Some New Features in LISREL 9

August 14, 2012

Contents

Structural equation modelling (SEM) was introduced initially as a way of analyzing a covariance or correlation matrix. Typically, one would read this matrix into LISREL and estimate the model by maximum likelihood. If raw data was available without missing values, one could also use PRELIS first to estimate an asymptotic covariance matrix to obtain robust estimates of standard errors and chi-squares.

Modern structural equation modeling is based on raw data. For examples based on covariance and correlation matrices, see Jöreskog & Sörbom (2003a-b). If raw data is available in a LISREL system file or in a text file one can read the data into LISREL and formulate the model using either SIMPLIS syntax or LISREL syntax and, if requested, LISREL 9 will automatically perform robust estimation of standard errors and chi-square goodness of fit measures under non-normality. It is no longer necessary to estimate an asymptotic covariance matrix with PRELIS and read this into LISREL. The estimation of the asymptotic covariance matrix and the model is now done in LISREL9. If the data contain missing values, LISREL 9 will automatically use FIML to estimate the model. Alternatively, users may choose to impute the missing values by EM or MCMC and estimate the model based on the imputed data. All this works for both continuous and ordinal variables.

Here SIMPLIS syntax will be used to illustrate all examples. Tha data files and syntax files for these examples are given in the folder `\LISREL 9 Examples\ls9ex`. The corresponding syntax in LISREL syntax are also given in that folder. Thus, npv1a.spl is a SIMPLIS syntax file and npv1b.lis is the corresponding input file in LISREL syntax.

1 Continuous Variables without Missing Values

1.1 Confirmatory Factor Analysis

1.1.1 Example: Nine Psychological Variables (NPV)

To illustrate all the different cases and the different steps in the analysis the classical example of confirmatory factor analysis of nine psychological variables (NPV) from the Holzinger-Swineford (1939) study will be used. The nine variables is a subset of 26 variables administered to 145 seventh- and eighth-grade children in the Grant-White school in Chicago. The nine tests are (with the original variable number in parenthesis):

VIS PERC Visual Perception (V1)

CUBES Cubes (V2)

LOZENGES Lozenges (V4)

PAR COMP Paragraph Comprehension (V6)

SEN COMP Sentence Completion (V7)

WORDMEAN Word meaning (V9)

ADDITION Addition (V10)

COUNTDOT Counting dots (V12)

SCCAPS Straight-curved capitals (V13)

It is hypothesized that these variables have three correlated common factors: visual perception here called Visual, verbal ability here called Verbal and speed her called Speed such that the first three variables measure Visual, the next three measure Verbal, and the last three measure Speed. A path diagram of the model to be estimated is given in Figure 1.

Suppose the data is available in a text file **npv.dat** with the names of the variables in the first line. The first few lines of the data file looks like this[1]

```
'VIS PERC' CUBES LOZENGES 'PAR COMP' 'SEN COMP' WORDMEAN ADDITION COUNTDOTS SCCAPS
      33       22       17        8       17       10       65       98      195
      34       24       22       11       19       19       50       86      228
      29       23        9        9       19       11      114      103      144
      16       25       10        8       25       24      112      122      160
      30       25       20       10       23       18       94      113      201
      36       33       36       17       25       41      129      139      333
      28       25        9       10       18       11       96       95      174
      30       25       11       11       21        8      103      114      197
      20       25        6        9       21       16       89      101      178
      27       26        6       10       16       13       88      107      137
      32       21        8        1        7       11      103      136      154
```

1.1.2 Creating a LISREL Data System File

For most analysis with LISREL it is convenient to to work with a LISREL **data system file** of the type **.lsf**. LISREL can import data from many formats such as SAS, SPSS, STATA, and EXCEL. LISREL can also import data in text format with spaces (*.dat or *.raw), commas (*.csv) or tab characters (*.txt) as delimiters between entries. The data is then stored as a LISREL data system file **.lsf**. First we illustrate how to import a text file with spaces as delimiters. The procedure is the same for all other types of files. Importation of data from external sources is described in the PRELIS Guide.

Since the data in this example is a text file with spaces as delimiters, an easy way to create a **.lsf** file is by running the following simple PRELIS syntax file

```
DA NI=9
RA=NPV.DAT LF
CO All
OU RA=NPV.LSF
```

LF is a new PRELIS option to tell PRELIS that the labels are in the first line(s) of the data file.

[1]If a name contains spaces or other special characters, put the name within ' ' as shown.

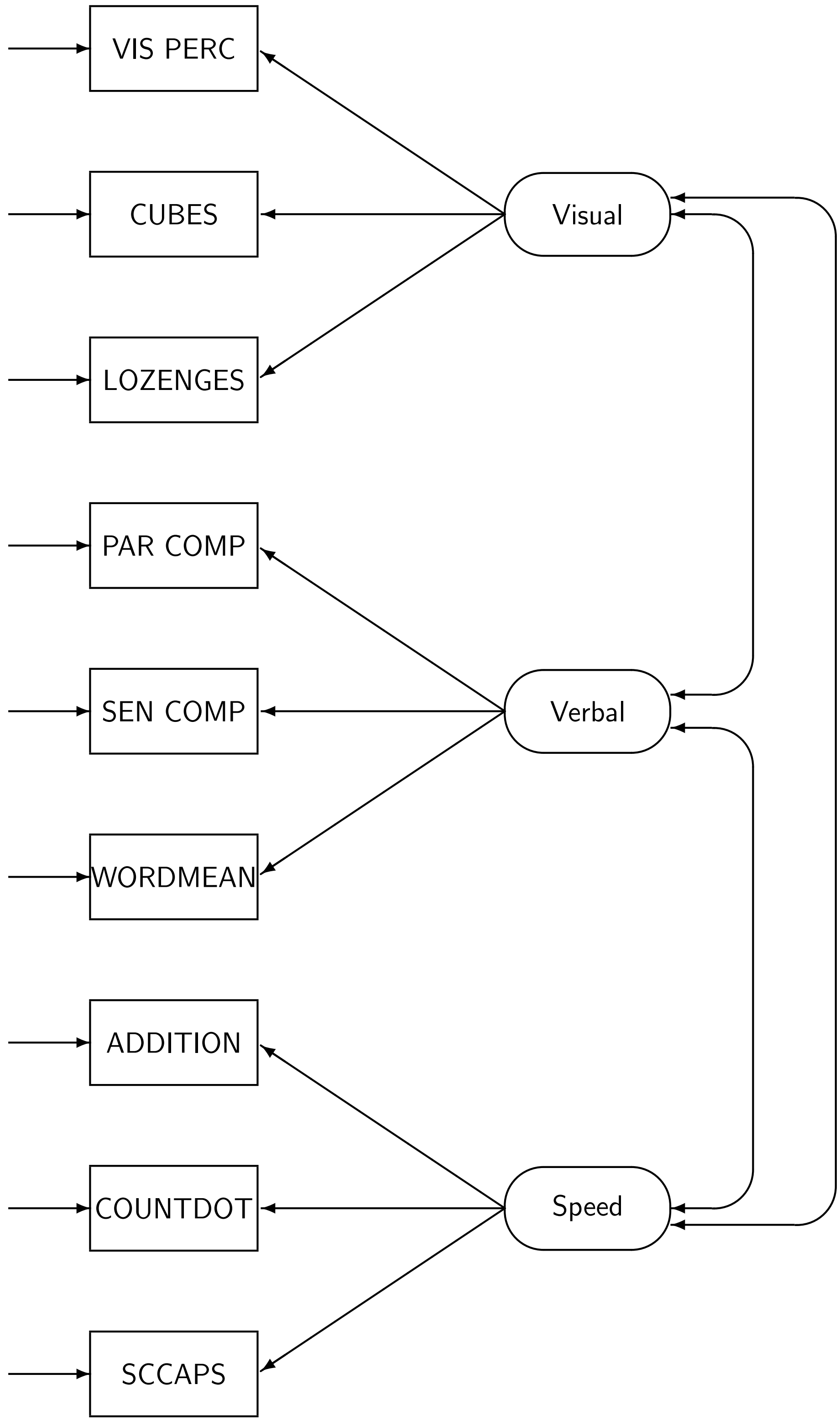


Figure 1: Confirmatory Factor Analysis Model for Nine Psychological Variables

1.1.3 Estimating the Model by Maximum Likelihood

With the **npv.lsf** file on hand, one can estimate the model by normal theory maximum likelihood[2]. The first SIMPLIS file is (see file **npv1a.spl**):

```
Estimation of the NPV Model by Maximum Likelihood
Raw Data from File npv.lsf
Latent Variables: Visual Verbal Speed
Relationships:
   'VIS PERC' - LOZENGES = Visual
   'PAR COMP' - WORDMEAN = Verbal
   ADDITION - SCCAPS = Speed
Path Diagram
End of Problem
```

One can also include a line

```
Path Diagram
```

to display path diagrams with parameter estimates, standard errors, and t-values.

The output file **npv1a.out** shows several sections which are new in LISREL9. The sample covariance matrix $\mathbf{S}$ is given as

```
        Covariance Matrix

           VIS PERC      CUBES   LOZENGES   PAR COMP   SEN COMP   WORDMEAN
           --------   --------   --------   --------   --------   --------
 VIS PERC    47.801
    CUBES    10.013     19.758
 LOZENGES    25.798     15.417     69.172
 PAR COMP     7.973      3.421      9.207     11.393
 SEN COMP     9.936      3.296     11.092     11.277     21.616
 WORDMEAN    17.425      6.876     22.954     19.167     25.321     63.163
 ADDITION    17.132      7.015     14.763     16.766     28.069     33.768
 COUNTDOT    44.651     15.675     41.659      7.357     19.311     20.213
   SCCAPS   124.657     40.803    114.763     39.309     61.230     79.993


        Covariance Matrix

           ADDITION   COUNTDOT     SCCAPS
           --------   --------   --------
 ADDITION   565.593
 COUNTDOT   293.126    440.792
   SCCAPS   368.436    410.823   1371.618
```

After the covariance matrix the following lines are given

[2]The term normal theory maximum likelihood is used to mean that the estimation of the model is based on the assumption that the variables have a multivariate normal distribution.

```
Total Variance = 2610.906 Generalized Variance = 0.106203D+17

Largest Eigenvalue = 1734.725 Smallest Eigenvalue = 3.665

Condition Number = 21.756
```

The total variance is the sum of the diagonal elements of **S** and the generalized variance is the determinant of **S** which equals the product of all the eigenvalues of **S**. The largest and smallest eigenvalues of **S** are also given. These quantities are useful in principal components analysis. The condition number is the square root of the ratio of the largest and smallest eigenvalue. A small condition number indicates multicollinearity in the data. If the condition number is very small LISREL gives a warning. This might indicate that one or more variables are linear or nearly linear combinations of other variables.

LISREL9 gives parameter estimates, standard errors, Z-values, P-values and R^2 for the measurement equations as follows

```
LISREL Estimates (Maximum Likelihood)

        Measurement Equations

VIS PERC = 4.678*Visual, Errorvar.= 25.915, R² = 0.458
Standerr  (0.622)                  (4.566)
Z-values   7.525                    5.675
P-values   0.000                    0.000

   CUBES = 2.296*Visual, Errorvar.= 14.487, R² = 0.267
Standerr  (0.407)                  (1.974)
Z-values   5.642                    7.339
P-values   0.000                    0.000

LOZENGES = 5.769*Visual, Errorvar.= 35.896, R² = 0.481
Standerr  (0.748)                  (6.637)
Z-values   7.711                    5.409
P-values   0.000                    0.000

PAR COMP = 2.922*Verbal, Errorvar.= 2.857 , R² = 0.749
Standerr  (0.236)                  (0.587)
Z-values   12.355                   4.870
P-values   0.000                    0.000

SEN COMP = 3.856*Verbal, Errorvar.= 6.749 , R² = 0.688
Standerr  (0.332)                  (1.161)
Z-values   11.630                   5.812
P-values   0.000                    0.000
```

```
WORDMEAN = 6.567*Verbal, Errorvar.= 20.034, R² = 0.683
Standerr  (0.568)                  (3.407)
Z-values   11.572                   5.880
P-values   0.000                    0.000

ADDITION = 15.676*Speed, Errorvar.= 319.868, R² = 0.434
Standerr  (2.005)                  (48.586)
Z-values   7.819                    6.584
P-values   0.000                    0.000

COUNTDOT = 16.709*Speed, Errorvar.= 161.588, R² = 0.633
Standerr  (1.746)                  (38.034)
Z-values   9.568                    4.248
P-values   0.000                    0.000

  SCCAPS = 25.956*Speed, Errorvar.= 697.900 , R² = 0.491
Standerr  (3.106)                  (116.121)
Z-values   8.357                    6.010
P-values   0.000                    0.000
```

By default LISREL standardizes the latent variables. This seems most resonable since the latent variables are unobservable and have no definite scale. The correlations among the latent variables, with standard errors and Z- values are given as follows

```
             Visual     Verbal      Speed
  Visual      1.000

  Verbal      0.541      1.000
            (0.085)
              6.377

   Speed      0.523      0.336      1.000
            (0.094)    (0.091)
              5.582      3.687
```

These estimates have been obtained by maximizing the likelihood function L under multivariate normality. Therefore it is possible to give the log-likelihood values at the maximum of the likelihood function. It is common the report the value of $-2ln(L)$, sometimes called deviance, instead of L. LISREL9 gives the value $-2\ln(L)$ for the estimated model and for a saturated model. A saturated model is a model where the mean vector and covariance matrix of the multivariate normal distribution are unconstrained.

