if there are optimization problems. If there are no optimization problems, the initial estimates are usually not of interest; they are not reproduced in the examples in this chapter.

You can write an equivalent model (labeled here as Model Form C) using a latent variable F_Y to represent the true value corresponding to Y .

Model Form C

$$\begin{aligned}
 Y &= F_Y + E_Y \\
 X &= F_X + E_X \\
 F_Y &= \alpha + \beta F_X \\
 \text{Cov}(F_X, E_X) &= \text{Cov}(F_X, E_X) = \text{Cov}(E_X, E_Y) = 0
 \end{aligned}$$

The first two of the three equations express the observed variables in terms of a true score plus error; these equations are called the measurement model. The third equation, expressing the relationship between the latent true-score variables, is called the structural or causal model. The decomposition of a model into a measurement model and a structural model (Keesling 1972; Wiley 1973; Jöreskog 1973) has been popularized by the program LISREL (Jöreskog and Sörbom 1988). The statements for fitting this model are

```

proc calis cov;
  lineqs y=fy + ey,
         x=fx + ex,
         fy=beta fx;
  std fx=vfx,
      ey=vey,
      ex=vex;
run;

```

You do not need to include the variance of F_Y in the STD statement because the variance of F_Y is determined by the structural model in terms of the variance of F_X , that is, $\text{Var}(F_Y) = \beta^2 \text{Var}(F_X)$.

Correlations involving endogenous variables are derived from the model. For example, the structural equation in Model Form C implies that F_Y and F_X are correlated unless β is zero. In all of the models discussed so far, the latent exogenous variables are assumed to be jointly uncorrelated. For example, in Model Form C, E_Y , E_X , and F_X are assumed to be uncorrelated. If you want to specify a model in which E_Y and E_X , say, are correlated, you can use the COV statement to specify the numeric value of the covariance $\text{Cov}(E_Y, E_X)$ between E_Y and E_X , or you can specify a name to make the covariance a parameter to be estimated. For example,

```

proc calis cov;
  lineqs y=fy + ey,
         x=fx + ex,
         fy=beta fx;
  std fy=vfy,
      fx=vfx,
      ey=vey,
      ex=vex;
  cov ey ex=ceyex;
run;

```

This COV statement specifies that the covariance between EY and EX is a parameter named CEYEX. All covariances that are not listed in the COV statement and that are not determined by the model are assumed to be zero. If the model contained two or more manifest exogenous variables, their covariances would be set to the observed sample values by default.

Identification of Models

Unfortunately, if you try to fit models of Form B or Form C without additional constraints, you cannot obtain unique estimates of the parameters. These models have four parameters (one coefficient and three variances). The covariance matrix of the observed variables Y and X has only three elements that are free to vary, since $\text{Cov}(Y,X)=\text{Cov}(X,Y)$. The covariance structure can, therefore, be expressed as three equations in four unknown parameters. Since there are fewer equations than unknowns, there are many different sets of values for the parameters that provide a solution for the equations. Such a model is said to be underidentified.

If the number of parameters equals the number of free elements in the covariance matrix, then there may exist a unique set of parameter estimates that exactly reproduce the observed covariance matrix. In this case, the model is said to be just identified or saturated.

If the number of parameters is less than the number of free elements in the covariance matrix, there may exist no set of parameter estimates that reproduces the observed covariance matrix. In this case, the model is said to be overidentified. Various statistical criteria, such as maximum likelihood, can be used to choose parameter estimates that approximately reproduce the observed covariance matrix. If you use ML or GLS estimation, PROC CALIS can perform a statistical test of the goodness of fit of the model under the assumption of multivariate normality of all variables and independence of the observations.

If the model is just identified or overidentified, it is said to be identified. If you use ML or GLS estimation for an identified model, PROC CALIS can compute approximate standard errors for the parameter estimates. For underidentified models, PROC CALIS obtains approximate standard errors by imposing additional constraints resulting from the use of a generalized inverse of the Hessian matrix.

You cannot guarantee that a model is identified simply by counting the parameters. For example, for any latent variable, you must specify a numeric value for the variance, or for some covariance involving the variable, or for a coefficient of the variable in at least one equation. Otherwise, the scale of the latent variable is indeterminate, and the model will be underidentified regardless of the number of parameters and the size of the covariance matrix. As another example, an exploratory factor analysis with two or more common factors is always underidentified because you can rotate the common factors without affecting the fit of the model.

PROC CALIS can usually detect an underidentified model by computing the approximate covariance matrix of the parameter estimates and checking whether any estimate is linearly related to other estimates (Bollen 1989, pp. 248–250), in which case PROC CALIS displays equations showing the linear relationships among the estimates. An-

other way to obtain empirical evidence regarding the identification of a model is to run the analysis several times with different initial estimates to see if the same final estimates are obtained.

Bollen (1989) provides detailed discussions of conditions for identification in a variety of models.

The following example is inspired by Fuller (1987, pp. 40–41). The hypothetical data are counts of two types of cells, cells forming rosettes and nucleated cells, in spleen samples. It is reasonable to assume that counts have a Poisson distribution; hence, the square roots of the counts should have a constant error variance of 0.25.

You can use PROC CALIS to fit a model of Form C to the square roots of the counts without constraints on the parameters, as displayed in following statements. The option OMETHOD=QUANEW is used in the PROC CALIS statement because in this case it produces more rapid convergence than the default optimization method.

```

data spleen;
  input rosette nucleate;
  sqrtrose=sqrt(rosette);
  sqrtnucl=sqrt(nucleate);
  datalines;
4 62
5 87
5 117
6 142
8 212
9 120
12 254
13 179
15 125
19 182
28 301
51 357
;

proc calis data=spleen cov omethod=quaneew;
  lineqs sqrtrose=factrose + err_rose,
         sqrtnucl=factnucl + err_nucl,
         factrose=beta factnucl;
  std err_rose=v_rose,
     err_nucl=v_nucl,
     factnucl=v_factnu;
run;

```

This model is underidentified. PROC CALIS displays the following warning:

```
WARNING: Problem not identified: More parameters to estimate ( 4 )
        than given values in data matrix ( 3 ).
```

and diagnoses the indeterminacy as follows:

```
NOTE: Hessian matrix is not full rank. Not all parameters are identified.
      Some parameter estimates are linearly related to other parameter
      estimates as shown in the following equations:
```

```
v_nucl  = -10.554977 - 0.036438 * beta + 1.00000 * v_factnu
          + 0.149564 * v_rose
```

The constraint that the error variances equal 0.25 can be imposed by modifying the STD statement:

```
proc calis data=spleen cov stderr;
  lineqs sqrtrose=factrose + err_rose,
         sqrtnucl=factnucl + err_nucl,
         factrose=beta factnucl;
  std err_rose=.25,
      err_nucl=.25,
      factnucl=v_factnu;
run;
```

The resulting parameter estimates are displayed in Figure 14.2.

| The CALIS Procedure | | | | |
|--|------------|-----------------|----------------|---------|
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | |
| | factrose = | 0.4034*factnucl | | |
| | Std Err | 0.0508 beta | | |
| | t Value | 7.9439 | | |
| Variances of Exogenous Variables | | | | |
| Variable | Parameter | Estimate | Standard Error | t Value |
| factnucl | v_factnu | 10.45846 | 4.56608 | 2.29 |
| err_rose | | 0.25000 | | |
| err_nucl | | 0.25000 | | |

Figure 14.2. Spleen Data: Parameter Estimates for Overidentified Model

This model is overidentified and the chi-square goodness-of-fit test yields a p -value of 0.0219, as displayed in Figure 14.3.

| The CALIS Procedure | |
|--|--------|
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.4775 |
| Goodness of Fit Index (GFI) | 0.7274 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.1821 |
| Root Mean Square Residual (RMR) | 0.1785 |
| Parsimonious GFI (Mulaik, 1989) | 0.7274 |
| Chi-Square | 5.2522 |
| Chi-Square DF | 1 |
| Pr > Chi-Square | 0.0219 |
| Independence Model Chi-Square | 13.273 |
| Independence Model Chi-Square DF | 1 |
| RMSEA Estimate | 0.6217 |
| RMSEA 90% Lower Confidence Limit | 0.1899 |
| RMSEA 90% Upper Confidence Limit | 1.1869 |
| ECVI Estimate | 0.9775 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 2.2444 |
| Probability of Close Fit | 0.0237 |
| Bentler's Comparative Fit Index | 0.6535 |
| Normal Theory Reweighted LS Chi-Square | 9.5588 |
| Akaike's Information Criterion | 3.2522 |
| Bozdogan's (1987) CAIC | 1.7673 |
| Schwarz's Bayesian Criterion | 2.7673 |
| McDonald's (1989) Centrality | 0.8376 |
| Bentler & Bonett's (1980) Non-normed Index | 0.6535 |
| Bentler & Bonett's (1980) NFI | 0.6043 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.6043 |
| Z-Test of Wilson & Hilferty (1931) | 2.0375 |
| Bollen (1986) Normed Index Rho1 | 0.6043 |
| Bollen (1988) Non-normed Index Delta2 | 0.6535 |
| Hoelter's (1983) Critical N | 10 |

Figure 14.3. Spleen Data: Fit Statistics for Overidentified Model

The sample size is so small that the p -value should not be taken to be accurate, but to get a small p -value with such a small sample indicates it is possible that the model is seriously deficient. The deficiency could be due to any of the following:

- The error variances are not both equal to 0.25.
- The error terms are correlated with each other or with the true scores.
- The observations are not independent.
- There is a disturbance in the linear relation between `factrose` and `factnucl`.
- The relation between `factrose` and `factnucl` is not linear.
- The actual distributions are not adequately approximated by the multivariate normal distribution.

A simple and plausible modification to the model is to add a “disturbance term” or “error in the equation” to the structural model, as follows.


```

proc calis data=spleen cov stderr;
  lineqs sqrtrose=factrose + err_rose,
        sqrtnucl=factnucl + err_nucl,
        factrose=beta factnucl + disturb;
  std err_rose=.25,
     err_nucl=.25,
     factnucl=v_factnu,
     disturb=v_dist;
run;

```

The following parameter estimates are produced.

| The CALIS Procedure | | | | |
|--|-------------------|----------|----------------|---------|
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | |
| factrose = | 0.3907*factnucl + | 1.0000 | disturb | |
| Std Err | 0.0771 | beta | | |
| t Value | 5.0692 | | | |
| Variances of Exogenous Variables | | | | |
| Variable | Parameter | Estimate | Standard Error | t Value |
| factnucl | v_factnu | 10.50458 | 4.58577 | 2.29 |
| err_rose | | 0.25000 | | |
| err_nucl | | 0.25000 | | |
| disturb | v_dist | 0.38153 | 0.28556 | 1.34 |

Figure 14.4. Spleen Data: Parameter Estimated for Just Identified Model

This model is just identified, so there are no degrees of freedom for the chi-square goodness-of-fit test.

Path Diagrams and the RAM Model

Complicated models are often easier to understand when they are expressed as path diagrams. One advantage of path diagrams over equations is that variances and covariances can be shown directly in the path diagram. Loehlin (1987) provides a detailed discussion of path diagrams.

It is customary to write the names of manifest variables in rectangles and names of latent variables in ovals. The coefficients in each equation are indicated by drawing arrows from the independent variables to the dependent variable. Covariances between exogenous variables are drawn as two-headed arrows. The variance of an exogenous variable can be displayed as a two-headed arrow with both heads pointing to the exogenous variable, since the variance of a variable is the covariance of the variable with itself. Here is a path diagram for the spleen data, explicitly showing all latent variables and variances of exogenous variables.

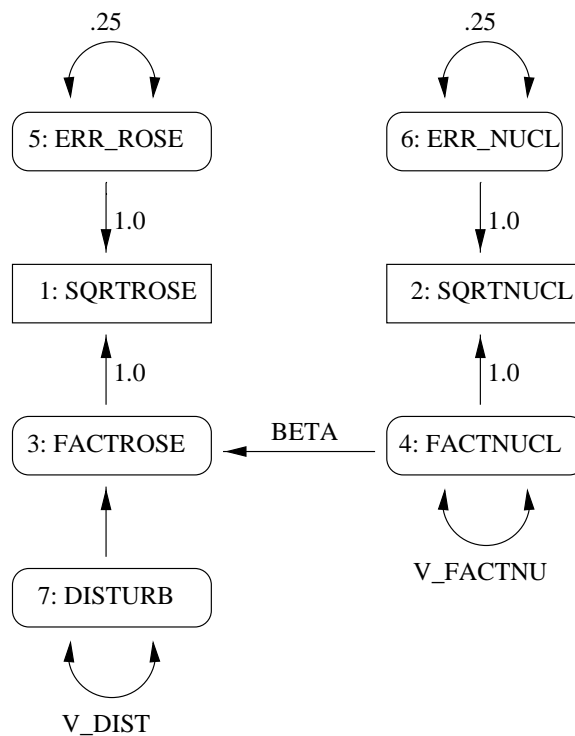


Figure 14.5. Path Diagram: Spleen

There is an easier way to draw the path diagram based on McArdle's reticular action model (RAM) (McArdle and McDonald 1984). McArdle uses the convention that a two-headed arrow that points to an endogenous variable actually refers to the error or disturbance term associated with that variable. A two-headed arrow with both heads pointing to the same endogenous variable represents the error or disturbance variance for the equation that determines the endogenous variable; there is no need to draw a separate oval for the error or disturbance term. Similarly, a two-headed arrow connecting two endogenous variables represents the covariance between the error of disturbance terms associated with the endogenous variables. The RAM conventions allow the previous path diagram to be simplified, as follows.

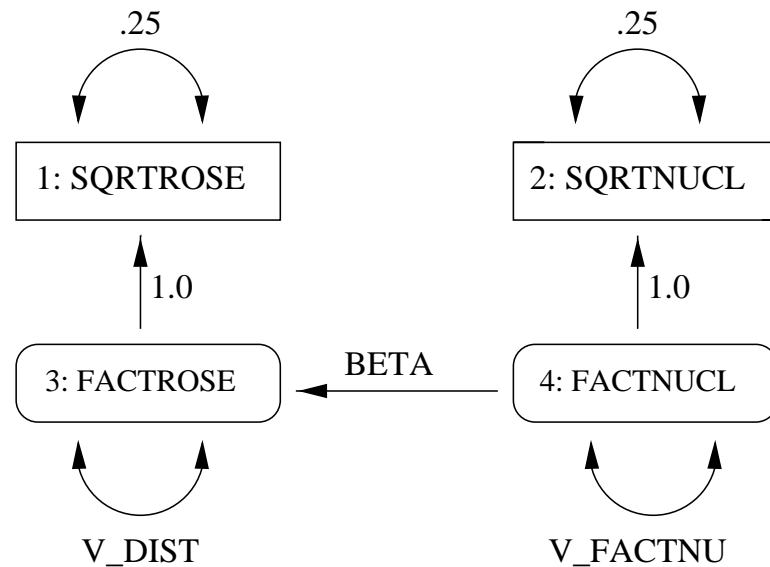


Figure 14.6. Path Diagram: Spleen

The RAM statement in PROC CALIS provides a simple way to transcribe a path diagram based on the reticular action model. Assign the integers 1, 2, 3, . . . to the variables in the order in which they appear in the SAS data set or in the VAR statement, if you use one. Assign subsequent consecutive integers to the latent variables displayed explicitly in the path diagram (excluding the error and disturbance terms implied by two-headed arrows) in any order. Each arrow in the path diagram can then be identified by two numbers indicating the variables connected by the path. The RAM statement consists of a list of descriptions of all the arrows in the path diagram. The descriptions are separated by commas. Each arrow description consists of three or four numbers and, optionally, a name in the following order:

1. The number of heads the arrow has.
2. The number of the variable the arrow points to, or either variable if the arrow is two-headed.
3. The number of the variable the arrow comes from, or the other variable if the arrow is two-headed.
4. The value of the coefficient or (co)variance that the arrow represents.