The log-likelihood values are given in the output as

```
                                   Log-likelihood Values

                            Estimated Model          Saturated Model
                            ---------------          ---------------
Number of free parameters(t)         21                          45
-2ln(L)                        6707.266                    6655.724
AIC (Akaike, 1974)*            6749.266                    6745.724
BIC (Schwarz, 1978)*           6811.777                    6879.677

*LISREL uses AIC= 2t - 2ln(L) and BIC = tln(N)- 2ln(L)
```

LISREL9 also give the values of AIC and BIC. These can be used for the problem of selecting the "best" model from several *a priori* specified models. One then chooses the model with the smallest AIC or BIC. The original papers of Akaike (1974) and Schwarz (1978) define AIC and BIC in terms of $\ln(L)$ but LISREL9 uses $-2\ln(L)$ and the formulas:

$$AIC = 2t - 2ln(L) \ , \tag{1}$$

$$BIC = tln(N) - 2ln(L) \ , \tag{2}$$

where t is the number of free parameters in the model and N is the total sample size.

1.1.4 Testing the Model

Various chi-square statistics are used for testing structural equation models. If normality holds and the model is fitted by the maximum likelihood (ML) method, one such chi-square statistic is obtained as N times the minimum of the ML fit function, where N is the sample size. An asymptotically equivalent chi-square statistic can be obtained from a general formula developed by Browne (1984) and using an asymptotic covariance matrix estimated under multivariate normality, see Section 5.1. These chi-square statistics are denoted C_1 and C_2(NT), respectively. They are valid under multivariate normality of the observed variables and if the model holds.

For this analysis, LISREL9 gives the two chi-square values `C1` and `C2_NT` as

```
Degrees of Freedom For (C1)-(C2)                        24
Maximum Likelihood Ratio Chi-Square (C1)                51.542 (P = 0.0009)
Browne's (1984) ADF Chi-Square (C2_NT)                  48.952 (P = 0.0019)
```

1.1.5 Robust Estimation

The analysis just described assumes that the variables have a multivariate normal distribution. This assumption is questionable in many cases. Although the maximum likelihood parameter estimates are considered to be robust against non-normality, their standard errors and chi-squares are affected by non-normality. It is therefore recommended to use the maximum likelihood method with robustified standard errors and chi-squares, which is called `Robust Maximum Likelihood`. To do so just include a line anywhere between the second line and the last line, see file **npv2a.spl**. This gives the following information about the distribution of the variables.

```
Total Sample Size =    145

Univariate Summary Statistics for Continuous Variables

Variable     Mean  St. Dev.  Skewness  Kurtosis  Minimum Freq.  Maximum Freq.
--------     ----  --------   -------  --------  ------- -----  ------- -----
VIS PERC   29.579     6.914    -0.119    -0.046   11.000     1   51.000     1
   CUBES   24.800     4.445     0.239     0.872    9.000     1   37.000     2
LOZENGES   15.966     8.317     0.623    -0.454    3.000     2   36.000     1
PAR COMP    9.952     3.375     0.405     0.252    1.000     1   19.000     1
SEN COMP   18.848     4.649    -0.550     0.221    4.000     1   28.000     1
WORDMEAN   17.283     7.947     0.729     0.233    2.000     1   41.000     1
ADDITION   90.179    23.782     0.163    -0.356   30.000     1  149.000     1
COUNTDOT  109.766    20.995     0.698     2.283   61.000     1  200.000     1
  SCCAPS  191.779    37.035     0.200     0.515  112.000     1  333.000     1
```

This shows that the range of the variables are quite different, reflecting the case that they are composed of different number of items. For example, PAR COMP has a range of 1 to 19, whereas SCCAPS has a range of 112 to 333. This is also reflected in the means and standard deviations.

LISREL9 also gives tests of univariate and multivariate skewness and kurtosis.

```
Test of Univariate Normality for Continuous Variables

             Skewness         Kurtosis      Skewness and Kurtosis

Variable Z-Score P-Value   Z-Score P-Value   Chi-Square P-Value

VIS PERC  -0.604   0.546     0.045   0.964        0.367   0.833
   CUBES   1.202   0.229     1.843   0.065        4.842   0.089
LOZENGES   2.958   0.003    -1.320   0.187       10.491   0.005
PAR COMP   1.995   0.046     0.761   0.447        4.559   0.102
SEN COMP  -2.646   0.008     0.693   0.489        7.483   0.024
WORDMEAN   3.385   0.001     0.720   0.472       11.977   0.003
ADDITION   0.826   0.409    -0.937   0.349        1.560   0.458
COUNTDOT   3.263   0.001     3.325   0.001       21.699   0.000
  SCCAPS   1.008   0.313     1.273   0.203        2.638   0.267

Relative Multivariate Kurtosis = 1.072
```

```
Test of Multivariate Normality for Continuous Variables
        Skewness                     Kurtosis               Skewness and Kurtosis
   Value  Z-Score P-Value     Value  Z-Score P-Value        Chi-Square P-Value
  ------  ------- -------   -------  ------- -------        ---------- -------
  11.733    5.426   0.000   106.098    3.023   0.003            38.579   0.000
```

It is seen that the hypothesis of zero skewness and kurtosis is rejected for LOZENGES, SEN COMP, WORDMEAN, and COUNTDOT.

The output file **npv2a.out** gives the same parameter estimate as before but different standard errors. As a consequence, also t-values and P-values are different. The parameter estimates and the two sets of standard errors are given in Table 1.1.5.

Table 1: Parameter Estimates, Normal Standard Errors, and Robust Standard Errors

	Parameter	Standard Errors	
Factor Loading	Estimates	Normal	Robust
VIS PERC on Visual	4.678	0.622	0.691
CUBES on Visual	2.296	0.407	0.374
LOZENGES on Visual	5.769	0.748	0.723
PAR COMP on Verbal	2.992	0.236	0.249
SEN COMP on Verbal	3.856	0.332	0.330
WORDMEAN on Verbal	6.567	0.568	0.571
ADDITION on Speed	15.676	2.005	1.824
COUNTDOT on Speed	16.709	1.746	1.769
SCCAPS on Speed	25.956	3.106	3.066
Factor Correlations			
Verbal vs Visual	0.541	0.085	0.093
Verbal vs Speed	0.523	0.094	0.099
Verbal vs Speed	0.336	0.091	0.114

If the observed variables are non-normal, one can use the same formula from Browne (1984) using an asymptotic covariance matrix (ACM)[3] estimated under non-normality. This chi-square, often called the ADF (Asymptotically Distribution Free) chi-square statistic, is denoted C_2(NNT) in LISREL 9[4]. It has been found in simulation studies that the ADF statistic does not work well because it is difficult to estimate the ACM accurately unless N is huge, see *e.g*, Curran, West, & Finch (1996).

Satorra & Bentler (1988) proposed another approximate chi-square statistic C_3, often called the SB chi-square statistic, which is C_1 multiplied by a scale factor which is estimated from the sample and involves estimates of the ACM both under normality and non-normality. The scale factor is estimated such that C_3 has an asymptotically correct

[3]The ACM is an estimate of the covariance matrix of the sample variances and covariances. Under non-normality this involves estimates of fourth-order moments.

[4]In previous versions of LISREL, C_2(NT) and C_2(NNT) was called C2 and C4, respectively.

mean even though it does not have an asymptotic chi-square distribution. In practice, C_3 is conceived of as a way of correcting C_1 for the effects of non-normality and C_3 is often used as it performs better than the ADF test C_2(NT) in LISREL, particularly if N is not very large, see e.g., Hu, Bentler, & Kano (1992).

Satorra & Bentler (1988) also mentioned the possibility of using a Satterthwaite (1941) type correction which adjusts C_1 such that the corrected value has the correct asymptotic mean and variance. This type of fit measure has not been much investigated, neither for continuous nor for ordinal variables. However, this type of chi-square fit statistic has been implemented in LISREL9, where it is denoted C_4. The formulas for C_1–C_4 are given in the Appendix.

For our present example, C_1–C_4 appear in the output as

```
Degrees of Freedom For (C1)-(C3)                       24
Maximum Likelihood Ratio Chi-Square (C1)               51.542 (P = 0.0009)
Browne's (1984) ADF Chi-Square (C2_NT)                 48.952 (P = 0.0019)
Browne's (1984) ADF Chi-Square (C2_NNT)                64.648 (P = 0.0000)
Satorra-Bentler (1988) Scaled Chi-square (C3)          50.061 (P = 0.0014)
Satorra-Bentler (1988) Adjusted Chi-square (C4)        35.134 (P = 0.0056)
Degrees of Freedom For C4                              16.844
```

C_1 and C_2(NT) are the same as before but with robust estimation LISREL9 also gives C_2(NNT), C_3 and C_4 so that one can see what the effect of non-normality is. In particular,the difference $C_2(\text{NNT}) - C_2(\text{NT})$ can be viewed as an effect of non-nomality.

Note that C_4 has its own degrees of freedom which is different from the model degrees of freedom. LISREL 9 gives the degrees of freedom for C_4 as a fractional number and uses this fractional degrees of freedom to compute the P-value for C_4.

1.1.6 Estimation Using Data in Text Form

Since the original data is given in text form in this example, it is not necessary to use a **lsf** file to analyze tha data. One can read the text data file **npv.dat** directly into LISREL using the following SIMPLIS syntax file, see file **npv3a.spl**.

```
Estimation of the NPV Model by Robust Maximum Likelihood
Using text data with Labels in the first line
Raw Data from File NPV.DAT
Continuous 'VIS PERC' - SCCAPS
Latent Variables: Visual Verbal Speed
Relationships:
   'VIS PERC' - LOZENGES = Visual
   'PAR COMP' - WORDMEAN = Verbal
   ADDITION - SCCAPS = Speed
Robust Estimation
Options: RS SC MI
Path Diagram
End of Problem
```

The `Options`line can be used to request additional output, see the **SIMPLIS Syntax Guide**. In this case, `RS` means residuals and standardized residuals, `SC` means completely standardized solution, and `MI` means modification indices.

1.1.7 Modifying the Model

The output file **npv3a.out** gives the following information about modification indices

```
        The Modification Indices Suggest to Add the
  Path to   from      Decrease in Chi-Square    New Estimate
 ADDITION  Visual              8.5                  -6.85
 COUNTDOT  Verbal              8.3                  -4.91
 SCCAPS    Visual             27.8                  23.94
 SCCAPS    Verbal             10.7                  11.01
```

This suggests that the fit cam be improved by adding a path from Visual to SCCAPS. If this makes sense, one can add this path, see file **npv3aa.spl** and rerun the model. This gives a solution where

```
   SCCAPS = 16.559*Visual + 16.274*Speed, Errorvar.= 620.929, R² = 0.547
 Standerr  (3.675)         (3.336)                   (97.603)
 Z-values   4.506           4.879                     6.362
 P-values   0.000           0.000                     0.000
```

and the chi-squares are now

```
 Degrees of Freedom for (C1)-(C3)                     23
 Maximum Likelihood Ratio Chi-Square (C1)             28.293 (P = 0.2049)
 Browne's (1984) ADF Chi-Square (C2_NT)               27.898 (P = 0.2197)
 Browne's (1984) ADF Chi-Square (C2_NNT)              31.701 (P = 0.1065)
 Satorra-Bentler (1988) Scaled Chi-square (C3)        28.221 (P = 0.2075)
 Satorra-Bentler (1988) Adjusted Chi-square (C4)      20.437 (P = 0.2342)
 Degrees of Freedom for C4                            16.656
```

indicating a good fit.

1.1.8 Analyzing Correlations

Factor analysis was mainly developed by psychologists for the purpose of identifying mental abilities by means of psychological testing. Various theories of mental abilities and various procedures for analyzing the correlations among psychological tests emerged.

Following this old tradition, users of LISREL might be tempted to analyze the correlation matrix of the nine psychological variables instead of the covariance matrix as we have done in the previous examples. However, analyzing the correlation matrix by maximum likelihood (ML) is problematic in several ways as pointed out by Cudeck (1989), see also Appendix C in Jöreskog,*et.al.* (2003). There are three ways to resolve this problem:

Approach 1 Use the covariance matrix and ML as before and request the completely standardized solution (SC)[5] as was done on the `Options` line in file **npv3a.spl**. This gives the completely standardized solution in matrix form as

```
Completely Standardized Solution

        LAMBDA-X

               Visual     Verbal      Speed
             --------   --------   --------
 VIS PERC       0.677       - -        - -
    Cubes       0.517       - -        - -
 LOZENGES       0.694       - -        - -
 PAR COMP        - -      0.866        - -
 SEN COMP        - -      0.829        - -
 WORDMEAN        - -      0.826        - -
 ADDITION        - -        - -      0.659
 COUNTDOT        - -        - -      0.796
   SCCAPS        - -        - -      0.701

        PHI

               Visual     Verbal      Speed
             --------   --------   --------
   Visual       1.000
   Verbal       0.541      1.000
    Speed       0.523      0.336      1.000



        THETA-DELTA

             VIS PERC      Cubes   LOZENGES   PAR COMP   SEN COMP   WORDMEAN
             --------   --------   --------   --------   --------   --------
                0.542      0.733      0.519      0.251      0.312      0.317
             ADDITION   COUNTDOT     SCCAPS
             --------   --------   --------
                0.566      0.367      0.509
```

The disadvantage with this alternative is that one does not get standard errors for the completely standardized solution.

[5] LISREL has two kinds of standardized solutions: the standardized solution (SS) in which only the latent variables are standardized and the completely standardized solution (SC) in which both the observed and the latent variables are standardized.

Approach 2 Use the following PRELIS syntax file to standardize the original variables (file **npv2.prl**):

```
RA=NPV.LSF
SV ALL
OU RA=NPVstd.LSF
```

SV is a new PRELIS command to standardize the variables. One can standardize some of the variables by listing them on the SV line. **npv2.prl** produces a new **lsf** file **NPVstd.lsf** in which all variables has sample means 0 and sample standard deviations 1.

Then use NPVstd.LSF instead of NPV.LSF in **npv2a.spl** to obtain a completely standardized solution with robust standard errors.

Approach 3 Use the sample correlation matrix with robust unweighted least squares (RULS) or with robust diagonally weighted least squares (RDWLS) This will use an estimate of the asymptotic covariance matrix of the sample correlations to obtain correct asymptotic standard errors and chi-squares under non-normality.

The following SIMPLIS command file demonstrates the Approach 3, see file **npv4a.spl**:

```
Estimation of the NPV Model
by Robust Diagonally Weighted Least Squares
Using Correlations
Raw Data from File NPV.LSF
Analyze Correlations
Latent Variables: Visual Verbal Speed
Relationships:
   'VIS PERC' - LOZENGES = Visual
   'PAR COMP' - WORDMEAN = Verbal
   ADDITION - SCCAPS = Speed
Robust Estimation
Options: DWLS
Path Diagram
End of Problem
```

Note the added line

```
Analyze Correlations
```

This gives the standadrdized solution as

```
LISREL Estimates (Robust Diagonally Weighted Least Squares)
VIS PERC = 0.726*Visual, Errorvar.= 0.472 , R² = 0.528
Standerr  (0.0707)                (0.195)
Z-values   10.272                  2.424
```

```
P-values   0.000                    0.015

   CUBES = 0.481*Visual, Errorvar.= 0.769 , R² = 0.231
Standerr  (0.0810)                 (0.183)
Z-values   5.938                    4.204
P-values   0.000                    0.000

LOZENGES = 0.677*Visual, Errorvar.= 0.541 , R² = 0.459
Standerr  (0.0687)                 (0.190)
Z-values   9.862                    2.851
P-values   0.000                    0.004

PAR COMP = 0.863*Verbal, Errorvar.= 0.255 , R² = 0.745
Standerr  (0.0326)                 (0.175)
Z-values   26.440                   1.459
P-values   0.000                    0.145

SEN COMP = 0.836*Verbal, Errorvar.= 0.302 , R² = 0.698
Standerr  (0.0350)                 (0.176)
Z-values   23.908                   1.719
P-values   0.000                    0.086

WORDMEAN = 0.823*Verbal, Errorvar.= 0.323 , R² = 0.677
Standerr  (0.0361)                 (0.176)
Z-values   22.777                   1.839
P-values   0.000                    0.066

ADDITION = 0.611*Speed, Errorvar.= 0.627 , R² = 0.373
Standerr  (0.0658)                (0.184)
Z-values   9.291                   3.407
P-values   0.000                   0.001

COUNTDOT = 0.711*Speed, Errorvar.= 0.494 , R² = 0.506
Standerr  (0.0584)                (0.185)
Z-values   12.178                  2.668
P-values   0.000                   0.008

  SCCAPS = 0.842*Speed, Errorvar.= 0.290 , R² = 0.710
Standerr  (0.0584)                (0.193)
Z-values   14.423                  1.508
P-values   0.000                   0.131
```

```
        Correlation Matrix of Independent Variables

                Visual      Verbal       Speed
              --------    --------    --------
  Visual        1.000

  Verbal        0.535       1.000
              (0.084)
                6.336

   Speed        0.571       0.379       1.000
              (0.086)     (0.086)
                6.637       4.384
```

1.2 Structural Equation Models for Latent Variables

1.2.1 Example: Attitudes to Drinking and Driving (DRINK)

In a study designed to determine the predictors of drinking and driving behavior among 18- to 24-year-old males, the model shown in the path diagram in Figure 1.2.1 was proposed.

The latent variables shown in the figure are as follows:

Attitude attitude toward drinking and driving

Norms social norms pertaining to drinking and driving

Control percieved control over drinking and driving

Intention intention to drink and drive

Behavior drinking and driving behavior

Attitude is measured by five indicators X_2-X_5, Norms is measured by three indicators X_6-X_8, Control is measured by four indicators X_9-X_{12}, Intension is measured by two indicators Y_1-Y_2, and behavior is measured by two indicators Y_3-Y_4. Fictitious data based on the theory of planned behavior (Ajzen, 1991) is given in file **drinkdata.lsf**.

This example illustrates a possible strategy of analysis. Begin by testing the measurement model for Attitude, see file **drink11a.spl**.

```
Drinking and Driving
Testing the Measurement Model for Attitude
Raw Data from File drinkdata.lsf
Latent Variables Attitude
Relationships
X1-X5 = Attitude
Path Diagram
End of Problem
```

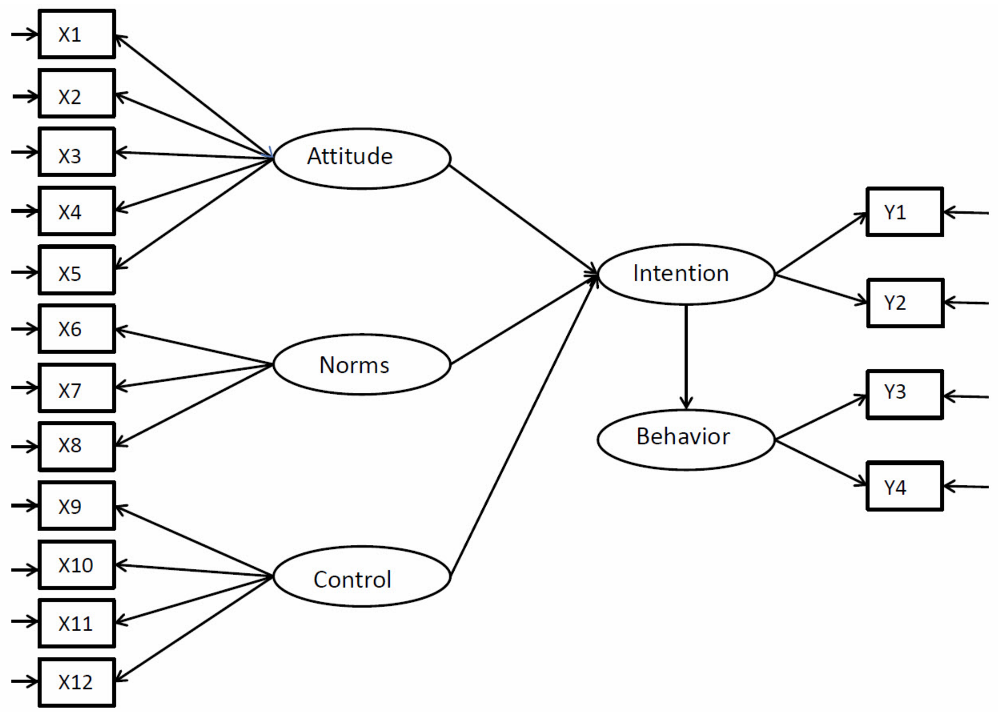


Figure 2: Conceptual Path Diagram for Attitudes to Drinking and Driving

Note that although the data file **drinkdata.lsf** contains many variables, LISREL automatically selects the subset of variables used in the model.

To test the measurement model for Attitude and Norms simultaneously, add Norms in the list of Latent Variables and add the line, see file **drink12a.spl**.

```
X6-X8 = Norms
```

To test the measurement model for Attitude, Norms, and Control simultaneously, add Control in the list of Latent Variables and add the line, see file **drink13a.spl**.

```
X9-X12= Control
```

Finally, to test the measurement model for all latent variables simultaneosly, add Intention and Behavior in the list of Latent Variables and add the two lines, see file **drink14a.spl**.

```
Y1-Y2 = Intention
Y3-Y4 = Behavior
```

If any of these analysis shows a large modification index for an indicator, the measurement model must be reconsidered and modified. For example, suppose there is a large modification index for the path from Attitude to X8. This might mean that X8 is not entirely an indicator of Norms but to some extent also a measure of Attitude. If this idea makes sense then the model should be modified by letting X8 be a composite measure of both Norms and Attitude.

One can now test the full model in Figure 1.2.1 by adding the two lines, see file **drink15a.spl**.

```
Intention = Attitude - Control
Behavior = Intention
```

defining the structural relationships among the latent variables The full SIMPLIS command file is now

```
Drinking and Driving
Raw Data from File drinkdata.lsf
Latent Variables Attitude Norms Control Intention Behavior
Relationships
Y1-Y2 = Intention
Y3-Y4 = Behavior
X1-X5 = Attitude
X6-X8 = Norms
X9-X12= Control
Intention = Attitude - Control
Behavior = Intention
Robust Estimation
Path Diagram
End of Problem
```

According to Browne & Cudeck (1993) one can use the following fit measures, see file **drink15a.out**:

```
 Root Mean Square Error of Approximation (RMSEA)       0.0282
 90 Percent Confidence Interval for RMSEA              (0.0196 ; 0.0363)
 P-Value for Test of Close Fit (RMSEA < 0.05)          1.00
```

to conclude that the model fits at least approximately.

2 Continuous Variables with Missing Values

2.1 Treatment of Missing Values

Missing values and incomplete data are almost unavoidable in in social, behavioral, medical and most other areas of investigation. One can distinguish between three types of incomplete data:

- Unit nonresponse, for example, a person does not respond at all to an item in a questionnaire.
- Subject attrition, for example, when a person falls out of a sample after some time in a longitudinal follow-up study.
- Item nonresponse, for example, a person respond to some but not all items in a questionnaire.

The literature, *e.g.*, Schafer (1997) distinguishes between three mechanisms of nonresponse.

MCAR Missing completely at random

MAR Missing at random

MNAR Missing not at random

Let z_{ij} be any element on the data matrix. Informally, one can define these concepts as

MCAR $Pr(z_{ij} = missing)$ does not depend on any variable in the data.

MAR $Pr(z_{ij} = missing)$ may depend on other variables in the data but not on z_{ij}. Example: A missing value of a person's income may depend on his/her age and education but not on his/her actual income.

MNAR $Pr(z_{ij} = missing)$ depends on z_{ij}. Example: In a questionnaire people with higher income tend not to report their income.

LISREL has several ways of dealing with missing values:

1. Listwise deletion
2. Pairwise deletion
3. Imputation by matching
4. Multiple imputation
    - EM
    - MCMC
5. Full Information Maximum Likelihood (FIML)[6]

Of these methods the first three are *ad hoc* procedures whereas the last two are based probability models for missingness. As a consequence, the *ad hoc* methods may lead to biased estimates under MAR and can only be recommended under MCAR.