5. A name if the arrow represents a parameter to be estimated, in which case the previous number is taken to be the initial estimate of the parameter. Omit the name if the arrow represents a constant. If you specify a name, the fourth number may be omitted.

The model for the spleen data can be specified with the RAM statement, as follows:

```

/* 1 sqrtrose */
/* 2 sqrtnucl */
/* 3 factrose */
/* 4 factnucl */
proc calis data=spleen cov stderr method=ml outram=splram1;
var sqrtrose sqrtnucl;
ram 1 1 3 1,          /* sqrtrose <- factrose */
    1 2 4 1,          /* sqrtnucl <- factnucl */
    1 3 4 beta,       /* factrose <- factnucl */
    2 1 1 .25,        /* error variance for sqrtrose */
    2 2 2 .25,        /* error variance for sqrtnucl */
    2 3 3 v_dist,    /* disturbance variance for factrose */
    2 4 4 v_factnu;  /* variance of factnucl */
run;

```

The resulting output in RAM form is displayed in Figure 14.7.

| The CALIS Procedure | | | | | | | |
|--|--------|-------------|--------------|------------|----------|-------------------|---------|
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | | | | |
| RAM Estimates | | | | | | | |
| Term | Matrix | ----Row---- | ---Column--- | Parameter | Estimate | Standard Error | t Value |
| 1 | 2 | sqrtrose | 1 F1 | 3 . | 1.00000 | | |
| 1 | 2 | sqrtnucl | 2 F2 | 4 . | 1.00000 | | |
| 1 | 2 | F1 | 3 F2 | 4 beta | 0.39074 | 0.07708 | 5.07 |
| 1 | 3 | E1 | 1 E1 | 1 . | 0.25000 | | |
| 1 | 3 | E2 | 2 E2 | 2 . | 0.25000 | | |
| 1 | 3 | D1 | 3 D1 | 3 v_dist | 0.38153 | 0.28556 | 1.34 |
| 1 | 3 | D2 | 4 D2 | 4 v_factnu | 10.50458 | 4.58577 | 2.29 |

Figure 14.7. Spleen Data: RAM Model

You can request an output data set containing the model specification by using the OUTRAM= option in the PROC CALIS statement. Names for the latent variables can be specified in a V NAMES statement.

```

proc calis data=spleen cov stderr method=ml outram=splram1;
  var sqrtrose sqrtnucl;
  vnames 1 factrose factnucl,
         2 err_rose err_nucl disturb factnucl;
  ram 1 1 3 1,          /* sqrtrose <- factrose */
      1 2 4 1,          /* sqrtnucl <- factnucl */
      1 3 4 beta,      /* factrose <- factnucl */
      2 1 1 .25,       /* error variance for sqrtrose */

                        2 2 2 .25,          /* error variance for sqrtnucl */
                        2 3 3 v_dist,      /* disturbance variance for factrose */
                        2 4 4 v_factnu;    /* variance of factnucl */
run;

proc print;
run;

```

The RAM output is displayed in Figure 14.8.

| The CALIS Procedure | | | | | | | |
|--|--------|--------------|--------------|-----------|------------|-------------------|--------------|
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | | | | |
| RAM Estimates | | | | | | | |
| Term | Matrix | ----Row----- | ---Column--- | Parameter | Estimate | Standard Error | t Value |
| 1 | 2 | sqrtrose | 1 | factrose | 3 . | 1.00000 | |
| 1 | 2 | sqrtnucl | 2 | factnucl | 4 . | 1.00000 | |
| 1 | 2 | factrose | 3 | factnucl | 4 beta | 0.39074 | 0.07708 5.07 |
| 1 | 3 | err_rose | 1 | err_rose | 1 . | 0.25000 | |
| 1 | 3 | err_nucl | 2 | err_nucl | 2 . | 0.25000 | |
| 1 | 3 | disturb | 3 | disturb | 3 v_dist | 0.38153 | 0.28556 1.34 |
| 1 | 3 | factnucl | 4 | factnucl | 4 v_factnu | 10.50458 | 4.58577 2.29 |

Figure 14.8. Spleen Data: RAM Model with Names for Latent Variables

The OUTRAM= data set contains the RAM model as you specified it in the RAM statement, but it contains the final parameter estimates and standard errors instead of the initial values.

| Obs | _TYPE_ | _NAME_ | _MATNR_ | _ROW_ | _COL_ | _ESTIM_ | _STDERR_ |
|-----|---------|----------|---------|-------|-------|---------|----------|
| 1 | MODEL | _IDE_ | 1 | 2 | 4 | 1.0000 | 0.00000 |
| 2 | MODEL | _A_ | 2 | 4 | 4 | 6.0000 | 2.00000 |
| 3 | MODEL | _P_ | 3 | 4 | 4 | 3.0000 | 0.00000 |
| 4 | VARNAME | sqrtrose | 2 | . | 1 | . | . |
| 5 | VARNAME | sqrtnucl | 2 | . | 2 | . | . |
| 6 | VARNAME | factrose | 2 | . | 3 | . | . |
| 7 | VARNAME | factnucl | 2 | . | 4 | . | . |
| 8 | VARNAME | err_rose | 3 | . | 1 | . | . |
| 9 | VARNAME | err_nucl | 3 | . | 2 | . | . |
| 10 | VARNAME | disturb | 3 | . | 3 | . | . |
| 11 | VARNAME | factnucl | 3 | . | 4 | . | . |
| 12 | METHOD | ML | . | . | . | . | . |
| 13 | STAT | N | . | . | . | 12.0000 | . |
| 14 | STAT | FIT | . | . | . | 0.0000 | . |
| 15 | STAT | GFI | . | . | . | 1.0000 | . |
| 16 | STAT | AGFI | . | . | . | . | . |
| 17 | STAT | RMR | . | . | . | 0.0000 | . |
| 18 | STAT | PGFI | . | . | . | 0.0000 | . |
| 19 | STAT | NPARAM | . | . | . | 3.0000 | . |
| 20 | STAT | DF | . | . | . | 0.0000 | . |
| 21 | STAT | N_ACT | . | . | . | 0.0000 | . |
| 22 | STAT | CHISQUAR | . | . | . | 0.0000 | . |
| 23 | STAT | P_CHISQ | . | . | . | 0.0000 | . |
| 24 | STAT | CHISQNUL | . | . | . | 13.2732 | . |
| 25 | STAT | RMSEAEST | . | . | . | 0.0000 | . |
| 26 | STAT | RMSEALOB | . | . | . | . | . |
| 27 | STAT | RMSEAUPB | . | . | . | . | . |
| 28 | STAT | P_CLOSFT | . | . | . | . | . |
| 29 | STAT | ECVI_EST | . | . | . | 0.7500 | . |
| 30 | STAT | ECVI_LOB | . | . | . | . | . |
| 31 | STAT | ECVI_UPB | . | . | . | . | . |
| 32 | STAT | COMPFITI | . | . | . | 1.0000 | . |
| 33 | STAT | ADJCHISQ | . | . | . | . | . |
| 34 | STAT | P_ACHISQ | . | . | . | . | . |
| 35 | STAT | RLSCHISQ | . | . | . | 0.0000 | . |
| 36 | STAT | AIC | . | . | . | 0.0000 | . |
| 37 | STAT | CAIC | . | . | . | 0.0000 | . |
| 38 | STAT | SBC | . | . | . | 0.0000 | . |
| 39 | STAT | CENTRALI | . | . | . | 1.0000 | . |
| 40 | STAT | BB_NONOR | . | . | . | . | . |
| 41 | STAT | BB_NORMD | . | . | . | 1.0000 | . |
| 42 | STAT | PARSIMON | . | . | . | 0.0000 | . |
| 43 | STAT | ZTESTWH | . | . | . | . | . |
| 44 | STAT | BOL_RHO1 | . | . | . | . | . |
| 45 | STAT | BOL_DEL2 | . | . | . | 1.0000 | . |
| 46 | STAT | CNHOELT | . | . | . | . | . |
| 47 | ESTIM | | 2 | 1 | 3 | 1.0000 | 0.00000 |
| 48 | ESTIM | | 2 | 2 | 4 | 1.0000 | 0.00000 |
| 49 | ESTIM | beta | 2 | 3 | 4 | 0.3907 | 0.07708 |
| 50 | ESTIM | | 3 | 1 | 1 | 0.2500 | 0.00000 |
| 51 | ESTIM | | 3 | 2 | 2 | 0.2500 | 0.00000 |
| 52 | ESTIM | v_dist | 3 | 3 | 3 | 0.3815 | 0.28556 |
| 53 | ESTIM | v_factnu | 3 | 4 | 4 | 10.5046 | 4.58577 |

Figure 14.9. Spleen Data: OUTRAM= Data Set with Final Parameter Estimates

This data set can be used as input to another run of PROC CALIS with the INRAM= option in the PROC CALIS statement. For example, if the iteration limit is exceeded, you can use the RAM data set to start a new run that begins with the final estimates from the last run. Or you can change the data set to add or remove constraints or modify the model in various other ways. The easiest way to change a RAM data set is to use the FSEDIT procedure, but you can also use a DATA step. For example, you

could set the variance of the disturbance term to zero, effectively removing the disturbance from the equation, by removing the parameter name `v_dist` in the `_NAME_` variable and setting the value of the estimate to zero in the `_ESTIM_` variable:

```
data splram2(type=ram);
  set splram1;
  if _name_='v_dist' then
    do;
      _name_=' ';
      _estim=0;
    end;
run;

proc calis data=spleen inram=splram2 cov stderr;
run;
```

The resulting RAM output is displayed in Figure 14.10.

| The CALIS Procedure | | | | | | | |
|--|--------|--------------|--------------|-----------|------------|-------------------|---------|
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | | | | |
| RAM Estimates | | | | | | | |
| Term | Matrix | ----Row----- | ---Column--- | Parameter | Estimate | Standard Error | t Value |
| 1 | 2 | sqrtrose | 1 | factrose | 3 . | 1.00000 | |
| 1 | 2 | sqrtnucl | 2 | factnucl | 4 . | 1.00000 | |
| 1 | 2 | factrose | 3 | factnucl | 4 beta | 0.40340 | 7.94 |
| 1 | 3 | err_rose | 1 | err_rose | 1 . | 0.25000 | |
| 1 | 3 | err_nucl | 2 | err_nucl | 2 . | 0.25000 | |
| 1 | 3 | disturb | 3 | disturb | 3 . | 0 | |
| 1 | 3 | factnucl | 4 | factnucl | 4 v_factnu | 10.45846 | 2.29 |

Figure 14.10. Spleen Data: RAM Model with INRAM= Data Set

Some Measurement Models

Psychometric test theory involves many kinds of models relating scores on psychological and educational tests to latent variables representing intelligence or various underlying abilities. The following example uses data on four vocabulary tests from Lord (1957). Tests W and X have 15 items each and are administered with very liberal time limits. Tests Y and Z have 75 items and are administered under time pressure. The covariance matrix is read by the following DATA step:

```
data lord(type=cov);
  input _type_ $ _name_ $ w x y z;
  datalines;
n . 649 . . .
cov w 86.3979 . . .
cov x 57.7751 86.2632 . .
cov y 56.8651 59.3177 97.2850 .
cov z 58.8986 59.6683 73.8201 97.8192
;
```

The psychometric model of interest states that W and X are determined by a single common factor F_{WX} , and Y and Z are determined by a single common factor F_{YZ} . The two common factors are expected to have a positive correlation, and it is desired to estimate this correlation. It is convenient to assume that the common factors have unit variance, so their correlation will be equal to their covariance. The error terms for all the manifest variables are assumed to be uncorrelated with each other and with the common factors. The model (labeled here as Model Form D) is as follows.

Model Form D

$$\begin{aligned}
 W &= \beta_W F_{WX} + E_W \\
 X &= \beta_X F_{WX} + E_X \\
 Y &= \beta_Y F_{YZ} + E_Y \\
 Z &= \beta_Z F_{YZ} + E_Z \\
 \text{Var}(F_{WX}) &= \text{Var}(F_{YZ}) = 1 \\
 \text{Cov}(F_{WX}, F_{YZ}) &= \rho \\
 \text{Cov}(E_W, E_X) &= \text{Cov}(E_W, E_Y) = \text{Cov}(E_W, E_Z) = \text{Cov}(E_X, E_Y) \\
 &= \text{Cov}(E_X, E_Z) = \text{Cov}(E_Y, E_Z) = \text{Cov}(E_W, F_{WX}) \\
 &= \text{Cov}(E_W, F_{YZ}) = \text{Cov}(E_X, F_{WX}) = \text{Cov}(E_X, F_{YZ}) \\
 &= \text{Cov}(E_Y, F_{WX}) = \text{Cov}(E_Y, F_{YZ}) = \text{Cov}(E_Z, F_{WX}) \\
 &= \text{Cov}(E_Z, F_{YZ}) = 0
 \end{aligned}$$

The corresponding path diagram is as follows.

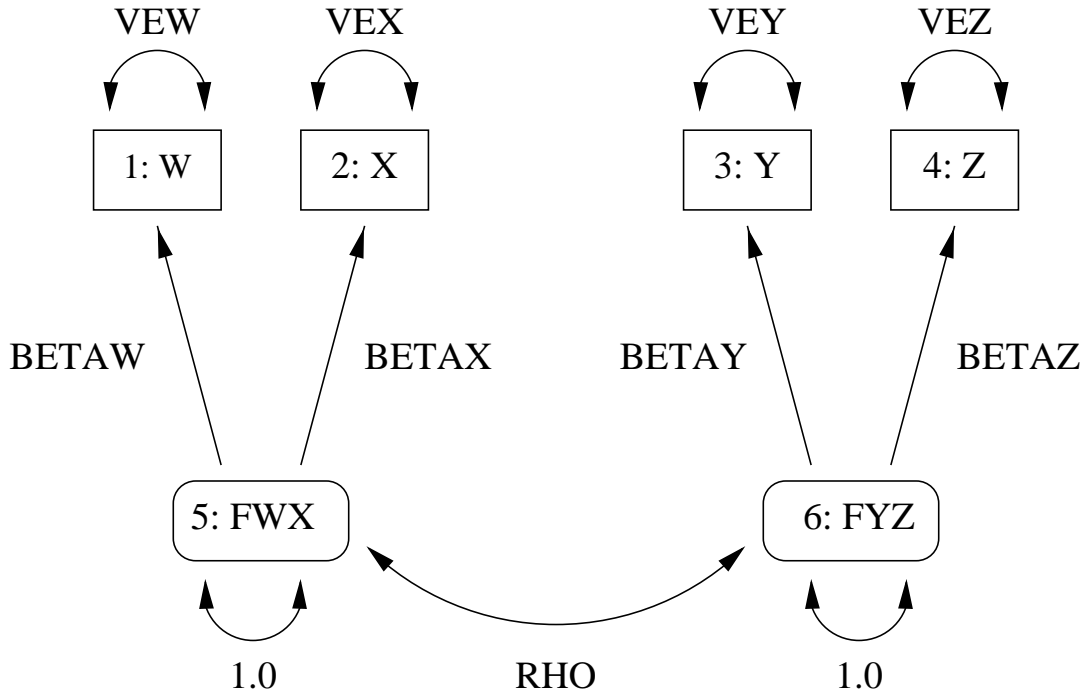


Figure 14.11. Path Diagram: Lord

This path diagram can be converted to a RAM model as follows:

```

/* 1=w 2=x 3=y 4=z 5=fwx 6=fyz */
title 'H4: unconstrained';
proc calis data=lord cov;
  ram 1 1 5 betaw,
      1 2 5 betax,
      1 3 6 betay,
      1 4 6 betaz,
      2 1 1 vew,
      2 2 2 vex,
      2 3 3 vey,
      2 4 4 vez,
      2 5 5 1,
      2 6 6 1,
      2 5 6 rho;
run;

```

Here are the major results.

```

H4: unconstrained

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Fit Function                                0.0011
Goodness of Fit Index (GFI)                 0.9995
GFI Adjusted for Degrees of Freedom (AGFI)  0.9946
Root Mean Square Residual (RMR)            0.2720
Parsimonious GFI (Mulaik, 1989)           0.1666
Chi-Square                                  0.7030
Chi-Square DF                              1
Pr > Chi-Square                             0.4018
Independence Model Chi-Square               1466.6
Independence Model Chi-Square DF           6
RMSEA Estimate                              0.0000
RMSEA 90% Lower Confidence Limit            .
RMSEA 90% Upper Confidence Limit           0.0974
ECVI Estimate                               0.0291
ECVI 90% Lower Confidence Limit            .
ECVI 90% Upper Confidence Limit           0.0391
Probability of Close Fit                    0.6854
Bentler's Comparative Fit Index            1.0000
Normal Theory Reweighted LS Chi-Square     0.7026
Akaike's Information Criterion              -1.2970
Bozdogan's (1987) CAIC                     -6.7725
Schwarz's Bayesian Criterion                -5.7725
McDonald's (1989) Centrality               1.0002
Bentler & Bonett's (1980) Non-normed Index 1.0012
Bentler & Bonett's (1980) NFI              0.9995
James, Mulaik, & Brett (1982) Parsimonious NFI 0.1666
Z-Test of Wilson & Hilferty (1931)        0.2363
Bollen (1986) Normed Index Rho1           0.9971
Bollen (1988) Non-normed Index Delta2      1.0002
Hoelter's (1983) Critical N                 3543

```

Figure 14.12. Lord Data: Major Results for RAM Model, Hypothesis H4

```

H4: unconstrained

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

RAM Estimates

Term Matrix --Row-- -Column- Parameter Estimate Standard Error t Value
1 2 w 1 F1 5 betaw 7.50066 0.32339 23.19
1 2 x 2 F1 5 betax 7.70266 0.32063 24.02
1 2 y 3 F2 6 betay 8.50947 0.32694 26.03
1 2 z 4 F2 6 betaz 8.67505 0.32560 26.64
1 3 E1 1 E1 1 vew 30.13796 2.47037 12.20
1 3 E2 2 E2 2 vex 26.93217 2.43065 11.08
1 3 E3 3 E3 3 vey 24.87396 2.35986 10.54
1 3 E4 4 E4 4 vez 22.56264 2.35028 9.60
1 3 D1 5 D1 5 . 1.00000
1 3 D2 6 D1 5 rho 0.89855 0.01865 48.18
1 3 D2 6 D2 6 . 1.00000

```

The same analysis can be performed with the LINEQS statement. Subsequent analyses are illustrated with the LINEQS statement rather than the RAM statement because it is slightly easier to understand the constraints as written in the LINEQS statement without constantly referring to the path diagram. The LINEQS and RAM statements may yield slightly different results due to the inexactness of the numerical optimization; the discrepancies can be reduced by specifying a more stringent convergence criterion such as GCONV=1E-4 or GCONV=1E-6. It is convenient to create an OUTRAM= data set for use in fitting other models with additional constraints.

```
title 'H4: unconstrained';
proc calis data=lord cov outram=ram4;
  lineqs w=betaw fwx + ew,
         x=betax fwx + ex,
         y=betay fyz + ey,
         z=betaz fyz + ez;
  std fwx fyz=1,
      ew ex ey ez=vew vex vey vez;
  cov fwx fyz=rho;
run;
```

The LINEQS displayed output is as follows.

| H4: unconstrained | | | | | |
|--|-----------|--------------|----------------|----------------|---------|
| The CALIS Procedure | | | | | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | | |
| w | = | 7.5007*fwx | + | 1.0000 ew | |
| Std Err | | 0.3234 betaw | | | |
| t Value | | 23.1939 | | | |
| x | = | 7.7027*fwx | + | 1.0000 ex | |
| Std Err | | 0.3206 betax | | | |
| t Value | | 24.0235 | | | |
| y | = | 8.5095*fyz | + | 1.0000 ey | |
| Std Err | | 0.3269 betay | | | |
| t Value | | 26.0273 | | | |
| z | = | 8.6751*fyz | + | 1.0000 ez | |
| Std Err | | 0.3256 betaz | | | |
| t Value | | 26.6430 | | | |
| Variances of Exogenous Variables | | | | | |
| Variable | Parameter | Estimate | Standard Error | t Value | |
| fwx | | 1.00000 | | | |
| fyz | | 1.00000 | | | |
| ew | vew | 30.13796 | 2.47037 | 12.20 | |
| ex | vex | 26.93217 | 2.43065 | 11.08 | |
| ey | vey | 24.87396 | 2.35986 | 10.54 | |
| ez | vez | 22.56264 | 2.35028 | 9.60 | |
| Covariances Among Exogenous Variables | | | | | |
| Var1 | Var2 | Parameter | Estimate | Standard Error | t Value |
| fwx | fyz | rho | 0.89855 | 0.01865 | 48.18 |

Figure 14.13. Lord Data: Using LINEQS Statement for RAM Model, Hypothesis H4

In an analysis of these data by Jöreskog and Sörbom (1979, pp. 54–56; Loehlin 1987, pp. 84–87), four hypotheses are considered:

$$H_1: \rho = 1,$$

$$\beta_W = \beta_X, \text{ Var}(E_W) = \text{Var}(E_X),$$

$$\beta_Y = \beta_Z, \text{ Var}(E_Y) = \text{Var}(E_Z)$$

H_2 : same as H_1 : except ρ is unconstrained

$$H_3: \rho = 1$$

H_4 : Model Form D without any additional constraints

The hypothesis H_3 says that there is really just one common factor instead of two; in the terminology of test theory, W, X, Y, and Z are said to be congeneric. The hypothesis H_2 says that W and X have the same true-scores and have equal error variance; such tests are said to be parallel. The hypothesis H_2 also requires Y and Z to be parallel. The hypothesis H_1 says that W and X are parallel tests, Y and Z are parallel tests, and all four tests are congeneric.

It is most convenient to fit the models in the opposite order from that in which they are numbered. The previous analysis fit the model for H_4 and created an OUTRAM= data set called ram4. The hypothesis H_3 can be fitted directly or by modifying the ram4 data set. Since H_3 differs from H_4 only in that ρ is constrained to equal 1, the ram4 data set can be modified by finding the observation for which `_NAME_='rho'` and changing the variable `_NAME_` to a blank value (meaning that the observation represents a constant rather than a parameter to be fitted) and setting the variable `_ESTIM_` to the value 1. Both of the following analyses produce the same results:

```

title 'H3: W, X, Y, and Z are congeneric';
proc calis data=lord cov;
  lineqs w=betaw f + ew,
         x=betax f + ex,
         y=betay f + ey,
         z=betaz f + ez;
  std f=1,
      ew ex ey ez=vew vex vey vez;
run;

data ram3(type=ram);
  set ram4;
  if _name_='rho' then
    do;
      _name_=' ';
      _estim_=1;
    end;
run;

proc calis data=lord inram=ram3 cov;
run;

```

The resulting output from either of these analyses is displayed in Figure 14.14.

| H3: W, X, Y, and Z are congeneric | |
|--|---------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0559 |
| Goodness of Fit Index (GFI) | 0.9714 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.8570 |
| Root Mean Square Residual (RMR) | 2.4636 |
| Parsimonious GFI (Mulaik, 1989) | 0.3238 |
| Chi-Square | 36.2095 |
| Chi-Square DF | 2 |
| Pr > Chi-Square | <.0001 |
| Independence Model Chi-Square | 1466.6 |
| Independence Model Chi-Square DF | 6 |
| RMSEA Estimate | 0.1625 |
| RMSEA 90% Lower Confidence Limit | 0.1187 |
| RMSEA 90% Upper Confidence Limit | 0.2108 |
| ECVI Estimate | 0.0808 |
| ECVI 90% Lower Confidence Limit | 0.0561 |
| ECVI 90% Upper Confidence Limit | 0.1170 |
| Probability of Close Fit | 0.0000 |
| Bentler's Comparative Fit Index | 0.9766 |
| Normal Theory Reweighted LS Chi-Square | 38.1432 |
| Akaike's Information Criterion | 32.2095 |
| Bozdogan's (1987) CAIC | 21.2586 |
| Schwarz's Bayesian Criterion | 23.2586 |
| McDonald's (1989) Centrality | 0.9740 |
| Bentler & Bonett's (1980) Non-normed Index | 0.9297 |
| Bentler & Bonett's (1980) NFI | 0.9753 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.3251 |
| Z-Test of Wilson & Hilferty (1931) | 5.2108 |
| Bollen (1986) Normed Index Rho1 | 0.9259 |
| Bollen (1988) Non-normed Index Delta2 | 0.9766 |
| Hoelter's (1983) Critical N | 109 |

Figure 14.14. Lord Data: Major Results for Hypothesis H3

| H3: W, X, Y, and Z are congeneric | | | | |
|--|-----------|--------------|----------------|-----------|
| The CALIS Procedure | | | | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | |
| w | = | 7.1047*fwx | + | 1.0000 ew |
| Std Err | | 0.3218 betaw | | |
| t Value | | 22.0802 | | |
| x | = | 7.2691*fwx | + | 1.0000 ex |
| Std Err | | 0.3183 betax | | |
| t Value | | 22.8397 | | |
| y | = | 8.3735*fyz | + | 1.0000 ey |
| Std Err | | 0.3254 betay | | |
| t Value | | 25.7316 | | |
| z | = | 8.5106*fyz | + | 1.0000 ez |
| Std Err | | 0.3241 betaz | | |
| t Value | | 26.2598 | | |
| Variances of Exogenous Variables | | | | |
| Variable | Parameter | Estimate | Standard Error | t Value |
| fwx | | 1.00000 | | |
| fyz | | 1.00000 | | |
| ew | vew | 35.92087 | 2.41466 | 14.88 |
| ex | vex | 33.42397 | 2.31038 | 14.47 |
| ey | vey | 27.16980 | 2.24619 | 12.10 |
| ez | vez | 25.38948 | 2.20839 | 11.50 |

The hypothesis H_2 requires that several pairs of parameters be constrained to have equal estimates. With PROC CALIS, you can impose this constraint by giving the same name to parameters that are constrained to be equal. This can be done directly in the LINEQS and STD statements or by using PROC FSEDIT or a DATA step to change the values in the ram4 data set:

```

title 'H2: W and X parallel, Y and Z parallel';
proc calis data=lord cov;
  lineqs w=betawx fwx + ew,
         x=betawx fwx + ex,
         y=betayz fyz + ey,
         z=betayz fyz + ez;
  std fwx fyz=1,
      ew ex ey ez=vewx vewx veyz veyz;
  cov fwx fyz=rho;
run;

data ram2(type=ram);
  set ram4;
  if _name_ = 'betaw' then _name_ = 'betawx';
  if _name_ = 'betax' then _name_ = 'betawx';
  if _name_ = 'betay' then _name_ = 'betayz';
  if _name_ = 'betaz' then _name_ = 'betayz';
  if _name_ = 'vew' then _name_ = 'vewx';

```

```

if _name_='vex' then _name_='vewx';
if _name_='vey' then _name_='veyz';
if _name_='vez' then _name_='veyz';
run;

proc calis data=lord inram=ram2 cov;
run;

```

The resulting output from either of these analyses is displayed in Figure 14.15.

| H2: W and X parallel, Y and Z parallel | |
|--|----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0030 |
| Goodness of Fit Index (GFI) | 0.9985 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9970 |
| Root Mean Square Residual (RMR) | 0.6983 |
| Parsimonious GFI (Mulaik, 1989) | 0.8321 |
| Chi-Square | 1.9335 |
| Chi-Square DF | 5 |
| Pr > Chi-Square | 0.8583 |
| Independence Model Chi-Square | 1466.6 |
| Independence Model Chi-Square DF | 6 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0293 |
| ECVI Estimate | 0.0185 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.0276 |
| Probability of Close Fit | 0.9936 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 1.9568 |
| Akaike's Information Criterion | -8.0665 |
| Bozdogan's (1987) CAIC | -35.4436 |
| Schwarz's Bayesian Criterion | -30.4436 |
| McDonald's (1989) Centrality | 1.0024 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0025 |
| Bentler & Bonett's (1980) NFI | 0.9987 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.8322 |
| Z-Test of Wilson & Hilferty (1931) | -1.0768 |
| Bollen (1986) Normed Index Rho1 | 0.9984 |
| Bollen (1988) Non-normed Index Delta2 | 1.0021 |
| Hoelter's (1983) Critical N | 3712 |

Figure 14.15. Lord Data: Major Results for Hypothesis H2

| H2: W and X parallel, Y and Z parallel | | | | | |
|--|-----------|---------------|----------------|----------------|---------|
| The CALIS Procedure | | | | | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | | | | | |
| w | = | 7.6010*fwx | + | 1.0000 ew | |
| Std Err | | 0.2684 betawx | | | |
| t Value | | 28.3158 | | | |
| x | = | 7.6010*fwx | + | 1.0000 ex | |
| Std Err | | 0.2684 betawx | | | |
| t Value | | 28.3158 | | | |
| y | = | 8.5919*fyz | + | 1.0000 ey | |
| Std Err | | 0.2797 betayz | | | |
| t Value | | 30.7215 | | | |
| z | = | 8.5919*fyz | + | 1.0000 ez | |
| Std Err | | 0.2797 betayz | | | |
| t Value | | 30.7215 | | | |
| Variances of Exogenous Variables | | | | | |
| Variable | Parameter | Estimate | Standard Error | t Value | |
| fwx | | 1.00000 | | | |
| fyz | | 1.00000 | | | |
| ew | vewx | 28.55545 | 1.58641 | 18.00 | |
| ex | vewx | 28.55545 | 1.58641 | 18.00 | |
| ey | veyz | 23.73200 | 1.31844 | 18.00 | |
| ez | veyz | 23.73200 | 1.31844 | 18.00 | |
| Covariances Among Exogenous Variables | | | | | |
| Var1 | Var2 | Parameter | Estimate | Standard Error | t Value |
| fwx | fyz | rho | 0.89864 | 0.01865 | 48.18 |