Listwise deletion means that all cases with missing values are deleted. This leads to a complete data matrix with no missing values which is used to estimate the model. This procedure can lead to a large loss of information in that the resulting sample size is much smaller than the original. Listwise deletion can give biased, inconsistent, and inefficient estimates under MAR. It should only be used under MCAR.

Pairwise deletion means that means and variances are estimated using all available data for each variable and covariances are estimated using all available data for each pair of

[6]Of course, the maximum likelihood (ML) used earlier is also a full information maximum likelihood method. However, it is convenient to use the term ML for the case of complete data and the term FIML for the case of missing data.

variables. These means, variances and covariances are then combined to form a mean vector and a covariance matrix which are used to estimate the model. While some efficiency is obtained compared to listwise deletion, it is difficult to specify a sample size N to be used in the estimation of the model, since the variances and covariances are all based on different sample sizes and there is no guaranty that the covariance matrix will be positive definite which is required by the maximum likelihood method. Although pairwise deletion is available in LISREL, it is not recommended. Its best use is for data screening for then it gives the most complete information about the missing values in the data.

Imputation means that real values are substituted for the missing values. Various *ad.hoc.* procedures for imputation have been suggested in the literature. One such is imputation by matching which is available in LISREL. It is based on the idea that individuals who have similar values on a set mathing variables may also be similar on a variable with missing values. This will work well if the matching variables are good predictors of the variable with missing values.

Methods 4 and 5 are both based on the assumption of multvariate normality and missingness under MAR. Method 4 uses multiple imputation methods to generate a complete data matrix. The multiple imputation procedure implemented in LISREL is described in details in Schafer (1997) and uses the EM algorithm and the method of generating random draws from probabilty distributions via Markov chains (MCMC). Formulas are given in Section **??**. The EM algorithm generates one single complete data matrix whereas the MCMC method generates several complete data matrices and uses the average of these. As a consequence, the MCMC method is more reliable than the EM algorithm. In both cases, the complete data matrix can be used to estimate the mean vector and the covariance matrix of the observed variables which can be used to estimate the model. However, in LISREL 9 it is not necessary to do these steps separately as they are done automatically as will be described in what follows.

Method 5 is the default method in LISREL9 when there are missing data. This is the recommended method for dealing with the problem of missing data. So this is described first.

If the variables have a multivariate normal distribution all subsets of the variables also have that distribution. So the likelihood function for the observed values can be evaluated for each observation without using any missing values. Formulas are given in Section **??**.

2.2 Latent Curve Models

2.2.1 Example: Treatment of Prostate Cancer (PSAVAR)

A medical doctor offered all his patients diagnosed with prostate cancer a treatment aimed at reducing the cancer activity in the prostate. The severity of prostate cancer is often assessed by a plasma component known as prostate specific antigen (PSA), an enzyme that is elevated in the presence of prostate cancer. The PSA level was measured regularly every three months. The data contains five repeated measurements of PSA. The age of the patient is also included in the data. Not every patient accepted the offer initially and several patients chose to enter the program after the first occasion. Some patients, who accepted the initial offer, are absent at some later occasions for various reasons. Thus there are missing values in the data.

The aim of this study is to answer the following questions: What is the average initial PSA value? Do all patients have the same initial PSA value? Is there an overall effect of treatment. Is there a decline of PSA values over time, and, if so, what is the average rate of decline? Do all patients have the same rate of decline? Does the individual initial PSA value and/or the rate of decline depend on the patient's age?

This is a typical example of repeated measurements data, the analysis of which is sometimes done within the framework of multilevel analysis. It represents the simplest type of two-level model but it can also be analyzed as a structural equation model, see Bollen & Curran (2006). In this context it illustates a mean and covariance structure model estimated from longitudinal data with missing values.

The data file for this example is **psavar.lsf**, where missing values are shown as -9.000[7].

LISREL for Windows - [psavar.PSF]

File Edit Data Transformation Statistics Graphs Multilevel SurveyGLIM View Window Help

	PSA0	PSA3	PSA6	PSA9	PSA12	Age
1	30.400	28.000	26.900	25.200	19.600	69.000
2	27.800	26.700	20.500	18.700	18.800	58.000
3	26.600	21.800	17.800	17.900	14.500	53.000
4	24.800	24.500	20.200	19.800	18.800	61.000
5	33.700	30.300	25.400	27.300	20.100	63.000
6	26.500	24.600	20.900	-9.000	18.900	49.000
7	26.200	24.400	21.800	22.200	18.400	63.000
8	24.800	19.500	18.000	16.100	12.500	49.000
9	28.400	-9.000	22.500	19.400	22.900	63.000
10	26.100	-9.000	23.300	22.000	14.600	56.000
11	28.800	31.300	-9.000	23.100	22.800	68.000
12	29.800	-9.000	25.600	24.500	21.000	67.000
13	22.900	23.900	-9.000	19.400	15.600	47.000
14	30.100	27.700	25.700	20.400	20.800	56.000
15	26.500	-9.000	-9.000	20.000	17.400	57.000
16	-9.000	-9.000	17.100	12.900	-9.000	43.000

In this kind of data it is inevitable that there are missing values. For example, a patient may be on vacation or ill or unable to come to the doctor for any reason at some occasion or a patient may die and therefore will not come to the doctor after a certain occasion. It is seen in that

- Patients 9 and 10 are missing at 3 months
- Patient 15 is missing at 3 and 6 months
- Patient 16 is missing at 0, 3, and 12 months

In the following analysis it is assumed that data are missing at random (MAR), although there may be a small probability that a patient will be missing because his PSA value is high.

[7] If the data is imported from an external source which already have a missing value code, the missing values will show up in the **lsf** file as -999999.000, which is the global missing data code in LISREL.

2.2.2 Data Screening

Whenever one starts an analysis of a new data set, it is recommended to begin with a data screening. To do so click on **Statistics** at the top of the screen and select **Data Screening** from the **Statistics** menu. This will reveal the follwing information about the data.

```
 Number of Missing Values per Variable

      PSA0      PSA3      PSA6      PSA9     PSA12       Age
  --------  --------  --------  --------  --------  --------
        17        14        13        12        11         0
```

This table says that there are 17 patients missing initially, 14 missing at 3 months, 13 at 6 months, etc.

```
 Distribution of Missing Values

 Total Sample Size =    100

 Number of Missing Values     0     1     2     3
          Number of Cases    46    43     9     2
```

This table says that there are only 46 patients with complete data on all six occasions. Thus, if one uses listwise deletion 54% of the sample will be lost. 43 patients are missing on one occasion, 9 patients are missing at two occasions, 2 patients are missing on three occasions. This table does not tell on which occasions the patients are missing. The next table gives gives more complete information about the missing data patterns.

```
Missing Data Map

Frequency PerCent   Pattern
       46    46.0   0 0 0 0 0 0
        9     9.0   1 0 0 0 0 0
        8     8.0   0 1 0 0 0 0
        2     2.0   1 1 0 0 0 0
        8     8.0   0 0 1 0 0 0
        2     2.0   1 0 1 0 0 0
        2     2.0   0 1 1 0 0 0
        9     9.0   0 0 0 1 0 0
        1     1.0   1 0 0 1 0 0
        1     1.0   1 1 0 1 0 0
        1     1.0   0 0 1 1 0 0
        9     9.0   0 0 0 0 1 0
        1     1.0   1 0 0 0 1 0
        1     1.0   1 1 0 0 1 0
```

The columns under `Pattern` correspond to the variables in the order they are in **psavar.lsf**. A 0 means a non-missing value and a 1 means a missing value. Recall that the last variable is the patient's age. This has no missing values. Here one can see for example that two patients are missing at both 0 and 3 months and another patient is missing at 6 and 9 months.

The following information about the univariate distributions of the variables have been obtained using all available data for each variable,*i.e.*, 83 patients for `PSA0`, 86 patients for `PSA3`, etc.

```
Univariate Summary Statistics for Continuous Variables

Variable     Mean  St. Dev.  Skewness  Kurtosis  Minimum Freq.  Maximum Freq.
--------     ----  --------  --------  --------  ------- -----  ------- -----
    PSA0   31.164     5.684     0.068    -0.852   19.900     1   44.100     1
    PSA3   30.036     6.025    -0.248    -0.732   14.500     1   42.100     1
    PSA6   27.443     6.084    -0.335    -0.961   13.700     1   37.600     1
    PSA9   25.333     6.391    -0.331    -1.066   10.600     1   36.200     1
   PSA12   23.406     6.306    -0.309    -1.069    9.600     1   35.800     1
     Age   55.450     7.896    -0.329    -0.234   32.000     1   70.000     1
```

It is seen that the mean age is 55.45 years and that average initial PSA value is 31.164 with a minimum at 19.9 and maximum at 44.1. At 12 months the corresponding values are 23.406, 9.6, and 35.8, respectively. Thus there is some evidence that the PSA values are decreasing over time.

The model to be estimated is

$$y_{it} = a_i + b_i T_t + e_{it} \tag{3}$$

$$i = 1, 2, \ldots, N \;\; \text{individuals} \tag{4}$$

$$T_t = \text{Time at occasion } t = 1, 2, \ldots, n_i \tag{5}$$

$$a_i = \alpha + \gamma_a z_i + u_i \tag{6}$$

$$b_i = \beta + \gamma_b z_i + v_i \tag{7}$$

$$z_i = \text{Covariate observed on individual } i \tag{8}$$

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N(\mathbf{0}, \mathbf{\Phi}) \tag{9}$$

$$e_{it} \sim N(0, \sigma_e^2) \tag{10}$$

$$y_{it} = (\alpha + \beta T_t + \gamma_a z_i + \gamma_b T_t z_i) + (u_i + v_i T_t + e_{it}) \tag{11}$$

An interpretation of this is as follows. Each patient has his own linear growth curve[8], represented by (3) which is the regression of y_{it} on time with intercept a_i and slope b_i varying across patients. In principle, the intercepts a_i and slopes b_i could all be different across patients. It is of interest to know if the intercepts and/or the slopes are equal across patients. The four cases are illustrated in Figure 3. If there is variation in intercepts and/or the slopes across patients, one is interested in whether a covariate z_i (in this case age) can predict the intercept and/or the slope.

Path diagrams for the models without and with a covariate are illustrated in Figures 4 and 5, respectively, with $T_t = t - 1$ for four occasions.

The model in Figure 4 can be estimated with FIML using the following SIMPLIS syntax file (**psavar1a.spl**):

```
Linear Growth Curve for psavar Data
Raw Data from File psavar.LSF
Latent Variables: a b
Relationships
PSA0  = 1*a 0*b
PSA3  = 1*a 3*b
PSA6  = 1*a 6*b
PSA9  = 1*a 9*b
PSA12 = 1*a 12*b
a b = CONST
Equal Error Variances: PSA0 - PSA12
Path Diagram
End of Problem
```

There are two latent variables `a` and `b` in the model. They represent the intercept and slope of the patients linear growth curves. The objective is to estimate the mean vector and covariance matrix of `a` and `b` and the error variance of the psa meaures. The error variance is assumed to be the same at all occasions.

In the current example, `a` and `b` are latent variables, and the line in the input file **psavar1a.spl**

[8] In general, the growth curves are not restricted to be linear, but can be quadratic, cubic, or other types of functions of time, see Jöreskog, Sörbom, Du Toit, & Du Toit (2003).