The hypothesis H_1 requires one more constraint in addition to those in H_2 :

```

title 'H1: W and X parallel, Y and Z parallel, all congeneric';
proc calis data=lord cov;
  lineqs w=betawx f + ew,
         x=betawx f + ex,
         y=betayz f + ey,
         z=betayz f + ez;
  std f=1,
      ew ex ey ez=vevx vewx veyz veyz;
run;

data ram1(type=ram);
  set ram2;
  if _name_='rho' then
  do;
    _name_=' ';
    _estim_=1;
  end;
run;

```

```
proc calis data=lord inram=raml cov;
run;
```

The resulting output from either of these analyses is displayed in Figure 14.16.

| H1: W and X parallel, Y and Z parallel, all congeneric | |
|--|---------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0576 |
| Goodness of Fit Index (GFI) | 0.9705 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9509 |
| Root Mean Square Residual (RMR) | 2.5430 |
| Parsimonious GFI (Mulaik, 1989) | 0.9705 |
| Chi-Square | 37.3337 |
| Chi-Square DF | 6 |
| Pr > Chi-Square | <.0001 |
| Independence Model Chi-Square | 1466.6 |
| Independence Model Chi-Square DF | 6 |
| RMSEA Estimate | 0.0898 |
| RMSEA 90% Lower Confidence Limit | 0.0635 |
| RMSEA 90% Upper Confidence Limit | 0.1184 |
| ECVI Estimate | 0.0701 |
| ECVI 90% Lower Confidence Limit | 0.0458 |
| ECVI 90% Upper Confidence Limit | 0.1059 |
| Probability of Close Fit | 0.0076 |
| Bentler's Comparative Fit Index | 0.9785 |
| Normal Theory Reweighted LS Chi-Square | 39.3380 |
| Akaike's Information Criterion | 25.3337 |
| Bozdogan's (1987) CAIC | -7.5189 |
| Schwarz's Bayesian Criterion | -1.5189 |
| McDonald's (1989) Centrality | 0.9761 |
| Bentler & Bonett's (1980) Non-normed Index | 0.9785 |
| Bentler & Bonett's (1980) NFI | 0.9745 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.9745 |
| Z-Test of Wilson & Hilferty (1931) | 4.5535 |
| Bollen (1986) Normed Index Rho1 | 0.9745 |
| Bollen (1988) Non-normed Index Delta2 | 0.9785 |
| Hoelter's (1983) Critical N | 220 |

Figure 14.16. Lord Data: Major Results for Hypothesis H1

```

H1: W and X parallel, Y and Z parallel, all congeneric

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

w      = 7.1862*fwx      + 1.0000 ew
Std Err 0.2660 betawx
t Value 27.0180
x      = 7.1862*fwx      + 1.0000 ex
Std Err 0.2660 betawx
t Value 27.0180
y      = 8.4420*fyz      + 1.0000 ey
Std Err 0.2800 betayz
t Value 30.1494
z      = 8.4420*fyz      + 1.0000 ez
Std Err 0.2800 betayz
t Value 30.1494

Variances of Exogenous Variables

Variable Parameter      Estimate      Standard
                                Error      t Value

fwx                                1.00000
fyz                                1.00000
ew      vewx      34.68865      1.64634      21.07
ex      vewx      34.68865      1.64634      21.07
ey      veyz      26.28513      1.39955      18.78
ez      veyz      26.28513      1.39955      18.78

Covariances Among Exogenous Variables

Var1 Var2 Parameter      Estimate      Standard
                                Error      t Value

fwx  fyz                                1.00000

```

The goodness-of-fit tests for the four hypotheses are summarized in the following table.

| Hypothesis | Number of Parameters | χ^2 | Degrees of Freedom | <i>p</i> -value | $\hat{\rho}$ |
|------------|----------------------|----------|--------------------|-----------------|--------------|
| H_1 | 4 | 37.33 | 6 | 0.0000 | 1.0 |
| H_2 | 5 | 1.93 | 5 | 0.8583 | 0.8986 |
| H_3 | 8 | 36.21 | 2 | 0.0000 | 1.0 |
| H_4 | 9 | 0.70 | 1 | 0.4018 | 0.8986 |

The hypotheses H_1 and H_3 , which posit $\rho = 1$, can be rejected. Hypotheses H_2 and H_4 seem to be consistent with the available data. Since H_2 is obtained by adding four constraints to H_4 , you can test H_2 versus H_4 by computing the differences of the chi-square statistics and their degrees of freedom, yielding a chi-square of 1.23 with four degrees of freedom, which is obviously not significant. So hypothesis H_2 is consistent with the available data.

The estimates of ρ for H_2 and H_4 are almost identical, about 0.90, indicating that the speeded and unspeeded tests are measuring almost the same latent variable, even

though the hypotheses that stated they measured exactly the same latent variable are rejected.

A Combined Measurement-Structural Model with Reciprocal Influence and Correlated Residuals

To illustrate a more complex model, this example uses some well-known data from Haller and Butterworth (1960). Various models and analyses of these data are given by Duncan, Haller, and Portes (1968), Jöreskog and Sörbom (1988), and Loehlin (1987).

The study is concerned with the career aspirations of high-school students and how these aspirations are affected by close friends. The data are collected from 442 seventeen-year-old boys in Michigan. There are 329 boys in the sample who named another boy in the sample as a best friend. The observations to be analyzed consist of the data from these 329 boys paired with the data from their best friends.

The method of data collection introduces two statistical problems. First, restricting the analysis to boys whose best friends are in the original sample causes the reduced sample to be biased. Second, since the data from a given boy may appear in two or more observations, the observations are not independent. Therefore, any statistical conclusions should be considered tentative. It is difficult to accurately assess the effects of the dependence of the observations on the analysis, but it could be argued on intuitive grounds that since each observation has data from two boys and since it seems likely that many of the boys will appear in the data set at least twice, the effective sample size may be as small as half of the reported 329 observations.

The correlation matrix is taken from Jöreskog and Sörbom (1988).

```

title 'Peer Influences on Aspiration: Haller & Butterworth (1960)';
data aspire(type=corr);
  _type_='corr';
  input _name_ $ riq rpa rses roa rea fiq fpa fses foa fea;
  label riq='Respondent: Intelligence'
        rpa='Respondent: Parental Aspiration'
        rses='Respondent: Family SES'
        roa='Respondent: Occupational Aspiration'
        rea='Respondent: Educational Aspiration'
        fiq='Friend: Intelligence'
        fpa='Friend: Parental Aspiration'
        fses='Friend: Family SES'
        foa='Friend: Occupational Aspiration'
        fea='Friend: Educational Aspiration';
  datalines;
riq  1.      .      .      .      .      .      .      .      .
rpa  .1839  1.      .      .      .      .      .      .      .
rses .2220  .0489  1.      .      .      .      .      .      .
roa  .4105  .2137  .3240  1.      .      .      .      .      .
rea  .4043  .2742  .4047  .6247  1.      .      .      .      .
fiq  .3355  .0782  .2302  .2995  .2863  1.      .      .      .
fpa  .1021  .1147  .0931  .0760  .0702  .2087  1.      .      .
fses .1861  .0186  .2707  .2930  .2407  .2950  -.0438  1.      .

```

```

foa .2598 .0839 .2786 .4216 .3275 .5007 .1988 .3607 1. .
fea .2903 .1124 .3054 .3269 .3669 .5191 .2784 .4105 .6404 1.
;
    
```

The model analyzed by Jöreskog and Sörbom (1988) is displayed in the following path diagram:

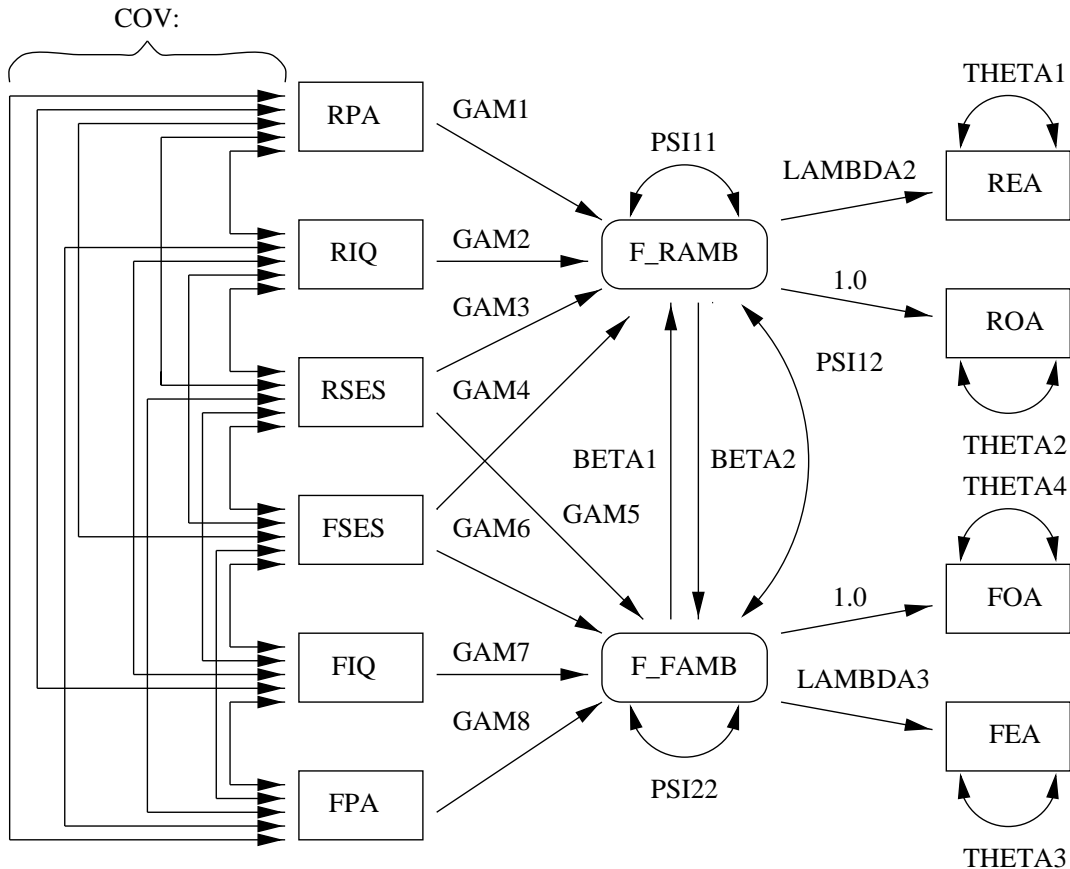


Figure 14.17. Path Diagram: Career Aspiration, Jöreskog and Sörbom

Two latent variables, f_ramb and f_famb , represent the respondent's level of ambition and his best friend's level of ambition, respectively. The model states that the respondent's ambition is determined by his intelligence and socioeconomic status, his perception of his parents' aspiration for him, and his friend's socioeconomic status and ambition. It is assumed that his friend's intelligence and socioeconomic status affect the respondent's ambition only indirectly through his friend's ambition. Ambition is indexed by the manifest variables of occupational and educational aspiration, which are assumed to have uncorrelated residuals. The path coefficient from ambition to occupational aspiration is set to 1.0 to determine the scale of the ambition latent variable.

This model can be analyzed with PROC CALIS using the LINEQS statement as follows, where the names of the parameters correspond to those used by Jöreskog and Sörbom (1988). Since this TYPE=CORR data set does not contain an observation with _TYPE_='N' giving the sample size, it is necessary to specify the degrees of freedom (sample size minus one) with the EDF= option in the PROC CALIS statement.

```

title2 'Joreskog-Sorbom (1988) analysis 1';
proc calis data=aspire edf=328;
  lineqs      /* measurement model for aspiration */
    rea=lambda2 f_ramb + e_rea,
    roa=f_ramb + e_roa,
    fea=lambda3 f_famb + e_fea,
    foa=f_famb + e_foa,
    /* structural model of influences */
    f_ramb=gam1 rpa + gam2 riq + gam3 rses +
      gam4 fses + beta1 f_famb + d_ramb,
    f_famb=gam8 fpa + gam7 fiq + gam6 fses +
      gam5 rses + beta2 f_ramb + d_famb;
  std d_ramb=psi11,
    d_famb=psi22,
    e_rea e_roa e_fea e_foa=theta;;
  cov d_ramb d_famb=psi12,
    rpa riq rses fpa fiq fses=cov;;
run;

```

Specify a name followed by a colon to represent a list of names formed by appending numbers to the specified name. For example, in the COV statement, the line

```
rpa riq rses fpa fiq fses=cov;;
```

is equivalent to

```
rpa riq rses fpa fiq fses=cov1-cov15;
```

The results from this analysis are as follows.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Joreskog-Sorbom (1988) analysis 1 | |
|---|----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0814 |
| Goodness of Fit Index (GFI) | 0.9844 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9428 |
| Root Mean Square Residual (RMR) | 0.0202 |
| Parsimonious GFI (Mulaik, 1989) | 0.3281 |
| Chi-Square | 26.6972 |
| Chi-Square DF | 15 |
| Pr > Chi-Square | 0.0313 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0488 |
| RMSEA 90% Lower Confidence Limit | 0.0145 |
| RMSEA 90% Upper Confidence Limit | 0.0783 |
| ECVI Estimate | 0.2959 |
| ECVI 90% Lower Confidence Limit | 0.2823 |
| ECVI 90% Upper Confidence Limit | 0.3721 |
| Probability of Close Fit | 0.4876 |
| Bentler's Comparative Fit Index | 0.9859 |
| Normal Theory Reweighted LS Chi-Square | 26.0113 |
| Akaike's Information Criterion | -3.3028 |
| Bozdogan's (1987) CAIC | -75.2437 |
| Schwarz's Bayesian Criterion | -60.2437 |
| McDonald's (1989) Centrality | 0.9824 |
| Bentler & Bonett's (1980) Non-normed Index | 0.9576 |
| Bentler & Bonett's (1980) NFI | 0.9694 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.3231 |
| Z-Test of Wilson & Hilferty (1931) | 1.8625 |
| Bollen (1986) Normed Index Rho1 | 0.9082 |
| Bollen (1988) Non-normed Index Delta2 | 0.9864 |
| Hoelter's (1983) Critical N | 309 |

Figure 14.18. Career Aspiration Data: J&S Analysis 1

Jöreskog and Sörbom (1988) present more detailed results from a second analysis in which two constraints are imposed:

- The coefficients connecting the latent ambition variables are equal.
- The covariance of the disturbances of the ambition variables is zero.

This analysis can be performed by changing the names `beta1` and `beta2` to `beta` and omitting the line from the COV statement for `psi12`:

```

title2 'Joreskog-Sorbom (1988) analysis 2';
proc calis data=aspire edf=328;
  lineqs /* measurement model for aspiration */
    rea=lambda2 f_ramb + e_rea,
    roa=f_ramb + e_roa,
    fea=lambda3 f_famb + e_fea,
    foa=f_famb + e_foa,
    /* structural model of influences */
    f_ramb=gam1 rpa + gam2 riq + gam3 rses +
    gam4 fses + beta f_famb + d_ramb,

```

```

          f_famb=gam8 fpa + gam7 fiq + gam6 fses +
          gam5 rses + beta f_ramb + d_famb;
    std d_ramb=psill1,
        d_famb=psi22,
        e_rea e_roa e_fea e_foa=theta;;
    cov rpa riq rses fpa fiq fses=cov;;
run;

```

The results are displayed in Figure 14.19.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Joreskog-Sorbom (1988) analysis 2 | |
|---|----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0820 |
| Goodness of Fit Index (GFI) | 0.9843 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9492 |
| Root Mean Square Residual (RMR) | 0.0203 |
| Parsimonious GFI (Mulaik, 1989) | 0.3718 |
| Chi-Square | 26.8987 |
| Chi-Square DF | 17 |
| Pr > Chi-Square | 0.0596 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0421 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0710 |
| ECVI Estimate | 0.2839 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.3592 |
| Probability of Close Fit | 0.6367 |
| Bentler's Comparative Fit Index | 0.9880 |
| Normal Theory Reweighted LS Chi-Square | 26.1595 |
| Akaike's Information Criterion | -7.1013 |
| Bozdogan's (1987) CAIC | -88.6343 |
| Schwarz's Bayesian Criterion | -71.6343 |
| McDonald's (1989) Centrality | 0.9851 |
| Bentler & Bonett's (1980) Non-normed Index | 0.9683 |
| Bentler & Bonett's (1980) NFI | 0.9692 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.3661 |
| Z-Test of Wilson & Hilferty (1931) | 1.5599 |
| Bollen (1986) Normed Index Rho1 | 0.9183 |
| Bollen (1988) Non-normed Index Delta2 | 0.9884 |
| Hoelter's (1983) Critical N | 338 |

Figure 14.19. Career Aspiration Data: J&S Analysis 2

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Joreskog-Sorbom (1988) analysis 2

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

```

roa      = 1.0000 f_ramb  + 1.0000 e_roa
rea      = 1.0610*f_ramb + 1.0000 e_rea
Std Err  0.0892 lambda2
t Value  11.8923
foa      = 1.0000 f_famb  + 1.0000 e_foa
fea      = 1.0736*f_famb + 1.0000 e_fea
Std Err  0.0806 lambda3
t Value  13.3150
    
```

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Joreskog-Sorbom (1988) analysis 2

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

```

f_ramb = 0.1801*f_famb  + 0.2540*riq      + 0.1637*rpa
Std Err 0.0391 beta      0.0419 gam2      0.0387 gam1
t Value 4.6031          6.0673          4.2274

      + 0.2211*rses      + 0.0773*fses      + 1.0000 d_ramb
      0.0419 gam3        0.0415 gam4
      5.2822            1.8626

f_famb = 0.1801*f_ramb  + 0.0684*rses      + 0.3306*fiq
Std Err 0.0391 beta      0.0387 gam5      0.0412 gam7
t Value 4.6031          1.7681          8.0331

      + 0.1520*fpa      + 0.2184*fses      + 1.0000 d_famb
      0.0364 gam8        0.0395 gam6
      4.1817            5.5320
    
```

```

Peer Influences on Aspiration: Haller & Butterworth (1960)
Joreskog-Sorbom (1988) analysis 2

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Variances of Exogenous Variables

Variable Parameter      Estimate      Standard
                        Error        t Value

riq                    1.00000
rpa                    1.00000
rses                   1.00000
fiq                    1.00000
fpa                    1.00000
fses                   1.00000
e_rea  theta1          0.33764      0.05178      6.52
e_roa  theta2          0.41205      0.05103      8.07
e_fea  theta3          0.31337      0.04574      6.85
e_foa  theta4          0.40381      0.04608      8.76
d_ramb psi11           0.28113      0.04640      6.06
d_famb psi22           0.22924      0.03889      5.89

Covariances Among Exogenous Variables

Var1  Var2  Parameter      Estimate      Standard
                        Error        t Value

riq  rpa  cov1           0.18390      0.05246      3.51
riq  rses cov3           0.22200      0.05110      4.34
rpa  rses cov2           0.04890      0.05493      0.89
riq  fiq  cov8           0.33550      0.04641      7.23
rpa  fiq  cov7           0.07820      0.05455      1.43
rses fiq  cov9           0.23020      0.05074      4.54
riq  fpa  cov5           0.10210      0.05415      1.89
rpa  fpa  cov4           0.11470      0.05412      2.12
rses fpa  cov6           0.09310      0.05438      1.71
fiq  fpa  cov10          0.20870      0.05163      4.04
riq  fses cov12          0.18610      0.05209      3.57
rpa  fses cov11          0.01860      0.05510      0.34
rses fses cov13          0.27070      0.04930      5.49
fiq  fses cov15          0.29500      0.04824      6.12
fpa  fses cov14          -0.04380     0.05476     -0.80

```

The difference between the chi-square values for the two preceding models is $26.8987 - 26.6972 = 0.2015$ with 2 degrees of freedom, which is far from significant. However, the chi-square test of the restricted model (analysis 2) against the alternative of a completely unrestricted covariance matrix yields a p -value of 0.0596, which indicates that the model may not be entirely satisfactory (p -values from these data are probably too small because of the dependence of the observations).

Loehlin (1987) points out that the models considered are unrealistic in at least two aspects. First, the variables of parental aspiration, intelligence, and socioeconomic status are assumed to be measured without error. Loehlin adds uncorrelated measurement errors to the model and assumes, for illustrative purposes, that the reliabilities of these variables are known to be 0.7, 0.8, and 0.9, respectively. In practice, these reliabilities would need to be obtained from a separate study of the same or a very similar population. If these constraints are omitted, the model is not identified. However,

constraining parameters to a constant in an analysis of a correlation matrix may make the chi-square goodness-of-fit test inaccurate, so there is more reason to be skeptical of the *p*-values. Second, the error terms for the respondent's aspiration are assumed to be uncorrelated with the corresponding terms for his friend. Loehlin introduces a correlation between the two educational aspiration error terms and between the two occupational aspiration error terms. These additions produce the following path diagram for Loehlin's model 1.

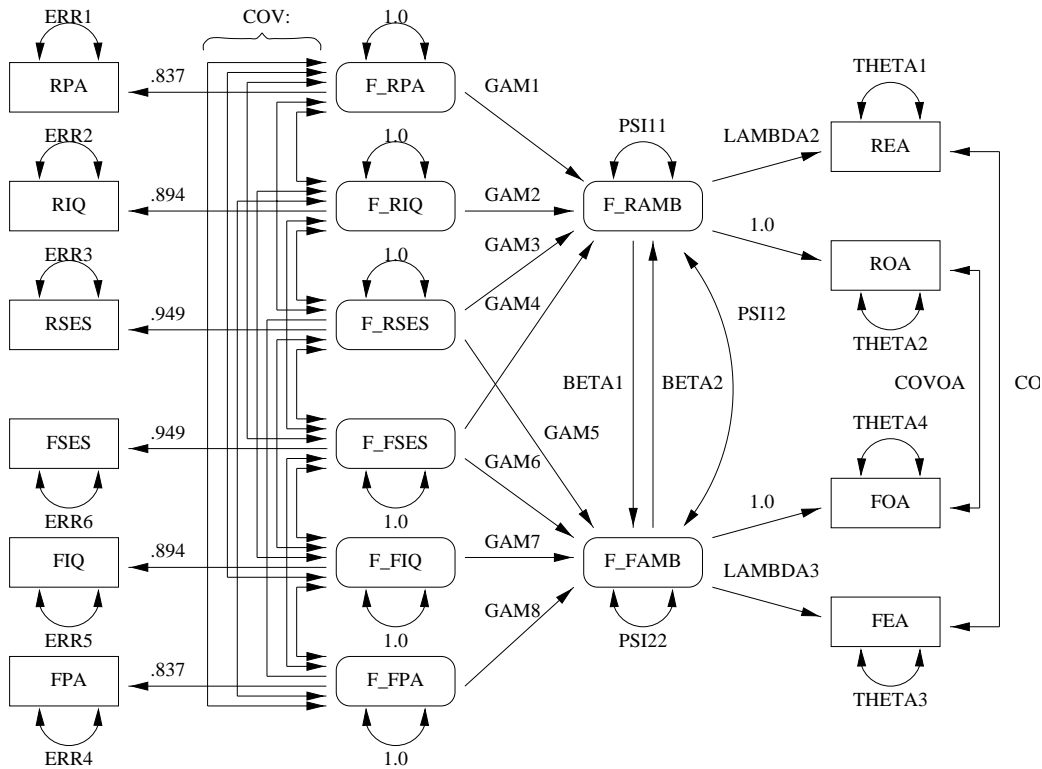


Figure 14.20. Path Diagram: Career Aspiration – Loehlin

The statements for fitting this model are as follows:

```

title2 'Loehlin (1987) analysis: Model 1';
proc calis data=aspire edf=328;
  lineqs      /* measurement model for aspiration */
    rea=lamba2 f_ramb + e_rea,
    roa=f_ramb + e_roa,
    fea=lamba3 f_famb + e_fea,
    foa=f_famb + e_foa,
  /* measurement model for intelligence and environment */
    rpa=.837 f_rpa + e_rpa,
    riq=.894 f_riq + e_riq,
    rses=.949 f_rses + e_rses,
    fpa=.837 f_fpa + e_fpa,
    fiq=.894 f_fiq + e_fiq,
    fses=.949 f_fses + e_fses,

```

```

/* structural model of influences */
f_ramb=gam1 f_rpa + gam2 f_riq + gam3 f_rses +
gam4 f_fses + bet1 f_famb + d_ramb,
f_famb=gam8 f_fpa + gam7 f_fiq + gam6 f_fses +
gam5 f_rses + bet2 f_ramb + d_famb;
std d_ramb=ps11,
d_famb=psi22,
f_rpa f_riq f_rses f_fpa f_fiq f_fses=1,
e_rea e_roa e_fea e_foa=theta:,
e_rpa e_riq e_rses e_fpa e_fiq e_fses=err:;
cov d_ramb d_famb=ps12,
e_rea e_fea=covea,
e_roa e_foa=covoa,
f_rpa f_riq f_rses f_fpa f_fiq f_fses=cov:;
run;

```

The results are displayed in Figure 14.21.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Loehlin (1987) analysis: Model 1 | |
|--|----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0366 |
| Goodness of Fit Index (GFI) | 0.9927 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9692 |
| Root Mean Square Residual (RMR) | 0.0149 |
| Parsimonious GFI (Mulaik, 1989) | 0.2868 |
| Chi-Square | 12.0132 |
| Chi-Square DF | 13 |
| Pr > Chi-Square | 0.5266 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0512 |
| ECVI Estimate | 0.3016 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.3392 |
| Probability of Close Fit | 0.9435 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 12.0168 |
| Akaike's Information Criterion | -13.9868 |
| Bozdogan's (1987) CAIC | -76.3356 |
| Schwarz's Bayesian Criterion | -63.3356 |
| McDonald's (1989) Centrality | 1.0015 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0041 |
| Bentler & Bonett's (1980) NFI | 0.9862 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.2849 |
| Z-Test of Wilson & Hilferty (1931) | -0.0679 |
| Bollen (1986) Normed Index Rho1 | 0.9523 |
| Bollen (1988) Non-normed Index Delta2 | 1.0011 |
| Hoelter's (1983) Critical N | 612 |

Figure 14.21. Career Aspiration Data: Loehlin Model 1

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 1

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

```

riq    = 0.8940 f_riq    + 1.0000 e_riq
rpa    = 0.8370 f_rpa    + 1.0000 e_rpa
rses   = 0.9490 f_rses   + 1.0000 e_rses
roa    = 1.0000 f_ramb   + 1.0000 e_roa
rea    = 1.0840*f_ramb   + 1.0000 e_rea
Std Err 0.0942 lambda2
t Value 11.5105

fiq    = 0.8940 f_fiq    + 1.0000 e_fiq
fpa    = 0.8370 f_fpa    + 1.0000 e_fpa
fses   = 0.9490 f_fses   + 1.0000 e_fses
foa    = 1.0000 f_famb   + 1.0000 e_foa
fea    = 1.1163*f_famb   + 1.0000 e_fea
Std Err 0.0863 lambda3
t Value 12.9394
    
```

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 1

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

```

f_ramb = 0.1190*f_famb + 0.1837*f_rpa + 0.2800*f_riq
Std Err 0.1140 bet1      0.0504 gam1      0.0614 gam2
t Value 1.0440          3.6420          4.5618

+ 0.2262*f_rses + 0.0870*f_fses + 1.0000 d_ramb
0.0522 gam3      0.0548 gam4
4.3300          1.5884

f_famb = 0.1302*f_ramb + 0.0633*f_rses + 0.1688*f_fpa
Std Err 0.1207 bet2      0.0522 gam5      0.0493 gam8
t Value 1.0792          1.2124          3.4205

+ 0.3539*f_fiq + 0.2154*f_fses + 1.0000 d_famb
0.0674 gam7      0.0512 gam6
5.2497          4.2060
    
```

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 1

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

VariANCES of Exogenous Variables

| Variable | Parameter | Estimate | Standard Error | t Value |
|----------|-----------|----------|----------------|---------|
| f_rpa | | 1.00000 | | |
| f_riq | | 1.00000 | | |
| f_rses | | 1.00000 | | |
| f_fpa | | 1.00000 | | |
| f_fiq | | 1.00000 | | |
| f_fses | | 1.00000 | | |
| e_rea | theta1 | 0.32707 | 0.05452 | 6.00 |
| e_roa | theta2 | 0.42307 | 0.05243 | 8.07 |
| e_fea | theta3 | 0.28715 | 0.04804 | 5.98 |
| e_foa | theta4 | 0.42240 | 0.04730 | 8.93 |
| e_rpa | err1 | 0.29584 | 0.07774 | 3.81 |
| e_riq | err2 | 0.20874 | 0.07832 | 2.67 |
| e_rses | err3 | 0.09887 | 0.07803 | 1.27 |
| e_fpa | err4 | 0.29987 | 0.07807 | 3.84 |
| e_fiq | err5 | 0.19988 | 0.07674 | 2.60 |
| e_fses | err6 | 0.10324 | 0.07824 | 1.32 |
| d_ramb | psi11 | 0.25418 | 0.04469 | 5.69 |
| d_famb | psi22 | 0.19698 | 0.03814 | 5.17 |

Covariances Among Exogenous Variables

| Var1 | Var2 | Parameter | Estimate | Standard Error | t Value |
|--------|--------|-----------|----------|----------------|---------|
| f_rpa | f_riq | cov1 | 0.24677 | 0.07519 | 3.28 |
| f_rpa | f_rses | cov2 | 0.06184 | 0.06945 | 0.89 |
| f_riq | f_rses | cov3 | 0.26351 | 0.06687 | 3.94 |
| f_rpa | f_fpa | cov4 | 0.15789 | 0.07873 | 2.01 |
| f_riq | f_fpa | cov5 | 0.13085 | 0.07418 | 1.76 |
| f_rses | f_fpa | cov6 | 0.11517 | 0.06978 | 1.65 |
| f_rpa | f_fiq | cov7 | 0.10853 | 0.07362 | 1.47 |
| f_riq | f_fiq | cov8 | 0.42476 | 0.07219 | 5.88 |
| f_rses | f_fiq | cov9 | 0.27250 | 0.06660 | 4.09 |
| f_fpa | f_fiq | cov10 | 0.27867 | 0.07530 | 3.70 |
| f_rpa | f_fses | cov11 | 0.02383 | 0.06952 | 0.34 |
| f_riq | f_fses | cov12 | 0.22135 | 0.06648 | 3.33 |
| f_rses | f_fses | cov13 | 0.30156 | 0.06359 | 4.74 |
| f_fpa | f_fses | cov14 | -0.05623 | 0.06971 | -0.81 |
| f_fiq | f_fses | cov15 | 0.34922 | 0.06771 | 5.16 |
| e_rea | e_fea | covea | 0.02308 | 0.03139 | 0.74 |
| e_roa | e_foa | covoa | 0.11206 | 0.03258 | 3.44 |
| d_ramb | d_famb | psi12 | -0.00935 | 0.05010 | -0.19 |

Since the p -value for the chi-square test is 0.5266, this model clearly cannot be rejected. However, Schwarz's Bayesian Criterion for this model (SBC = -63.3356) is somewhat larger than for Jöreskog and Sörbom's (1988) analysis 2 (SBC = -71.6343), suggesting that a more parsimonious model would be desirable.