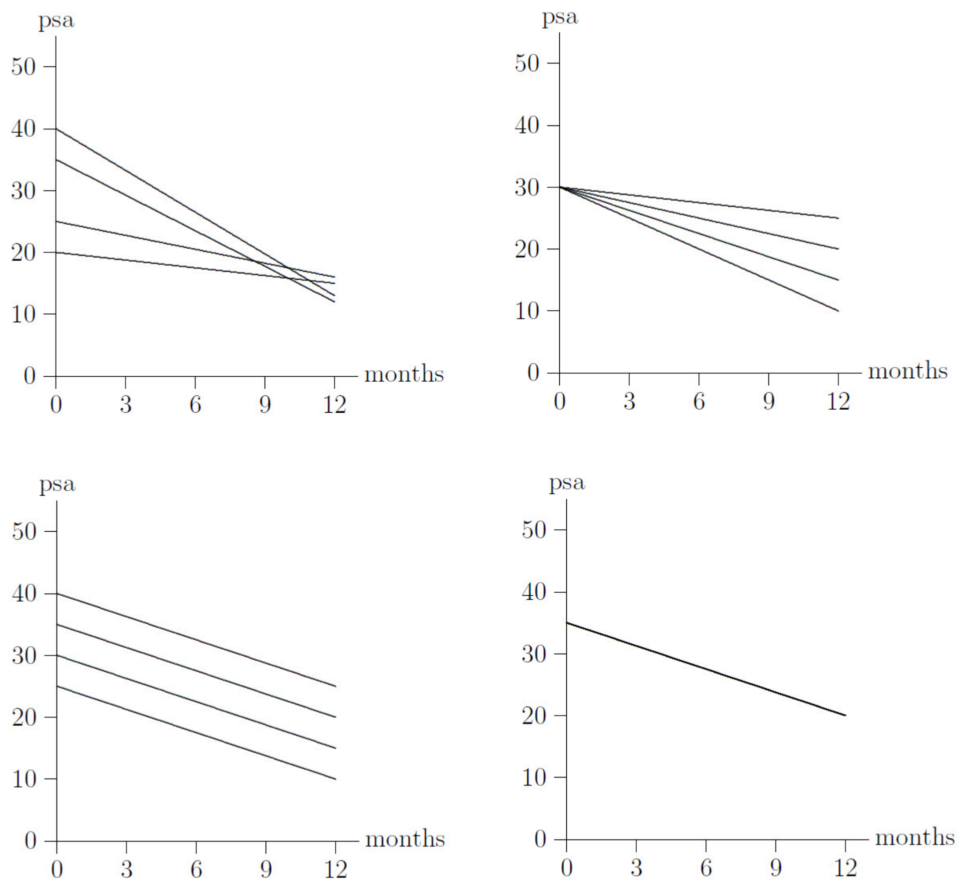


Figure 3: Four Cases of Intercepts and Slopes

```
a b = CONST
```

specifies that the means of `a` and `b` should be estimated.

The output gives the following information

```
        --------------------------------
         EM Algorithm for missing Data:
        --------------------------------

    Number of different missing-value patterns=       14
    Effective sample size:      100

    Convergence of EM-algorithm in     9 iterations
    -2 Ln(L) =     1997.49237
    Percentage missing values=  13.40
```

The EM algorithm is first used to estimate a saturated model where both the mean vector and covariance matrix are unconstrained. This also gives the value $-2\ln(L) = 1997.4924$.

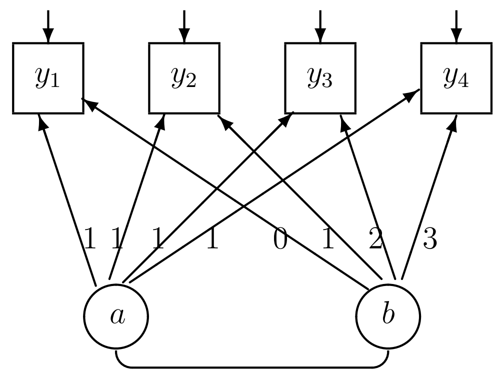


Figure 4: Path Diagram for a Linear Curve Model with Four Occasions

These are used to obtain starting values for the FIML method. After convergence the FIML method gives the following information about the fit of the model.

```
              Global Goodness of Fit Statistics, FIML case


                -2ln(L) for the saturated model =        1997.492
                -2ln(L) for the fitted model    =        2008.601

 Degrees of Freedom = 14
 Full Information ML Chi-Square                        11.108 (P = 0.6775)
 Root Mean Square Error of Approximation (RMSEA)       0.0
 90 Percent Confidence Interval for RMSEA              (0.0 ; 0.0775)
 P-Value for Test of Close Fit (RMSEA < 0.05)          0.844
```

The FIML estimates of the model parameters are given as

```
         Covariance Matrix of Independent Variables

                    a           b
               --------    --------
        a        30.899
                (4.613)
                  6.698

        b         0.302       0.004
                (0.108)     (0.005)
                  2.811       0.728
```

```
          Mean Vector of Independent Variables

                   a           b
              --------    --------
                31.934      -0.742
               (0.571)     (0.019)
                55.920     -39.869
```

The conclusions from this analysis are

- The average initial PSA value is 31.9 with a variance of 30.9.
- Thus, the initial PSA value varies considerably from patient to patient
- The effect of treatment is highly significant.
- The PSA value decreases by 0.7 per quarter (0.23 per year) and this rate of decrease is the same for all patients.

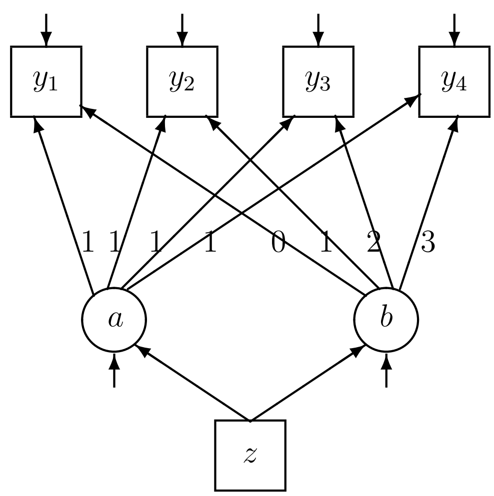


Figure 5: The Linear Curve Model with Covariate

To estimate the model in Figure 5 one can just add Age on the lines for a and b. The SIMPLIS syntax file is **psavar2a.spl**:

```
Linear Model with Covariate for psavar Data
Raw Data from File psavar.LSF
Latent Variables: a b
Relationships
PSA0 = 1*a 0*b
PSA3 = 1*a 3*b
PSA6 = 1*a 6*b
```

```
PSA9 = 1*a 9*b
PSA12 = 1*a 12*b
a b = CONST Age
Let the Errors on a and b correlate
Equal Error Variances: PSA0 - PSA12
Path Diagram
End of Problem
```

However, since we already know that all patients have the same slope b, it is not meaningful to predict b from Age. Thus instead of the line

```
a b = CONST Age
```

one should use, see file **psavar2aa.spl**

```
a = CONST Age
b = CONST 0*Age
```

The prediction equation for the intercept a is estimated as

```
        a = 15.288 + 0.300*Age, Errorvar.= 25.818, R² = 0.177
 Standerr  (3.709)  (0.0662)              (3.911)
 Z-values   4.121    4.533                 6.601
 P-values   0.000    0.000                 0.000
```

Thus, the intercept a depends on age. The intercept increases by 0.30 per year of age, on average.

There is an alternative method of estimation, based on the same two assumptions. One can use multiple imputation to obtain a complete data set and then analyze this by maximum likelihood or robust maximum likelihood method. Since the sample size $N = 100$ is small it is best to use maximum likelihood.

For the model in the last analysis the SIMPLIS syntax will be, see file **psavar3a.spl**:

```
Linear Model with Covariate for psavar Data
Estimated by ML using Multiple Imputation
Raw Data from File psavar.lsf
Multiple Imputation with MC
Latent Variables: a b
Relationships
PSA0  = 1*a 0*b
PSA3  = 1*a 3*b
PSA6  = 1*a 6*b
PSA9  = 1*a 9*b
PSA12 = 1*a 12*b
a = CONST Age
b = CONST 0*Age
```

```
Let the Errors of a and b correlate
Equal Error Variances: PSA0 - PSA12
Path Diagram
End of Problem
```

The only difference between this input file and **psavar2aa.spl** is the line

```
Multiple Imputation
```

which has been added. The output gives the following estimated equation for a:

```
        a = 15.000 + 0.306*Age, Errorvar.= 26.020, R² = 0.183
 Standerr  (3.570)  (0.0637)              (3.849)
 Z-values   4.202    4.798                 6.761
 P-values   0.000    0.000                 0.000
```

which is very similar to previous results.

An advantage of this approach is the one can get more measures of goodness of fit:

```
                              Log-likelihood Values
                     Estimated Model          Saturated Model
 Number of free parameters(t)       9                      27
 -2ln(L)                     1787.430                1764.336
 AIC (Akaike, 1974)*         1805.430                1818.336
 BIC (Schwarz, 1978)*        1828.877                1888.675

                           Goodness of Fit Statistics
 Degrees of Freedom for (C1)-(C2)                      18
 Maximum Likelihood Ratio Chi-Square (C1)              23.095 (P = 0.1870)
 Browne's (1984) ADF Chi-Square (C2_NT)                22.807 (P = 0.1981)
 Estimated Non-centrality Parameter (NCP)              5.095
 90 Percent Confidence Interval for NCP                (0.0 ; 21.634)

 Minimum Fit Function Value                            0.231
 Population Discrepancy Function Value (F0)            0.0509
 90 Percent Confidence Interval for F0                 (0.0 ; 0.216)
 Root Mean Square Error of Approximation (RMSEA)       0.0532
 90 Percent Confidence Interval for RMSEA              (0.0 ; 0.110)
 P-Value for Test of Close Fit (RMSEA < 0.05)          0.426
```

3 Ordinal Variables without Missing Values

3.1 Ordinal Variables

Observations on an ordinal variable represent responses to a set of ordered categories, such as a five-category Likert scale. It is only assumed that a person who selected one category has more of a characteristic than if he/she had chosen a lower category, but we do not know how much more. Ordinal variables are not continuous variables and should not be treated as if they are. It is common practice to treat scores 1, 2, 3, ... assigned to categories as if they have metric properties but this is wrong. Ordinal variables do not have origins or units of measurements. Means, variances, and covariances of ordinal variables have no meaning. The only information we have are counts of cases in response vector. To use ordinal variables in structural equation models requires other techniques than those that are traditionally employed with continuous variables.

There are many methods available for estimating structural equation models with ordinal variables, see *e.g.*, Jöreskog & Moustaki (2001), Yang-Wallentin, Jöreskog, & Luo (2010) or Forero, Maydeu-Olivares, & Gallardo-Pujol (2011) and references given there. There are essentially two types of methods. One is a full information maximum likelihood method (FIML) using a probit, logit, or other link function. The other method fits the model to a matrix of polychoric correlation or covariance matrix using some fit function like like ULS or DWLS. The focus is here on the two methods: FIML and DWLS. These methods are illustrated using exploratory and confirmatory factor analysis models, but any LISREL model can be used, even those involving a mean structure.

3.1.1 Example: Attitudes Toward Science and Technology (SCITECH)

In the Eurobarometer Survey 1992, citizens of Great Britain were asked questions about Science and Technology. The questions are given below.

1. Science and technology are making our lives healthier, easier and more comfortable [COMFORT].
2. Scientific and technological research cannot play an important role in protecting the environment and repairing it [ENVIRON].
3. The application of science and new technology, will make work more interesting [WORK].
4. Thanks to science and technology, there will be more opportunities for the future generations [FUTURE].
5. New technology does not depend on basic scientific research [TECHNOL].
6. Scientific and technological research do not play an important role in industrial development [INDUSTRY].
7. The benefits of science are greater than any harmful effects it may have [BENEFIT].

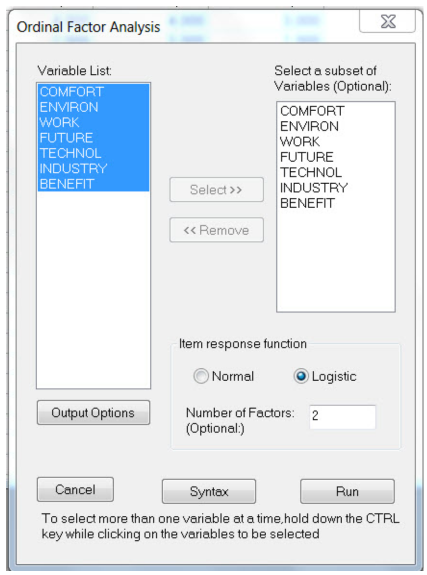
Note that the items COMFORT, WORK, FUTURE, and BENEFIT have a positive question wording, whereas the items ENVIRON, TECHNOL, and INDUSTRY have a negative question wording.

The response alternatives were strongly disagree, disagree to some extent, agree to some extent, and, strongly agree. These were coded as 1, 2, 3, 4, respectively. The data file is

scitech.lsf. Missing values have been deleted beforehand, see Batholomew *et.al.* (2002). This data set is used here to illustrate the case of no missing values. The sample size is 392. Exploratory factor analysis is illustrated first.

3.1.2 Exploratory Factor Analysis

To perform an exploratory factor analysis of ordinal variables by the FIML method, open the file **scitech.lsf** and select **EFA of Ordinal Variables** in the **Statistics Menu**. This shows the following window



Then select the variables to be analyzed, tick the box **Logistic**, for example, and insert 2 for **Number of Factors**. Then click **Run**.