Since it is assumed that the same model applies to all the boys in the sample, the path diagram should be symmetric with respect to the respondent and friend. In particular, the corresponding coefficients should be equal. By imposing equality constraints

on the 15 pairs of corresponding coefficients, this example obtains Loehlin's model 2. The LINEQS model is as follows, where an OUTRAM= data set is created to facilitate subsequent hypothesis tests:

```

title2 'Loehlin (1987) analysis: Model 2';
proc calis data=aspire edf=328 outram=ram2;
  lineqs      /* measurement model for aspiration */
    rea=lambda f_reamb + e_rea,          /* 1 ec! */
    roa=f_reamb + e_roa,
    fea=lambda f_famb + e_fea,
    foa=f_famb + e_foa,
    /* measurement model for intelligence and environment */
    rpa=.837 f_rpa + e_rpa,
    riq=.894 f_riq + e_riq,
    rses=.949 f_rses + e_rses,
    fpa=.837 f_fpa + e_fpa,
    fiq=.894 f_fiq + e_fiq,
    fses=.949 f_fses + e_fses,
    /* structural model of influences */      /* 5 ec! */
    f_reamb=gam1 f_rpa + gam2 f_riq + gam3 f_rses +
      gam4 f_fses + beta f_famb + d_reamb,
    f_famb=gam1 f_fpa + gam2 f_fiq + gam3 f_fses +
      gam4 f_rses + beta f_reamb + d_famb;
  std d_reamb=psi,          /* 1 ec! */
    d_famb=psi,
    f_rpa f_riq f_rses f_fpa f_fiq f_fses=1,
    e_rea e_fea=thetaea thetaea,          /* 2 ec! */
    e_roa e_foa=thetaoa thetaoa,
    e_rpa e_fpa=errpa1 errpa2,
    e_riq e_fiq=erriq1 erriq2,
    e_rses e_fses=errses1 errses2;
  cov d_reamb d_famb=psi12,
    e_rea e_fea=covea,
    e_roa e_foa = covoa,
    f_rpa f_riq f_rses=cov1-cov3,          /* 3 ec! */
    f_fpa f_fiq f_fses=cov1-cov3,
    f_rpa f_riq f_rses * f_fpa f_fiq f_fses = /* 3 ec! */
      cov4 cov5 cov6
      cov5 cov7 cov8
      cov6 cov8 cov9;
run;

```

The results are displayed in Figure 14.22.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Loehlin (1987) analysis: Model 2 | |
|--|-----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0581 |
| Goodness of Fit Index (GFI) | 0.9884 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9772 |
| Root Mean Square Residual (RMR) | 0.0276 |
| Parsimonious GFI (Mulaik, 1989) | 0.6150 |
| Chi-Square | 19.0697 |
| Chi-Square DF | 28 |
| Pr > Chi-Square | 0.8960 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0194 |
| ECVI Estimate | 0.2285 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.2664 |
| Probability of Close Fit | 0.9996 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 19.2372 |
| Akaike's Information Criterion | -36.9303 |
| Bozdogan's (1987) CAIC | -171.2200 |
| Schwarz's Bayesian Criterion | -143.2200 |
| McDonald's (1989) Centrality | 1.0137 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0174 |
| Bentler & Bonett's (1980) NFI | 0.9781 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.6086 |
| Z-Test of Wilson & Hilferty (1931) | -1.2599 |
| Bollen (1986) Normed Index Rho1 | 0.9649 |
| Bollen (1988) Non-normed Index Delta2 | 1.0106 |
| Hoelter's (1983) Critical N | 713 |

Figure 14.22. Career Aspiration Data: Loehlin Model 2

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 2

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

| | | | | | | |
|---------|---|---------|--------|---|--------|--------|
| riq | = | 0.8940 | f_riq | + | 1.0000 | e_riq |
| rpa | = | 0.8370 | f_rpa | + | 1.0000 | e_rpa |
| rses | = | 0.9490 | f_rses | + | 1.0000 | e_rses |
| roa | = | 1.0000 | f_ramb | + | 1.0000 | e_roa |
| rea | = | 1.1007* | f_ramb | + | 1.0000 | e_rea |
| Std Err | | 0.0684 | lambda | | | |
| t Value | | 16.0879 | | | | |
| fiq | = | 0.8940 | f_fiq | + | 1.0000 | e_fiq |
| fpa | = | 0.8370 | f_fpa | + | 1.0000 | e_fpa |
| fses | = | 0.9490 | f_fses | + | 1.0000 | e_fses |
| foa | = | 1.0000 | f_famb | + | 1.0000 | e_foa |
| fea | = | 1.1007* | f_famb | + | 1.0000 | e_fea |
| Std Err | | 0.0684 | lambda | | | |
| t Value | | 16.0879 | | | | |

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 2

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

| | | | | | | |
|---------|---|---------------|---|---------------|---|---------------|
| f_ramb | = | 0.1158*f_famb | + | 0.1758*f_rpa | + | 0.3223*f_riq |
| Std Err | | 0.0839 beta | | 0.0351 gam1 | | 0.0470 gam2 |
| t Value | | 1.3801 | | 5.0130 | | 6.8557 |
| | + | 0.2227*f_rses | + | 0.0756*f_fses | + | 1.0000 d_ramb |
| | | 0.0363 gam3 | | 0.0375 gam4 | | |
| | | 6.1373 | | 2.0170 | | |
| f_famb | = | 0.1158*f_ramb | + | 0.0756*f_rses | + | 0.1758*f_fpa |
| Std Err | | 0.0839 beta | | 0.0375 gam4 | | 0.0351 gam1 |
| t Value | | 1.3801 | | 2.0170 | | 5.0130 |
| | + | 0.3223*f_fiq | + | 0.2227*f_fses | + | 1.0000 d_famb |
| | | 0.0470 gam2 | | 0.0363 gam3 | | |
| | | 6.8557 | | 6.1373 | | |

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 2

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Variances of Exogenous Variables

| Variable | Parameter | Estimate | Standard Error | t Value |
|----------|-----------|----------|----------------|---------|
| f_rpa | | 1.00000 | | |
| f_riq | | 1.00000 | | |
| f_rses | | 1.00000 | | |
| f_fpa | | 1.00000 | | |
| f_fiq | | 1.00000 | | |
| f_fses | | 1.00000 | | |
| e_rea | thetaea | 0.30662 | 0.03726 | 8.23 |
| e_roa | thetaoa | 0.42295 | 0.03651 | 11.58 |
| e_fea | thetaea | 0.30662 | 0.03726 | 8.23 |
| e_foa | thetaoa | 0.42295 | 0.03651 | 11.58 |
| e_rpa | errpa1 | 0.30758 | 0.07511 | 4.09 |
| e_riq | errriq1 | 0.26656 | 0.07389 | 3.61 |
| e_rses | errses1 | 0.11467 | 0.07267 | 1.58 |
| e_fpa | errpa2 | 0.28834 | 0.07369 | 3.91 |
| e_fiq | errriq2 | 0.15573 | 0.06700 | 2.32 |
| e_fses | errses2 | 0.08814 | 0.07089 | 1.24 |
| d_ramb | psi | 0.22456 | 0.02971 | 7.56 |
| d_famb | psi | 0.22456 | 0.02971 | 7.56 |

Covariances Among Exogenous Variables

| Var1 | Var2 | Parameter | Estimate | Standard Error | t Value |
|--------|--------|-----------|----------|----------------|---------|
| f_rpa | f_riq | cov1 | 0.26470 | 0.05442 | 4.86 |
| f_rpa | f_rses | cov2 | 0.00176 | 0.04996 | 0.04 |
| f_riq | f_rses | cov3 | 0.31129 | 0.05057 | 6.16 |
| f_rpa | f_fpa | cov4 | 0.15784 | 0.07872 | 2.01 |
| f_riq | f_fpa | cov5 | 0.11837 | 0.05447 | 2.17 |
| f_rses | f_fpa | cov6 | 0.06910 | 0.04996 | 1.38 |
| f_rpa | f_fiq | cov5 | 0.11837 | 0.05447 | 2.17 |
| f_riq | f_fiq | cov7 | 0.43061 | 0.07258 | 5.93 |
| f_rses | f_fiq | cov8 | 0.24967 | 0.05060 | 4.93 |
| f_fpa | f_fiq | cov1 | 0.26470 | 0.05442 | 4.86 |
| f_rpa | f_fses | cov6 | 0.06910 | 0.04996 | 1.38 |
| f_riq | f_fses | cov8 | 0.24967 | 0.05060 | 4.93 |
| f_rses | f_fses | cov9 | 0.30190 | 0.06362 | 4.75 |
| f_fpa | f_fses | cov2 | 0.00176 | 0.04996 | 0.04 |
| f_fiq | f_fses | cov3 | 0.31129 | 0.05057 | 6.16 |
| e_rea | e_fea | covea | 0.02160 | 0.03144 | 0.69 |
| e_roa | e_foa | covoa | 0.11208 | 0.03257 | 3.44 |
| d_ramb | d_famb | psi12 | -0.00344 | 0.04931 | -0.07 |

The test of Loehlin's model 2 against model 1 yields a chi-square of $19.0697 - 12.0132 = 7.0565$ with 15 degrees of freedom, which is clearly not significant. Schwarz's Bayesian Criterion (SBC) is also much lower for model 2 (-143.2200) than model 1 (-63.3356). Hence, model 2 seems preferable on both substantive and statistical grounds.

A question of substantive interest is whether the friend's socioeconomic status (SES) has a significant direct influence on a boy's ambition. This can be addressed by

omitting the paths from `f_fses` to `f_ramb` and from `f_rses` to `f_famb` designated by the parameter name `gam4`, yielding Loehlin's model 3:

```

title2 'Loehlin (1987) analysis: Model 3';
data ram3(type=ram);
  set ram2;
  if _name_='gam4' then
  do;
    _name_=' ';
    _estim_=0;
  end;
run;
proc calis data=aspire edf=328 inram=ram3;
run;
    
```

The output is displayed in Figure 14.23.

| Peer Influences on Aspiration: Haller & Butterworth (1960) | |
|--|-----------|
| Loehlin (1987) analysis: Model 3 | |
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0702 |
| Goodness of Fit Index (GFI) | 0.9858 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9731 |
| Root Mean Square Residual (RMR) | 0.0304 |
| Parsimonious GFI (Mulaik, 1989) | 0.6353 |
| Chi-Square | 23.0365 |
| Chi-Square DF | 29 |
| Pr > Chi-Square | 0.7749 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0295 |
| ECVI Estimate | 0.2343 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.2780 |
| Probability of Close Fit | 0.9984 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 23.5027 |
| Akaike's Information Criterion | -34.9635 |
| Bozdogan's (1987) CAIC | -174.0492 |
| Schwarz's Bayesian Criterion | -145.0492 |
| McDonald's (1989) Centrality | 1.0091 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0112 |
| Bentler & Bonett's (1980) NFI | 0.9736 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.6274 |
| Z-Test of Wilson & Hilferty (1931) | -0.7563 |
| Bollen (1986) Normed Index Rho1 | 0.9590 |
| Bollen (1988) Non-normed Index Delta2 | 1.0071 |
| Hoelter's (1983) Critical N | 607 |

Figure 14.23. Career Aspiration Data: Loehlin Model 3

The chi-square value for testing model 3 versus model 2 is $23.0365 - 19.0697 = 3.9668$ with 1 degree of freedom and a p -value of 0.0464. Although the parameter is of marginal significance, the estimate in model 2 (0.0756) is small compared to the other coefficients, and SBC indicates that model 3 is preferable to model 2.

Another important question is whether the reciprocal influences between the respondent's and friend's ambitions are needed in the model. To test whether these paths are zero, set the parameter `beta` for the paths linking `f_ramb` and `f_famb` to zero to obtain Loehlin's model 4:

```

title2 'Loehlin (1987) analysis: Model 4';
data ram4(type=ram);
  set ram2;
  if _name_='beta' then
    do;
      _name_=' ';
      _estim_=0;
    end;
run;

proc calis data=aspire edf=328 inram=ram4;
run;

```

The output is displayed in Figure 14.24.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Loehlin (1987) analysis: Model 4 | |
|--|-----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0640 |
| Goodness of Fit Index (GFI) | 0.9873 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9760 |
| Root Mean Square Residual (RMR) | 0.0304 |
| Parsimonious GFI (Mulaik, 1989) | 0.6363 |
| Chi-Square | 20.9981 |
| Chi-Square DF | 29 |
| Pr > Chi-Square | 0.8592 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0234 |
| ECVI Estimate | 0.2281 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.2685 |
| Probability of Close Fit | 0.9994 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 20.8040 |
| Akaike's Information Criterion | -37.0019 |
| Bozdogan's (1987) CAIC | -176.0876 |
| Schwarz's Bayesian Criterion | -147.0876 |
| McDonald's (1989) Centrality | 1.0122 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0150 |
| Bentler & Bonett's (1980) NFI | 0.9759 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.6289 |
| Z-Test of Wilson & Hilferty (1931) | -1.0780 |
| Bollen (1986) Normed Index Rho1 | 0.9626 |
| Bollen (1988) Non-normed Index Delta2 | 1.0095 |
| Hoelter's (1983) Critical N | 666 |

Figure 14.24. Career Aspiration Data: Loehlin Model 4

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 4

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

| | | | | | | |
|---------|---|---------|--------|---|--------|--------|
| riq | = | 0.8940 | f_riq | + | 1.0000 | e_riq |
| rpa | = | 0.8370 | f_rpa | + | 1.0000 | e_rpa |
| rses | = | 0.9490 | f_rses | + | 1.0000 | e_rses |
| roa | = | 1.0000 | f_ramb | + | 1.0000 | e_roa |
| rea | = | 1.1051* | f_ramb | + | 1.0000 | e_rea |
| Std Err | | 0.0680 | lambda | | | |
| t Value | | 16.2416 | | | | |
| fiq | = | 0.8940 | f_fiq | + | 1.0000 | e_fiq |
| fpa | = | 0.8370 | f_fpa | + | 1.0000 | e_fpa |
| fses | = | 0.9490 | f_fses | + | 1.0000 | e_fses |
| foa | = | 1.0000 | f_famb | + | 1.0000 | e_foa |
| fea | = | 1.1051* | f_famb | + | 1.0000 | e_fea |
| Std Err | | 0.0680 | lambda | | | |
| t Value | | 16.2416 | | | | |

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 4

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

| | | | | | | | | | |
|---------|---|---------|--------|---|---------|--------|---|---------|--------|
| f_ramb | = | 0 | f_famb | + | 0.1776* | f_rpa | + | 0.3486* | f_riq |
| Std Err | | | | | 0.0361 | gam1 | | 0.0463 | gam2 |
| t Value | | | | | 4.9195 | | | 7.5362 | |
| | + | 0.2383* | f_rses | + | 0.1081* | f_fses | + | 1.0000 | d_ramb |
| | | 0.0355 | gam3 | | 0.0299 | gam4 | | | |
| | | 6.7158 | | | 3.6134 | | | | |
| f_famb | = | 0 | f_ramb | + | 0.1081* | f_rses | + | 0.1776* | f_fpa |
| Std Err | | | | | 0.0299 | gam4 | | 0.0361 | gam1 |
| t Value | | | | | 3.6134 | | | 4.9195 | |
| | + | 0.3486* | f_fiq | + | 0.2383* | f_fses | + | 1.0000 | d_famb |
| | | 0.0463 | gam2 | | 0.0355 | gam3 | | | |
| | | 7.5362 | | | 6.7158 | | | | |

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 4

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

VariANCES of Exogenous Variables

| Variable | Parameter | Estimate | Standard Error | t Value |
|----------|-----------|----------|----------------|---------|
| f_rpa | | 1.00000 | | |
| f_riq | | 1.00000 | | |
| f_rses | | 1.00000 | | |
| f_fpa | | 1.00000 | | |
| f_fiq | | 1.00000 | | |
| f_fses | | 1.00000 | | |
| e_rea | thetaea | 0.30502 | 0.03728 | 8.18 |
| e_roa | thetaoa | 0.42429 | 0.03645 | 11.64 |
| e_fea | thetaea | 0.30502 | 0.03728 | 8.18 |
| e_foa | thetaoa | 0.42429 | 0.03645 | 11.64 |
| e_rpa | errpa1 | 0.31354 | 0.07543 | 4.16 |
| e_riq | errriq1 | 0.29611 | 0.07299 | 4.06 |
| e_rses | errses1 | 0.12320 | 0.07273 | 1.69 |
| e_fpa | errpa2 | 0.29051 | 0.07374 | 3.94 |
| e_fiq | errriq2 | 0.18181 | 0.06611 | 2.75 |
| e_fses | errses2 | 0.09873 | 0.07109 | 1.39 |
| d_ramb | psi | 0.22738 | 0.03140 | 7.24 |
| d_famb | psi | 0.22738 | 0.03140 | 7.24 |

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 4

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Covariances Among Exogenous Variables

| Var1 | Var2 | Parameter | Estimate | Standard Error | t Value |
|--------|--------|-----------|----------|----------------|---------|
| f_rpa | f_riq | cov1 | 0.27241 | 0.05520 | 4.94 |
| f_rpa | f_rses | cov2 | 0.00476 | 0.05032 | 0.09 |
| f_riq | f_rses | cov3 | 0.32463 | 0.05089 | 6.38 |
| f_rpa | f_fpa | cov4 | 0.16949 | 0.07863 | 2.16 |
| f_riq | f_fpa | cov5 | 0.13539 | 0.05407 | 2.50 |
| f_rses | f_fpa | cov6 | 0.07362 | 0.05027 | 1.46 |
| f_rpa | f_fiq | cov5 | 0.13539 | 0.05407 | 2.50 |
| f_riq | f_fiq | cov7 | 0.46893 | 0.06980 | 6.72 |
| f_rses | f_fiq | cov8 | 0.26289 | 0.05093 | 5.16 |
| f_fpa | f_fiq | cov1 | 0.27241 | 0.05520 | 4.94 |
| f_rpa | f_fses | cov6 | 0.07362 | 0.05027 | 1.46 |
| f_riq | f_fses | cov8 | 0.26289 | 0.05093 | 5.16 |
| f_rses | f_fses | cov9 | 0.30880 | 0.06409 | 4.82 |
| f_fpa | f_fses | cov2 | 0.00476 | 0.05032 | 0.09 |
| f_fiq | f_fses | cov3 | 0.32463 | 0.05089 | 6.38 |
| e_rea | e_fea | covea | 0.02127 | 0.03150 | 0.68 |
| e_roa | e_foa | covoa | 0.11245 | 0.03258 | 3.45 |
| d_ramb | d_famb | psi12 | 0.05479 | 0.02699 | 2.03 |

The chi-square value for testing model 4 versus model 2 is $20.9981 - 19.0697 = 1.9284$ with 1 degree of freedom and a p -value of 0.1649. Hence, there is little evidence of reciprocal influence.

Loehlin's model 2 has not only the direct paths connecting the latent ambition variables `f_amb` and `f_famb` but also a covariance between the disturbance terms `d_amb` and `d_famb` to allow for other variables omitted from the model that might jointly influence the respondent and his friend. To test the hypothesis that this covariance is zero, set the parameter `psi12` to zero, yielding Loehlin's model 5:

```
title2 'Loehlin (1987) analysis: Model 5';  
data ram5(type=ram);  
  set ram2;  
  if _name_='psi12' then  
    do;  
      _name_='';  
      _estim_=0;  
    end;  
run;  
  
proc calis data=aspire edf=328 inram=ram5;  
run;
```

The output is displayed in Figure 14.25.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Loehlin (1987) analysis: Model 5 | |
|--|-----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.0582 |
| Goodness of Fit Index (GFI) | 0.9884 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9780 |
| Root Mean Square Residual (RMR) | 0.0276 |
| Parsimonious GFI (Mulaik, 1989) | 0.6370 |
| Chi-Square | 19.0745 |
| Chi-Square DF | 29 |
| Pr > Chi-Square | 0.9194 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0152 |
| ECVI Estimate | 0.2222 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.2592 |
| Probability of Close Fit | 0.9998 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 19.2269 |
| Akaike's Information Criterion | -38.9255 |
| Bozdogan's (1987) CAIC | -178.0111 |
| Schwarz's Bayesian Criterion | -149.0111 |
| McDonald's (1989) Centrality | 1.0152 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0186 |
| Bentler & Bonett's (1980) NFI | 0.9781 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.6303 |
| Z-Test of Wilson & Hilferty (1931) | -1.4014 |
| Bollen (1986) Normed Index Rho1 | 0.9661 |
| Bollen (1988) Non-normed Index Delta2 | 1.0118 |
| Hoelter's (1983) Critical N | 733 |

Figure 14.25. Career Aspiration Data: Loehlin Model 5

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 5

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

```

riq    = 0.8940 f_riq    + 1.0000 e_riq
rpa    = 0.8370 f_rpa    + 1.0000 e_rpa
rses   = 0.9490 f_rses   + 1.0000 e_rses
roa    = 1.0000 f_ramb   + 1.0000 e_roa
rea    = 1.1009*f_ramb   + 1.0000 e_rea
Std Err 0.0684 lambda
t Value 16.1041

fiq    = 0.8940 f_fiq    + 1.0000 e_fiq
fpa    = 0.8370 f_fpa    + 1.0000 e_fpa
fses   = 0.9490 f_fses   + 1.0000 e_fses
foa    = 1.0000 f_famb   + 1.0000 e_foa
fea    = 1.1009*f_famb   + 1.0000 e_fea
Std Err 0.0684 lambda
t Value 16.1041
    
```

Peer Influences on Aspiration: Haller & Butterworth (1960)
 Loehlin (1987) analysis: Model 5

The CALIS Procedure
 Covariance Structure Analysis: Maximum Likelihood Estimation

```

f_ramb = 0.1107*f_famb + 0.1762*f_rpa + 0.3235*f_riq
Std Err 0.0428 beta      0.0350 gam1      0.0435 gam2
t Value 2.5854          5.0308          7.4435

+ 0.2233*f_rses + 0.0770*f_fses + 1.0000 d_ramb
0.0353 gam3      0.0323 gam4
6.3215          2.3870

f_famb = 0.1107*f_ramb + 0.0770*f_rses + 0.1762*f_fpa
Std Err 0.0428 beta      0.0323 gam4      0.0350 gam1
t Value 2.5854          2.3870          5.0308

+ 0.3235*f_fiq + 0.2233*f_fses + 1.0000 d_famb
0.0435 gam2      0.0353 gam3
7.4435          6.3215
    
```

```

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 5

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Variances of Exogenous Variables

Variable Parameter      Estimate      Standard
                        Error          t Value

f_rpa                    1.00000
f_riq                    1.00000
f_rses                   1.00000
f_fpa                    1.00000
f_fiq                    1.00000
f_fses                   1.00000
e_rea  thetaea          0.30645      0.03721      8.24
e_roa  thetaoa          0.42304      0.03650     11.59
e_fea  thetaea          0.30645      0.03721      8.24
e_foa  thetaoa          0.42304      0.03650     11.59
e_rpa  errpa1           0.30781      0.07510      4.10
e_riq  erriq1           0.26748      0.07295      3.67
e_rses errs1            0.11477      0.07265      1.58
e_fpa  errpa2           0.28837      0.07366      3.91
e_fiq  erriq2           0.15653      0.06614      2.37
e_fses errs2            0.08832      0.07088      1.25
d_ramb psi              0.22453      0.02973      7.55
d_famb psi              0.22453      0.02973      7.55

Covariances Among Exogenous Variables

Var1  Var2  Parameter      Estimate      Standard
                        Error          t Value

f_rpa f_riq  cov1           0.26494      0.05436      4.87
f_rpa f_rses cov2           0.00185      0.04995      0.04
f_riq f_rses cov3           0.31164      0.05039      6.18
f_rpa f_fpa  cov4           0.15828      0.07846      2.02
f_riq f_fpa  cov5           0.11895      0.05383      2.21
f_rses f_fpa cov6           0.06924      0.04993      1.39
f_rpa f_fiq  cov5           0.11895      0.05383      2.21
f_riq f_fiq  cov7           0.43180      0.07084      6.10
f_rses f_fiq  cov8           0.25004      0.05039      4.96
f_fpa f_fiq  cov1           0.26494      0.05436      4.87
f_rpa f_fses cov6           0.06924      0.04993      1.39
f_riq f_fses cov8           0.25004      0.05039      4.96
f_rses f_fses cov9           0.30203      0.06360      4.75
f_fpa f_fses cov2           0.00185      0.04995      0.04
f_fiq f_fses cov3           0.31164      0.05039      6.18
e_rea e_fea  covea          0.02120      0.03094      0.69
e_roa e_foa  covoa          0.11197      0.03254      3.44
d_ramb d_famb 0

```

The chi-square value for testing model 5 versus model 2 is $19.0745 - 19.0697 = 0.0048$ with 1 degree of freedom. Omitting the covariance between the disturbance terms, therefore, causes hardly any deterioration in the fit of the model.

These data fail to provide evidence of direct reciprocal influence between the respondent's and friend's ambitions or of a covariance between the disturbance terms when these hypotheses are considered separately. Notice, however, that the covariance ψ_{12} between the disturbance terms increases from -0.003344 for model 2 to

0.05479 for model 4. Before you conclude that all of these paths can be omitted from the model, it is important to test both hypotheses together by setting both beta and psi12 to zero as in Loehlin's model 7:

```

title2 'Loehlin (1987) analysis: Model 7';
data ram7(type=ram);
  set ram2;
  if _name_='psi12'|_name_='beta' then
    do;
      _name_=' ';
      _estim_=0;
    end;
run;

proc calis data=aspire edf=328 inram=ram7;
run;
    
```

The relevant output is displayed in Figure 14.26.

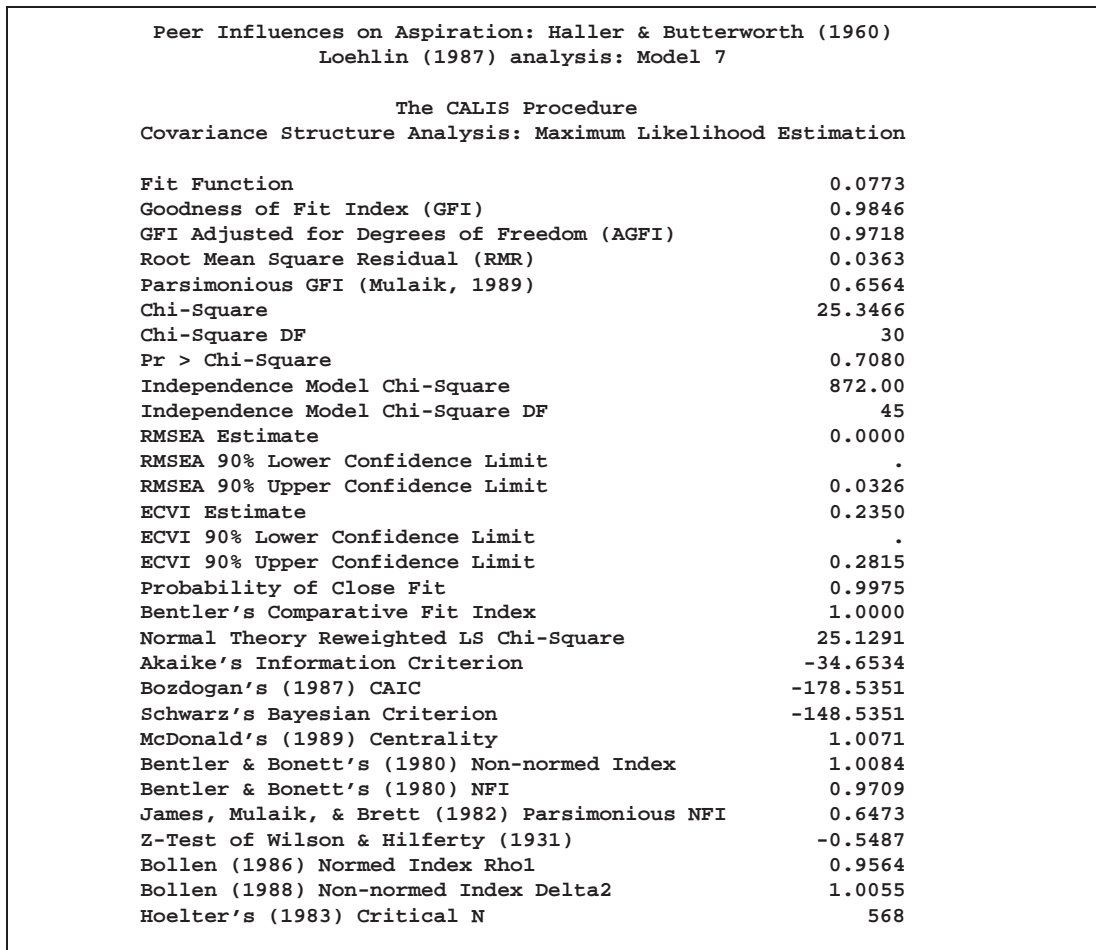


Figure 14.26. Career Aspiration Data: Loehlin Model 7

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 7

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

```

riq    = 0.8940 f_riq  + 1.0000 e_riq
rpa    = 0.8370 f_rpa  + 1.0000 e_rpa
rses   = 0.9490 f_rses + 1.0000 e_rses
roa    = 1.0000 f_ramb + 1.0000 e_roa
rea    = 1.1037*f_ramb + 1.0000 e_rea
Std Err 0.0678 lambda
t Value 16.2701
fiq    = 0.8940 f_fiq  + 1.0000 e_fiq
fpa    = 0.8370 f_fpa  + 1.0000 e_fpa
fses   = 0.9490 f_fses + 1.0000 e_fses
foa    = 1.0000 f_famb + 1.0000 e_foa
fea    = 1.1037*f_famb + 1.0000 e_fea
Std Err 0.0678 lambda
t Value 16.2701