The output file **scitech.out** gives the following unrotated and rotated factor loadings:

Unrotated Factor Loadings

	Factor 1	Factor 2	Unique Var
COMFORT	0.764	0.000	0.416
ENVIRON	0.237	0.817	0.277
WORK	0.669	-0.404	0.388
FUTURE	0.847	-0.323	0.179
TECHNOL	0.234	0.832	0.252
INDUSTRY	0.462	0.716	0.274
BENEFIT	0.713	-0.225	0.441

Varimax-Rotated Factor Loadings

	Factor 1	Factor 2	Unique Var
COMFORT	0.719	0.259	0.416
ENVIRON	-0.054	0.849	0.277
WORK	0.767	-0.153	0.388
FUTURE	0.906	-0.016	0.179
TECHNOL	-0.062	0.862	0.252
INDUSTRY	0.192	0.830	0.274
BENEFIT	0.747	0.030	0.441

Promax-Rotated Factor Loadings

	Factor 1	Factor 2	Unique Var
COMFORT	0.710	0.240	0.416
ENVIRON	-0.086	0.852	0.277
WORK	0.774	-0.175	0.388
FUTURE	0.908	-0.041	0.179
TECHNOL	-0.094	0.866	0.252
INDUSTRY	0.161	0.826	0.274
BENEFIT	0.747	0.010	0.441

Factor Correlations

	Factor 1	Factor 2
Factor 1	1.000	
Factor 2	0.064	1.000

Reference Variables Factor Loadings

	Factor 1	Factor 2	Unique Var
COMFORT	0.738	0.273	0.416
ENVIRON	0.006	0.851	0.277
WORK	0.757	-0.140	0.388
FUTURE	0.906	0.000	0.179
TECHNOL	0.000	0.865	0.252
INDUSTRY	0.251	0.837	0.274
BENEFIT	0.750	0.043	0.441

```
Factor Correlations

           Factor 1   Factor 2
           --------   --------
Factor 1      1.000
Factor 2     -0.089      1.000
```

The first solution is the standardized unrotated solution which is a transformation of the standardized solution obtained by the FIML procedure. The second solution is the varimax solution of Kaiser (1958). Both of these are orthogonal solutions, i.e., the factors are uncorrelated. The third solution is the promax solution of Hendrickson & White (1964). This is an oblique solution, i.e., the factors are correlated. The varimax and the promax solutions are transformations of the standardized unrotated solution and as such they are still maximum likelihood solutions. The fourth solution is the TSLS solution obtained in reference variables form as described the Technical Appendix. The reference variables are chosen as those variables in the promax solution that have the largest factor loadings in each column. In this case the reference variables are FUTURE and TECHNOL.

The Promax-Rotated Factor Loadings and the Reference Variables Factor Loadings suggest that there are two nearly uncorrelated factors and that Factor 1 is a Positive factor, having large loadings on the positively worded items COMFORT, WORK, FUTURE, and BENEFIT and Factor 2 is a Negative factor having large loadings on the negatively worded items ENVIRON, TECHNOL, and INDUSTRY. Three other files are also produced in this analysis: **SCITECH.POM**, **BIVFITS.POM** and **MULFITS.POM**. The first of these gives estimates of unstandardized factor loadings and thresholds and their standard errors. The other two give detailed information about the fit of the model.

The file **SCITECH.POM** give the following unstandardized parameters and standard errors

```
Unstandardized Thresholds Alpha^(i)_a
 COMFORT   -5.015   -2.744    1.535
 ENVIRON   -3.429   -1.243    1.002
    WORK   -2.941   -0.905    2.279
  FUTURE   -4.998   -2.125    1.885
 TECHNOL   -4.164   -1.482    1.086
INDUSTRY   -4.695   -2.507    0.461
 BENEFIT   -3.379   -1.006    1.705

Unstandardized Factor Loadings Beta_ij
 COMFORT    1.184    0.000
 ENVIRON    0.450    1.552
    WORK    1.074   -0.649
  FUTURE    2.002   -0.763
 TECHNOL    0.467    1.656
INDUSTRY    0.882    1.366
 BENEFIT    1.073   -0.339
```

```
 Standard Errors for Unstandardized Thresholds Alpha^(i)_a
  COMFORT    0.589    0.240    0.173
  ENVIRON    0.304    0.173    0.172
     WORK    0.251    0.147    0.227
   FUTURE    0.600    0.331    0.307
  TECHNOL    0.376    0.195    0.195
 INDUSTRY    0.444    0.243    0.168
  BENEFIT    0.290    0.143    0.179

 Standard Errors for Unstandardized Factor Loadings Beta_ij
  COMFORT    0.187    0.000
  ENVIRON    0.222    0.228
     WORK    0.192    0.191
   FUTURE    0.434    0.298
  TECHNOL    0.253    0.272
 INDUSTRY    0.252    0.233
  BENEFIT    0.172    0.188
```

3.1.3 Estimation Using Adaptive Quadrature

The above FIML approach to analyze ordinal variables is only intended for exploratory factor analysis with one or two factors. For other cases one can use the following SIMPLIS syntax file instead, see file **scitech2a.spl**:

```
Raw Data from file SCITECH.lsf
$ADAPQ(8) LOGIT
Latent Variables  Factor1 Factor2
Relationships
COMFORT - BENEFIT = Factor1
ENVIRON - BENEFIT = Factor2
Set the covariance between Factor1 and Factor2 to 0
End of Problem
```

This uses adaptive qudrature to evaluate the integrals involved.

The line

```
$ADAPQ(8) LOGIT
```

specifies that the adaptive quadrature procedure is to be used. One can specify the number of quadrature points to be used in the adaptive quadrature procedure and the link function. Follow the following guidelines to specify the number of quadrature points:

- For models with one latent variable, use 8 quadrature points
- For models with two or three latent variables, use 5-10 quadrature points

- For models with four or five quadrature points, use 5-6 quadrature points
- For models with six to ten latent variables use 3-4 quadrature points

The following link functions are available

- LOGIT
- PROBIT
- LOGLOG
- CLL (Complimentary log-log)

The adaptave quadrature gives a more accurate solution than **scitech.out**. But the unstandardized parameters are essentially the same in this case.

This output file also gives the following information:

```
Number of quadrature points =              8
Number of free parameters =               34
Number of iterations used =                8

-2lnL (deviance statistic) =      5822.09128
Akaike Information Criterion      5890.09128
Schwarz Criterion                 6025.11419
```

3.1.4 Confirmatory Factor Analysis

This FIML approach will now be used to illustrate a confirmatory factor analysis. Use the the file **scitech3a.spl**:

```
Raw Data from file scitech.lsf
$ADAPQ(8) LOGIT
Latent Variables Positive Negative
Relationships
COMFORT WORK FUTURE BENEFIT = Positive
ENVIRON TECHNOL INDUSTRY    = Negative
End of Problem
```

This gives

```
Number of quadrature points =              8
Number of free parameters =               29
Number of iterations used =               24

-2lnL (deviance statistic) =      5841.79954
Akaike Information Criterion      5899.79954
Schwarz Criterion                 6014.96613
```

The difference between the two deviance statistics is $5841.800 - 5822.091 = 18.29$. If this is used as a chi-square with 5 degrees of freedom, the confirmatory factor model is rejected in favor of the exploratory two-factor model. This means that not all of the seven specified zero factor loadings are zero. One can relax the loading of COMFORT on Negative, see file **scitech4a.spl**. This gives a good two-factor solution with a small but significant loading of COMFORT on Negative. The deviance for this model is $-2 \ln L = 5833.559$, corresponding to a chi-square of 11.468 with 4 degrees of freedom which is significant at the 1% level. This result can be interpreted as follows. Some fraction of people who generally have a positive attitude to science and technology will respond in the Strongly Disagree or Disagree to Some Extent categories to the COMFORT item.

4 Ordinal Variables with Missing Values

To illustrate the analysis of ordinal variables in this section some data from the Political Action Survey will be used. This was a cross-national survey designed and carried out to obtain information on conventional and unconventional forms of political participation in industrial societies (Barnes & Kaase, 1979).

The first Political Action Survey was conducted between 1973 and 1975 in eight countries: Britain, West Germany, The Netherlands, Austria, the USA, Italy, Switzerland, and Finland. New cross-sections including a panel were obtained during 1980–81 in three of the original countries: West Germany, The Netherlands, and the USA. All data was collected through personal interviews on representative samples of the population 16 years and older[9].

The Political Action Survey contains several hundred variables. For the present purpose of illustration the six variables representing the operational definition of *political efficacy* will be used. These item have been previously analyzed by Aish & Jöreskog (1990), Jöreskog (1990), and Jöreskog & Moustaki (2001, 2006), among others To begin with we use the data from the first cross-section of the USA sample.

4.1 Example: Measurement of Political Efficacy (EFFICACY)

The conceptual definition of political efficacy *is the feeling that individual political action does have, or can have, an impact upon the political process* (Campbell, *et al.*, 1954). The operational definition of political efficacy is based on the responses to the following six items:[10]

`NOSAY` People like me have no say in what the government does

[9]The data was made available by the Zentralarchiv für Empirische Sozialforschung, University of Cologne. The data was originally collected by independent institutions in different countries. Neither the original collectors nor the Zentralarchiv bear any responsibility for the analysis reported here.

[10]These are the questions that were used in the USA. In Britain, the same questions were used with *Congress in Washington* replaced by *Parliament*. In the other countries the corresponding questions were used in other languages.

`VOTING` Voting is the only way that people like me can have any say about how the government runs things

`COMPLEX` Sometimes politics and government seem so complicated that a person like me cannot really understand what is going on

`NOCARE` we don't think that public officials care much about what people like me think

`TOUCH` Generally speaking, those we elect to Congress in Washington lose touch with the people pretty quickly

`INTEREST` Parties are only interested in people's votes but not in their opinions

Permitted responses to these statements were

AS agree strongly

A agree

D disagree

DS disagree strongly

DK don't know

NA no answer

These responses were coded 1, 2, 3, 4, 8, 9, respectively. The data used here is the USA sample from the 1973 the Political Action Survey which was a cross-national survey designed and carried out to obtain information on conventional and unconventional forms of political participation in industrial societies (Barnes & Kaase, 1979). The data file is **efficacy.dat**, a text file with spaces as delimiters.

4.1.1 Data Screening

Most raw data from surveys are downloaded from large files at data archives and stored on media like diskettes or tapes for analysis. The data file may contain many variables on many cases. Before doing more elaborate analysis of the data, it is important to do a careful data screening to check for coding errors and other mistakes in the data. Such a data screening will also reveal outliers and other anomalies, and detect if there are specific patterns of missing values in the data. The data screening gives a general idea of the character and quality of the data. To get a complete data screening of all values in the data file, use the following PRELIS command file, see **efficacy1.prl**:

```
Screening of Efficacy Data
DA NI=6
LA
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
```

```
RA=EFFICACY.DAT
CL NOSAY-INTEREST 1=AS 2=A 3=D 4=DS 8=DK 9=NA
OU
```

PRELIS determines the sample size, all distinct data values for each variable and the absolute and relative frequency of occurrence of each value. The output file shows that there are 1719 cases in the data, that there are six distinct values on each variable, labeled `AS, A, D, DS, DK`, and `NA`, and the distribution of the data values over these categories.

The results are presented in compact form in Table 2.

Table 2: Univariate Marginal Distributions

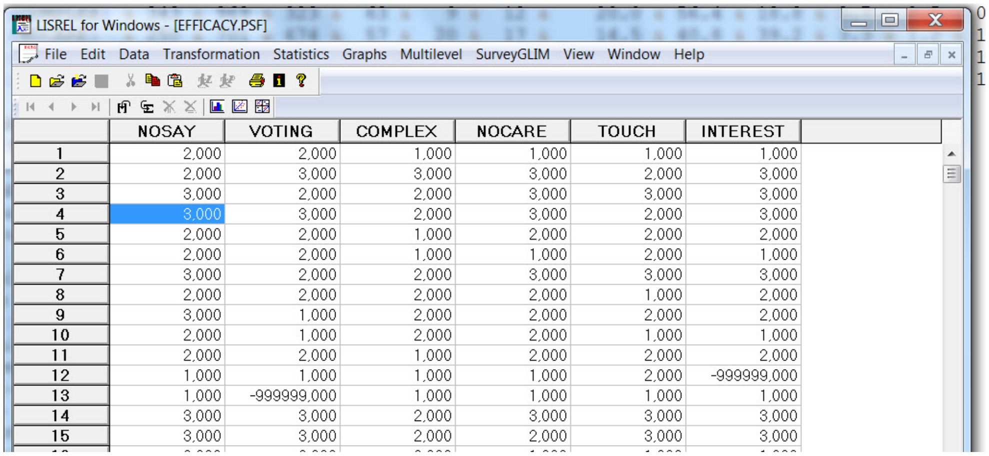
	Frequency						Percentage					
	AS	**A**	**D**	**DS**	**DK**	**NA**	**AS**	**A**	**D**	**DS**	**DK**	**NA**
NOSAY	175	518	857	130	29	10	10.2	30.1	49.9	7.6	1.7	0.6
VOTING	283	710	609	80	26	11	16.5	41.3	35.4	4.7	1.5	0.6
COMPLEX	343	969	323	63	9	12	20.0	56.4	18.8	3.7	0.5	0.7
NOCARE	250	701	674	57	20	17	14.5	40.8	39.2	3.3	1.2	1.0
TOUCH	273	881	462	26	60	17	15.9	51.3	26.9	1.5	3.5	1.0
INTEREST	264	762	581	31	62	19	15.4	44.3	33.8	1.8	3.6	1.1

Note that there are more people responding *Don't Know* on Touch and Interest.