```

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 7

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

```

f_ramb = 0 f_famb + 0.1765*f_rpa + 0.3573*f_riq
Std Err 0.0360 gam1 0.0461 gam2
t Value 4.8981 7.7520

+ 0.2419*f_rses + 0.1109*f_fses + 1.0000 d_ramb
0.0363 gam3 0.0306 gam4
6.6671 3.6280

f_famb = 0 f_ramb + 0.1109*f_rses + 0.1765*f_fpa
Std Err 0.0306 gam4 0.0360 gam1
t Value 3.6280 4.8981

+ 0.3573*f_fiq + 0.2419*f_fses + 1.0000 d_famb
0.0461 gam2 0.0363 gam3
7.7520 6.6671

```

```

Peer Influences on Aspiration: Haller & Butterworth (1960)
Loehlin (1987) analysis: Model 7

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Variances of Exogenous Variables

Variable Parameter      Estimate      Standard
                        Error      t Value

f_rpa                    1.00000
f_riq                    1.00000
f_rses                   1.00000
f_fpa                    1.00000
f_fiq                    1.00000
f_fses                   1.00000
e_rea  thetaea          0.31633      0.03648      8.67
e_roa  thetaoa          0.42656      0.03610     11.82
e_fea  thetaea          0.31633      0.03648      8.67
e_foa  thetaoa          0.42656      0.03610     11.82
e_rpa  errpa1          0.31329      0.07538      4.16
e_riq  erriq1          0.30776      0.07307      4.21
e_rses errs1           0.14303      0.07313      1.96
e_fpa  errpa2          0.29286      0.07389      3.96
e_fiq  erriq2          0.19193      0.06613      2.90
e_fses errs2           0.11804      0.07147      1.65
d_ramb psi              0.21011      0.02940      7.15
d_famb psi              0.21011      0.02940      7.15

Covariances Among Exogenous Variables

Var1  Var2  Parameter      Estimate      Standard
                        Error      t Value

f_rpa  f_riq  cov1           0.27533      0.05552      4.96
f_rpa  f_rses cov2           0.00611      0.05085      0.12
f_riq  f_rses cov3           0.33510      0.05150      6.51
f_rpa  f_fpa  cov4           0.17099      0.07872      2.17
f_riq  f_fpa  cov5           0.13859      0.05431      2.55
f_rses f_fpa  cov6           0.07563      0.05077      1.49
f_rpa  f_fiq  cov5           0.13859      0.05431      2.55
f_riq  f_fiq  cov7           0.48105      0.06993      6.88
f_rses f_fiq  cov8           0.27235      0.05157      5.28
f_fpa  f_fiq  cov1           0.27533      0.05552      4.96
f_rpa  f_fses cov6           0.07563      0.05077      1.49
f_riq  f_fses cov8           0.27235      0.05157      5.28
f_rses f_fses cov9           0.32046      0.06517      4.92
f_fpa  f_fses cov2           0.00611      0.05085      0.12
f_fiq  f_fses cov3           0.33510      0.05150      6.51
e_rea  e_fea  covea          0.04535      0.02918      1.55
e_roa  e_foa  covoa          0.12085      0.03214      3.76
d_ramb d_famb          0
    
```

When model 7 is tested against models 2, 4, and 5, the *p*-values are respectively 0.0433, 0.0370, and 0.0123, indicating that the combined effect of the reciprocal influence and the covariance of the disturbance terms is statistically significant. Thus, the hypothesis tests indicate that it is acceptable to omit either the reciprocal influences or the covariance of the disturbances but not both.

It is also of interest to test the covariances between the error terms for educational (COVEA) and occupational aspiration (COVOA), since these terms are omitted from

Jöreskog and Sörbom's models. Constraining COVEA and COVOA to zero produces Loehlin's model 6:

```

title2 'Loehlin (1987) analysis: Model 6';
data ram6(type=ram);
  set ram2;
  if _name_='covea' | _name_='covoa' then
  do;
    _name_=' ';
    _estim_=0;
  end;
run;

proc calis data=aspire edf=328 inram=ram6;
run;

```

The relevant output is displayed in Figure 14.27.

| Peer Influences on Aspiration: Haller & Butterworth (1960) Loehlin (1987) analysis: Model 6 | |
|--|-----------|
| The CALIS Procedure | |
| Covariance Structure Analysis: Maximum Likelihood Estimation | |
| Fit Function | 0.1020 |
| Goodness of Fit Index (GFI) | 0.9802 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 0.9638 |
| Root Mean Square Residual (RMR) | 0.0306 |
| Parsimonious GFI (Mulaik, 1989) | 0.6535 |
| Chi-Square | 33.4475 |
| Chi-Square DF | 30 |
| Pr > Chi-Square | 0.3035 |
| Independence Model Chi-Square | 872.00 |
| Independence Model Chi-Square DF | 45 |
| RMSEA Estimate | 0.0187 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | 0.0471 |
| ECVI Estimate | 0.2597 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | 0.3164 |
| Probability of Close Fit | 0.9686 |
| Bentler's Comparative Fit Index | 0.9958 |
| Normal Theory Reweighted LS Chi-Square | 32.9974 |
| Akaike's Information Criterion | -26.5525 |
| Bozdogan's (1987) CAIC | -170.4342 |
| Schwarz's Bayesian Criterion | -140.4342 |
| McDonald's (1989) Centrality | 0.9948 |
| Bentler & Bonett's (1980) Non-normed Index | 0.9937 |
| Bentler & Bonett's (1980) NFI | 0.9616 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.6411 |
| Z-Test of Wilson & Hilferty (1931) | 0.5151 |
| Bollen (1986) Normed Index Rho1 | 0.9425 |
| Bollen (1988) Non-normed Index Delta2 | 0.9959 |
| Hoelter's (1983) Critical N | 431 |

Figure 14.27. Career Aspiration Data: Loehlin Model 6

The chi-square value for testing model 6 versus model 2 is $33.4476 - 19.0697 = 14.3779$ with 2 degrees of freedom and a p -value of 0.0008, indicating that there is considerable evidence of correlation between the error terms.

The following table summarizes the results from Loehlin's seven models.

| Model | χ^2 | df | p-value | SBC |
|--------------------------------|----------|----|---------|-----------|
| 1. Full model | 12.0132 | 13 | 0.5266 | -63.3356 |
| 2. Equality constraints | 19.0697 | 28 | 0.8960 | -143.2200 |
| 3. No SES path | 23.0365 | 29 | 0.7749 | -145.0492 |
| 4. No reciprocal influence | 20.9981 | 29 | 0.8592 | -147.0876 |
| 5. No disturbance correlation | 19.0745 | 29 | 0.9194 | -149.0111 |
| 6. No error correlation | 33.4475 | 30 | 0.3035 | -140.4342 |
| 7. Constraints from both 4 & 5 | 25.3466 | 30 | 0.7080 | -148.5351 |

For comparing models, you can use a DATA step to compute the differences of the chi-square statistics and p-values.

```

title 'Comparisons among Loehlin's models';
data _null_;
  array achisq[7] _temporary_
    (12.0132 19.0697 23.0365 20.9981 19.0745 33.4475 25.3466);
  array adf[7] _temporary_
    (13 28 29 29 29 30 30);
  retain indent 16;
  file print;
  input ho ha @@;
  chisq = achisq[ho] - achisq[ha];
  df = adf[ho] - adf[ha];
  p = 1 - probchi( chisq, df);
  if _n_ = 1 then put
    / +indent 'model comparison   chi**2   df   p-value'
    / +indent '-----';
  put +indent +3 ho ' versus ' ha @18 +indent chisq 8.4 df 5. p 9.4;
datalines;
2 1   3 2   4 2   5 2   7 2   7 4   7 5   6 2
;

```

The DATA step displays the following table.

| Comparisons among Loehlin's models | | | | |
|------------------------------------|------------|---------|----|---------|
| model | comparison | chi**2 | df | p-value |
| 2 | versus 1 | 7.0565 | 15 | 0.9561 |
| 3 | versus 2 | 3.9668 | 1 | 0.0464 |
| 4 | versus 2 | 1.9284 | 1 | 0.1649 |
| 5 | versus 2 | 0.0048 | 1 | 0.9448 |
| 7 | versus 2 | 6.2769 | 2 | 0.0433 |
| 7 | versus 4 | 4.3485 | 1 | 0.0370 |
| 7 | versus 5 | 6.2721 | 1 | 0.0123 |
| 6 | versus 2 | 14.3778 | 2 | 0.0008 |

Although none of the seven models can be rejected when tested against the alternative of an unrestricted covariance matrix, the model comparisons make it clear that

there are important differences among the models. Schwarz's Bayesian Criterion indicates model 5 as the model of choice. The constraints added to model 5 in model 7 can be rejected ($p=0.0123$), while model 5 cannot be rejected when tested against the less-constrained model 2 ($p=0.9448$). Hence, among the small number of models considered, model 5 has strong statistical support. However, as Loehlin (1987, p. 106) points out, many other models for these data could be constructed. Further analysis should consider, in addition to simple modifications of the models, the possibility that more than one friend could influence a boy's aspirations, and that a boy's ambition might have some effect on his choice of friends. Pursuing such theories would be statistically challenging.

References

- Akaike, H. (1987), "Factor Analysis and AIC," *Psychometrika*, 52, 317–332.
- Bentler, P.M. and Bonett, D.G. (1980), "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures," *Psychological Bulletin*, 88, 588–606.
- Bollen, K.A. (1986), "Sample Size and Bentler and Bonett's Nonnormed Fit Index," *Psychometrika*, 51, 375–377.
- Bollen, K.A. (1989), *Structural Equations with Latent Variables*, New York: John Wiley & Sons, Inc.
- Boomsma, A. (1983), *On the Robustness of LISREL (Maximum Likelihood Estimation) against Small Sample Size and Nonnormality*, Amsterdam: Sociometric Research Foundation.
- Duncan, O.D., Haller, A.O., and Portes, A. (1968), "Peer Influences on Aspirations: A Reinterpretation," *American Journal of Sociology*, 74, 119–137.
- Fuller, W.A. (1987), *Measurement Error Models*, New York: John Wiley & Sons, Inc.
- Haller, A.O., and Butterworth, C.E. (1960), "Peer Influences on Levels of Occupational and Educational Aspiration," *Social Forces*, 38, 289–295.
- Hampel F.R., Ronchetti E.M., Rousseeuw P.J. and Stahel W.A. (1986), *Robust Statistics*, New York: John Wiley & Sons, Inc.
- Hoelter, J.W. (1983), "The Analysis of Covariance Structures: Goodness-of-Fit Indices," *Sociological Methods and Research*, 11, 325–344.
- Huber, P.J. (1981), *Robust Statistics*, New York: John Wiley & Sons, Inc.
- James, L.R., Mulaik, S.A., and Brett, J.M. (1982), *Causal Analysis*, Beverly Hills: Sage Publications.
- Jöreskog, K.G. (1973), "A General Method for Estimating a Linear Structural Equation System," in Goldberger, A.S. and Duncan, O.D., eds., *Structural Equation Models in the Social Sciences*, New York: Academic Press.
- Jöreskog, K.G. and Sörbom, D. (1979), *Advances in Factor Analysis and Structural Equation Models*, Cambridge, MA: Abt Books.

- Jöreskog, K.G. and Sörbom, D. (1988), *LISREL 7: A Guide to the Program and Applications*, Chicago: SPSS.
- Keesling, J.W. (1972), “Maximum Likelihood Approaches to Causal Analysis,” Ph.D. dissertation, Chicago: University of Chicago.
- Lee, S.Y. (1985), “Analysis of Covariance and Correlation Structures,” *Computational Statistics and Data Analysis*, 2, 279–295.
- Loehlin, J.C. (1987), *Latent Variable Models*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lord, F.M. (1957), “A Significance Test for the Hypothesis that Two Variables Measure the Same Trait Except for Errors of Measurement,” *Psychometrika*, 22, 207–220.
- McArdle, J.J. and McDonald, R.P. (1984), “Some Algebraic Properties of the Reticular Action Model,” *British Journal of Mathematical and Statistical Psychology*, 37, 234–251.
- McDonald, R.P. (1989), “An Index of Goodness-of-Fit Based on Noncentrality,” *Journal of Classification*, 6, 97–103.
- Schwarz, G. (1978), “Estimating the Dimension of a Model,” *Annals of Statistics*, 6, 461–464.
- Voss, R.E. (1969), “Response by Corn to NPK Fertilization on Marshall and Monona Soils as Influenced by Management and Meteorological Factors,” Ph.D. dissertation, Ames, Iowa: Iowa State University.
- Wiley, D.E. (1973), “The Identification Problem for Structural Equation Models with Unmeasured Variables,” in Goldberger, A.S. and Duncan, O.D., eds., *Structural Equation Models in the Social Sciences*, New York: Academic Press.
- Wilson, E.B. and Hilferty, M.M. (1931), “The Distribution of Chi-Square,” *Proceedings of the National Academy of Sciences*, 17, 694.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *SAS/STAT® User's Guide, Version 8*, Cary, NC: SAS Institute Inc., 1999.

SAS/STAT® User's Guide, Version 8

Copyright © 1999 by SAS Institute Inc., Cary, NC, USA.

ISBN 1-58025-494-2

All rights reserved. Produced in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

U.S. Government Restricted Rights Notice. Use, duplication, or disclosure of the software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

The Institute is a private company devoted to the support and further development of its software and related services.