Obviously, the responses *Don't Know* and *No Answer* cannot be used as categories for the ordinal scale that goes from *Agree Strongly* to *Disagree Strongly*. To proceed with the analysis, one must first define the *Don't Know* and *No Answer* responses as missing values. This can be done by adding `MI=8,9` on the `DA` line. In additon by adding `RA=EFFICAY.LSF` on the `OU` line, one will obtain a LISREL data system file **EFFICACY.LSF** which will serves as a basis for furter analysis. The PRELIS command file now looks like this, see file **efficay2.prl**:

```
Creation of a lsf file for Efficacy Data
DA NI=6 MI=8,9
LA
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
RA=EFFICACY.DAT
CL NOSAY-INTEREST 1=AS 2=A 3=D 4=DS
OU RA=EFFICACY.LSF
```

The first 15 lines of **efficacy.lsf** looks like tihs



	NOSAY	VOTING	COMPLEX	NOCARE	TOUCH	INTEREST
1	2.000	2.000	1.000	1.000	1.000	1.000
2	2.000	3.000	3.000	3.000	2.000	3.000
3	3.000	2.000	2.000	3.000	3.000	3.000
4	3.000	3.000	2.000	3.000	2.000	3.000
5	2.000	2.000	1.000	2.000	2.000	2.000
6	2.000	2.000	1.000	1.000	2.000	1.000
7	3.000	2.000	2.000	3.000	3.000	3.000
8	2.000	2.000	2.000	2.000	1.000	2.000
9	3.000	1.000	2.000	2.000	2.000	2.000
10	2.000	1.000	2.000	2.000	1.000	1.000
11	2.000	2.000	1.000	2.000	2.000	2.000
12	1.000	1.000	1.000	1.000	2.000	-999999.000
13	1.000	-999999.000	1.000	1.000	1.000	1.000
14	3.000	3.000	2.000	3.000	3.000	3.000
15	3.000	3.000	2.000	2.000	3.000	3.000

Note that missing values now appear as -999999.000 which is the global missing value code in LISREL.

To perform a data screening of **efficacy.lsf**, select **Data Screening** in the **Statistics** menu. This gives the following results.

The distribution of missing values over variables are given first.

```
Number of Missing Values per Variable

    NOSAY    VOTING    COMPLEX     NOCARE     TOUCH  INTEREST
       39        37         21         37        77        81
```

It is seen that there are only 21 missing values on `COMPLEX` whereas there are 77 and 81 on `TOUCH` and `INTEREST`, respectively. As we already know that most of the missing values on `TOUCH` and `INTEREST` are *Don't Know* rather than *No Answer* responses, it seems that these items are considered by the respondents to be more difficult to answer.

Next in the output is the distribution of missing values over cases.

```
Distribution of Missing Values

Total Sample Size =   1719

Number of Missing Values      0      1      2      3      4      5      6
         Number of Cases   1554    106     26     18      4      2      9
```

It is seen that there are only 1554 out of 1719 cases without any missing values. The other 165 cases have one or more missing values. With listwise deletion this is the loss of sample size that will occur. Most, or 106, of the 165 cases with missing values have only one missing value. But note that there are 9 cases with 6 missing values, *i.e.*, these cases have either not responded or have responded *Don't Know* to all of the six items. These 9 cases are of course useless for any purpose considered here.

The next two tables of output give information about sample sizes for all variables and all pairs of variables. These sample sizes are given in abslute numbers as well as in percentages.

```
Effective Sample Sizes
Univariate (in Diagonal) and Pairwise Bivariate (off Diagonal)

                 NOSAY     VOTING    COMPLEX     NOCARE      TOUCH   INTEREST
              --------   --------   --------   --------   --------   --------
    NOSAY         1680
   VOTING         1658       1682
  COMPLEX         1670       1674       1698
   NOCARE         1655       1656       1675       1682
    TOUCH         1620       1627       1635       1622       1642
 INTEREST         1619       1621       1632       1622       1598       1638
```

This table gives the univariate and bivariate sample sizes. Thus, there are 1680 cases with complete data on `NOSAY` but only 1638 cases with complete data on `INTEREST`. There are 1658 cases with complete data on *both* `NOSAY` and `VOTING` but only 1598 cases with complete data on *both* `TOUCH` and `INTEREST`.

The same kind of information, but in terms of percentage of missing data instead, is given in the following table.

```
Percentage of Missing Values
Univariate (in Diagonal) and Pairwise Bivariate (off Diagonal)

                 NOSAY     VOTING    COMPLEX     NOCARE      TOUCH   INTEREST
              --------   --------   --------   --------   --------   --------
    NOSAY         2.27
   VOTING         3.55       2.15
  COMPLEX         2.85       2.62       1.22
   NOCARE         3.72       3.66       2.56       2.15
    TOUCH         5.76       5.35       4.89       5.64       4.48
 INTEREST         5.82       5.70       5.06       5.64       7.04       4.71
```

The next lines give all possible patterns of missing data and their sample frequencies. Each column under `Pattern` corresponds to a variable. A 0 means a complete data and a 1 means a missing data.

```
Missing Data Map

Frequency PerCent   Pattern
     1554    90.4   0 0 0 0 0 0
       16     0.9   1 0 0 0 0 0
       12     0.7   0 1 0 0 0 0
        1     0.1   1 1 0 0 0 0
        4     0.2   0 0 1 0 0 0
       11     0.6   0 0 0 1 0 0
       31     1.8   0 0 0 0 1 0
```

```
 1      0.1    0 1 0 0 1 0
 2      0.1    1 1 0 0 1 0
 1      0.1    0 1 1 0 1 0
 4      0.2    0 0 0 1 1 0
 1      0.1    0 0 1 1 1 0
32      1.9    0 0 0 0 0 1
 1      0.1    0 1 0 0 0 1
 1      0.1    1 1 0 0 0 1
 1      0.1    1 0 1 0 0 1
 5      0.3    0 0 0 1 0 1
 2      0.1    1 0 0 1 0 1
 1      0.1    0 0 1 1 0 1
 1      0.1    1 0 1 1 0 1
14      0.8    0 0 0 0 1 1
 4      0.2    1 0 0 0 1 1
 4      0.2    0 1 0 0 1 1
 2      0.1    1 1 0 0 1 1
 1      0.1    0 1 1 0 1 1
 1      0.1    0 0 0 1 1 1
 2      0.1    0 1 1 1 1 1
 9      0.5    1 1 1 1 1 1
```

Thus, there are 1554 cases or 90.4% with no missing data, there are 16 cases or 0.9% with missing values on NOSAY only, and 1 case with missing values on both NOSAY and VOTING, etc. Note again that there are 9 cases with missing values on all 6 variables.

This kind of information is very effective in detecting specific patterns of missingness in the data. In this example there are no particular patterns of missingness. The only striking feature is that there are more missing values on `TOUCH` and `INTEREST`. We know from the first run that these are mainly *Don't know* responses.

The rest of the output (not shown here) gives the distribution of the 1554 cases of the listwise sample over the four ordinal categories for each variable. This shows that most people answer either *agree* or *disagree*. Fewer people answer with the *stronger* alternatives.

4.1.2 Estimating Models by FIML Using Adaptive Quadrature

First we investigate whether the six items meusure one unidimensional latent variable and we begin with FIML, see **efficacy2a.spl**:

```
Efficacy: Model 1 Estimated by FIML
Raw Data from file EFFICACY.LSF
$ADAPQ(8) PROBIT
Latent Variable Efficacy
Relationships
NOSAY - INTEREST = Efficacy
End of Problem
```

The output gives the following factor loadings

```
   NOSAY = 0.739*Efficacy, Errorvar.= 1.000, R² = 0.353
Standerr  (0.0407)
Z-values   18.154
P-values   0.000

  VOTING = 0.377*Efficacy, Errorvar.= 1.000, R² = 0.124
Standerr  (0.0324)
Z-values   11.643
P-values   0.000

 COMPLEX = 0.601*Efficacy, Errorvar.= 1.000, R² = 0.265
Standerr  (0.0375)
Z-values   16.042
P-values   0.000

  NOCARE = 1.656*Efficacy, Errorvar.= 1.000, R² = 0.733
Standerr  (0.103)
Z-values   16.007
P-values   0.000

   TOUCH = 1.185*Efficacy, Errorvar.= 1.000, R² = 0.584
Standerr  (0.0632)
Z-values   18.754
P-values   0.000

INTEREST = 1.361*Efficacy, Errorvar.= 1.000, R² = 0.649
Standerr  (0.0744)
Z-values   18.290
P-values   0.000
```

Note the small loading on VOTING. This indicates very low validity and reliability of the VOTING item which might be explained as follows. If the six items really measure one unidimensional trait Efficacy, then people who are high on Efficacy are supposed to disagree or disagree strongly and people who are low on Efficacy should agree or agree strongly to all items. If this is the case, there would be a positive association between the latent variable Efficacy and each ordinal variable. But isn't something wrong with VOTING? If one is high on Efficacy and one believes that voting is the only way one can influence politics, then one would agree or agree strongly to the VOTING statement. This fact in itself is sufficient to suggest that the VOTING item should be excluded from further consideration.

The output also gives the following information:

```
Number of quadrature points =              8
Number of free parameters =               24
```

```
Number of iterations used =                7

-2lnL (deviance statistic) =      19934.56514
Akaike Information Criterion      19982.56514
Schwarz Criterion                 20113.22711
```

For the moment we note the value of the deviance statistic $-2 \ln L = 19934.465$. Since there is no value of $-2 \ln L$ for a saturated model, it is impossible to say wether this is large or small in some absolute sense. The deviance statistic can therefore only be used to compare different models for the same data.

The output also gives estiamates of the thresholds, their standard errors and z-values. The thresholds are parameters of the model but are seldom useful in analysis of a single sample.

Threshold estimates and standard deviations

Threshold	Estimates	S.E.	Est./S.E.
TH1_NOSAY	-1.57282	0.05484	-28.67862
TH2_NOSAY	-0.26243	0.03847	-6.82165
TH3_NOSAY	1.74605	0.05834	29.92866
TH1_VOTING	-1.02350	0.03901	-26.23749
TH2_VOTING	0.24600	0.03301	7.45288
TH3_VOTING	1.78347	0.05658	31.52036
TH1_COMPLEX	-0.96727	0.04174	-23.17335
TH2_COMPLEX	0.87494	0.04015	21.79122
TH3_COMPLEX	2.04308	0.06656	30.69345
TH1_NOCARE	-2.01404	0.10803	-18.64379
TH2_NOCARE	0.35551	0.06034	5.89170
TH3_NOCARE	3.39121	0.16957	19.99895
TH1_TOUCH	-1.49811	0.06756	-22.17305
TH2_TOUCH	0.84076	0.05430	15.48431
TH3_TOUCH	3.21173	0.13297	24.15415
TH1_INTEREST	-1.65353	0.07801	-21.19712
TH2_INTEREST	0.56685	0.05494	10.31724
TH3_INTEREST	3.42422	0.15211	22.51147

It has been suggested in the political science literature that there are two components of Political Efficacy: Internal Efficacy (here called *Efficacy*) indicating *individuals self-perceptions that they are capable of understanding politics and competent enough to participate in political acts such as voting*, and External Efficacy (here called *Responsiveness* and abbreviated *Respons*) indicating *the belief that the public cannot influence political outcomes because government leaders and institutions are unresponsive* (Miller, *et al.*, 1980; Craig & Maggiotto, 1982). With this view, NOSAY and COMPLEX are indicators of Efficacy and TOUCH and INTEREST are indicators of Respons. The statement NOCARE contains two referents: *public officials* and *people like me*. This statement might elicit perceptions of the responsiveness of *government officials* to public opinion generally, in which case the emphasis is on the political actors, or it might express the opinions of *people like me* in which case the emphasis is on the respondent. In the first case, NOCARE measures Respons; in the second case, it measures Efficacy. I will therefore consider NOCARE as a complex variable, *i.e.*, as a variable measuring both Efficacy and Respons or a mixture of them. This is Model 2. A SIMPLIS command file for Model 2 is, see file **efficacy3a.spl**:

```
Efficacy: Model 2 Estimated by FIML
Raw Data from file EFFICACY.LSF
$ADAPQ(8) PROBIT GR(5)
Latent Variables Efficacy Respons
Relationships
NOSAY - NOCARE = Efficacy
NOCARE - INTEREST = Respons
End of Problem
```

This gives the following estimated factor loadings

```
   NOSAY = 0.916*Efficacy, Errorvar.= 1.000, R² = 0.456
Standerr  (0.0601)
Z-values   15.253
P-values   0.000

  VOTING = 0.461*Efficacy, Errorvar.= 1.000, R² = 0.175
Standerr  (0.0385)
Z-values   11.981
P-values   0.000

 COMPLEX = 0.686*Efficacy, Errorvar.= 1.000, R² = 0.320
Standerr  (0.0455)
Z-values   15.091
P-values   0.000

  NOCARE = 0.821*Efficacy + 0.903*Respons, Errorvar.= 1.000, R² = 0.723
Standerr  (0.131)          (0.108)
Z-values   6.275            8.339
P-values   0.000            0.000

   TOUCH = 1.333*Respons, Errorvar.= 1.000, R² = 0.640
Standerr  (0.0778)
Z-values   17.138
P-values   0.000
```

```
INTEREST = 1.607*Respons, Errorvar.= 1.000, R² = 0.721
Standerr  (0.107)
Z-values   14.996
P-values   0.000
```

Note that the loading of NOCARE on Efficacy is almost as large as that on Respons.

The correlation between the two components Efficacy and Respons is estimated as 0.75 as shown in the next part of the output:

```
         Correlation Matrix of Independent Variables

            Efficacy    Respons
            --------   --------
 Efficacy      1.000

  Respons      0.752      1.000
             (0.030)
              25.051
```

The correlation 0.75 is highly significant meaning that it is significant from 0. But more interestingly it is also significant from 1. An aproximate 95% confidence interval for the correlation is from 0.69 to 0.81.

This model has two more parameters than the previous model. The deviance statistic for this model is 19858.061 as shown in the next part of the output:

```
    Number of quadrature points =                8
    Number of free parameters =                 26
    Number of iterations used =                 17

    -2lnL (deviance statistic) =       19858.06108
    Akaike Information Criterion       19910.06108
    Schwarz Criterion                  20051.61154
```

The difference between the deviance statistic for this model and the deviance statistic for the unidimensional model is 19934.565-19858.061=76.504, which suggests that Model 2 fits the data much better than Model 1.

4.1.3 Estimation by Robust Diagonally Weighted Least Squares (RDWLS)

An alternative approach to estimate models for ordinal variables from data with missing values is to impute the missing values first and then estimate the model by *Robust Unweighted Least Squares* (RULS) or *Robust Diagonally Weighted Least Squares* (RDWLS), see Yang-Wallentin, Jöreskog, & Luo (2010) or Forero, Maydeu-Olivares, & Gallardo-Pujol (2011). This can be done for Model 2 using the following SIMPLIS syntax file, see file **efficacy4a.spl**:

```
Efficacy: Model 2 Estimated by Robust Diagonally Weighted Least Squares
Raw Data from file EFFICACY.LSF
Multiple Imputation with MC
Latent Variables Efficacy Respons
Relationships
NOSAY COMPLEX NOCARE = Efficacy
NOCARE - INTEREST = Respons
Robust Estimation
Method of Estimation: Diagonally Weighted Least Squares
Path Diagram
End of Problem
```

This approach is much faster than the FIML approach especially for large number of variables.

5 Appendix

5.1 General Covariance Structures

The majority of LISREL models are estimated using maximum likelihood (ML) method. This estimation method is appropriate for variables that are continuous, and at least approximately normally distributed. In many research problems the variables under study are neither normal nor even approximately continuous and the use of ML is not valid. An important technical development has been in extending the class of estimation methods to procedures that are correct when used with many different kinds of variables. This more general approach to estimation includes ML, GLS, and ULS and other methods as special cases. But it also applies to very different statistical distributions.

The statistical inference problem associated with all kinds of structural equation models, including factor analysis models, can be formulated very generally and compactly as follows. For the orginal formulation see, *e.g.*, Browne (1984) and Satorra (1989).

Let $\mathsf{Vec}(\mathbf{S})$ be the column vector formed by the columns of $\mathbf{S}$ stringed under each other. Thus, $\mathsf{Vec}(\mathbf{S})$ is of order $k^2 \times 1$. Since $\mathbf{S}$ is a symmetric matrix, the covariance matrix of $\mathsf{Vec}(\mathbf{S})$ is singular. Therefore, it is convenient to work with $\mathbf{s}$ instead, where $\mathbf{s}$ is a vector of the non-duplicated elements of $\mathbf{S}$:

$$\mathbf{s}' = (s_{11}, s_{21}, s_{22}, s_{31}, s_{32}, \ldots,) \ . \tag{12}$$

The relationship between $\mathbf{s}$ and $\mathsf{Vec}(\mathbf{S})$ is

$$\mathbf{s} = \mathbf{K}'\mathsf{Vec}(\mathbf{S}) \ , \tag{13}$$

where $\mathbf{K}$ is a matrix of order $k^2 \times \frac{1}{2}k(k+1)$. Each column of $\mathbf{K}$ has one nonzero value which is 1 for a diagonal element and $\frac{1}{2}$ for a non-diagonal element. The reverse relationship of (13) is

$$\mathsf{Vec}(\mathbf{S}) = \mathbf{K}(\mathbf{K}'\mathbf{K})^{-1}\mathbf{s} = \mathbf{D}\mathbf{s} \ , \tag{14}$$

where $\mathbf{D} = \mathbf{K}(\mathbf{K}'\mathbf{K})^{-1}$ is the duplication matrix, see Magnus & Neudecker (1988).

Similarly, let $\boldsymbol{\sigma} = \mathbf{K}'\mathsf{Vec}(\boldsymbol{\Sigma})$ be the corresponding vector of the non-duplicated elements of $\boldsymbol{\Sigma}$ and suppose that

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\theta}) \ , \tag{15}$$

is a differentiable function of a parameter vector $\boldsymbol{\theta}$. For example, in a LISREL model, $\boldsymbol{\theta}$ is a vector of all independent parameters in all parameter matrices $\boldsymbol{\Lambda}_y$, $\boldsymbol{\Lambda}_x$, $\mathbf{B}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$, $\boldsymbol{\Theta}_\epsilon$, $\boldsymbol{\Theta}_\delta$, and $\boldsymbol{\Theta}_{\delta\epsilon}$ but models which are not LISREL models can also be used. For an example see Jöreskog & Sörbom (1999), pp 347–348).

The sample vector $\mathbf{s}$ has a limiting normal distribution when $N \rightarrow \infty$. Browne (1984) showed that

$$N^{\frac{1}{2}}(\mathbf{s} - \boldsymbol{\sigma}) \stackrel{d}{\longrightarrow} N(\mathbf{0}, \boldsymbol{\Omega}) \ , \tag{16}$$

where $\stackrel{d}{\rightarrow}$ denotes convergence in distribution. Under general assumptions about the distribution of the observed variables, the elements of the covariance matrix $\boldsymbol{\Omega}$ are given by (Browne, 1984, eq. 2.2)

$$\omega_{ghij} = \sigma_{ghij} - \sigma_{gh}\sigma_{ij} \ , \tag{17}$$

where

$$\sigma_{ghij} = E[(z_g - \mu_g)(z_h - \mu_h)(z_i - \mu_i)(z_j - \mu_j)] \ , \tag{18}$$

is a fourth order central moment, and

$$\sigma_{gh} = E[(z_g - \mu_g)(z_h - \mu_h)] \ . \tag{19}$$

Under normality

$$\omega_{ghij} = \sigma_{gi}\sigma_{hj} + \sigma_{gj}\sigma_{hi} \ , \tag{20}$$

which can be written in matrix form as

$$\boldsymbol{\Omega} = 2\mathbf{K}'(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\mathbf{K} \ , \tag{21}$$

where $\otimes$ denotes a Kronecker product. Note that under normality $\boldsymbol{\Omega}$ is a function of $\boldsymbol{\Sigma}$ only so that no fourth order moments are involved.

Let $\mathbf{W}$ be a consistent estimate of $\boldsymbol{\Omega}$. In analogy with (17) – (19), the elements of $\mathbf{W}$ are obtained as

$$w_{gh,ij} = m_{ghij} - s_{gh}s_{ij} \ , \tag{22}$$

where

$$m_{ghij} = (1/N)\sum_{a=1}^{N}(z_{ag} - \bar{z}_g)(z_{ah} - \bar{z}_h)(z_{ai} - \bar{z}_i)(z_{aj} - \bar{z}_j) \tag{23}$$

is a fourth-order central sample moment.

Under normality we use the estimate

$$\mathbf{W} = 2\mathbf{K}'(\hat{\boldsymbol{\Sigma}} \otimes \hat{\boldsymbol{\Sigma}})\mathbf{K} \ , \tag{24}$$

in analogy with (21). Whenever it is necessary to distinguish the general $\mathbf{W}$ defined by (22) from the specific $\mathbf{W}$ in (24), the notation $\mathbf{W}_{\mathrm{NT}}$ is used for $\mathbf{W}$ in (24) and $\mathbf{W}_{\mathrm{NNT}}$ for the general $\mathbf{W}$ in (22).

To estimate the model, consider the minimization of the fit function

$$F(\mathbf{s}, \boldsymbol{\theta}) = [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})]'\mathbf{V}[\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})] \tag{25}$$

where $\mathbf{V}$ is either a fixed positive definite matrix or a random matrix converging in probability to a positive definite matrix $\overline{\mathbf{V}}$. The fit functions available in LISREL are Unweighted Least Squares (ULS)[11], Generalized Least Squares (GLS), Maximum Likelihood (ML), Diagonally Weighted Least Squares (DWLS), and Weighted Least Squares (WLS). They correspond to taking $\mathbf{V}$ in (25) as

$$\mathsf{ULS}: \quad \mathbf{V} = \mathbf{K}'(\mathbf{I} \otimes \mathbf{I})\mathbf{K} \tag{26}$$

$$\mathsf{GLS}: \quad \mathbf{V} = \mathbf{D}'(\mathbf{S}^{-1} \otimes \mathbf{S}^{-1})\mathbf{D} \tag{27}$$

$$\mathsf{ML}: \quad \mathbf{V} = \mathbf{D}'(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \hat{\boldsymbol{\Sigma}}^{-1})\mathbf{D} \tag{28}$$

$$\mathsf{DWLS}: \quad \mathbf{V} = \mathbf{D}_{\mathrm{W}} = [diag\mathbf{W}]^{-1} \tag{29}$$

$$\mathsf{WLS}: \quad \mathbf{V} = \mathbf{W}^{-1} \tag{30}$$

The matrix $\bar{\mathbf{V}}$ is unkown for all methods except ULS but can be estimated by $\hat{\mathbf{V}}$, where

$$\mathsf{ULS}: \quad \hat{\mathbf{V}} = \mathbf{V} \tag{31}$$

$$\mathsf{GLS}: \quad \hat{\mathbf{V}} = \mathbf{W}_{\mathrm{NT}}^{-1} \tag{32}$$

$$\mathsf{ML}: \quad \hat{\mathbf{V}} = \mathbf{W}_{\mathrm{NT}}^{-1} \tag{33}$$

$$\mathsf{DWLS}: \quad \hat{\mathbf{V}} = [diag\mathbf{W}_{\mathrm{NNT}}]^{-1} \tag{34}$$

$$\mathsf{WLS}: \quad \hat{\mathbf{V}} = \mathbf{W}_{\mathrm{NNT}}^{-1} \tag{35}$$

Note that

$$\mathbf{W}_{\mathrm{NT}}^{-1} = [2\mathbf{K}'(\hat{\boldsymbol{\Sigma}} \otimes \hat{\boldsymbol{\Sigma}})\mathbf{K}]^{-1} = \frac{1}{2}\mathbf{D}'(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \hat{\boldsymbol{\Sigma}}^{-1})\mathbf{D} = \frac{1}{2}\mathbf{V}_{\mathrm{ML}} \,, \tag{36}$$

as shown by Browne (1977, equations 20 and 21).

The fit function for ML is usually written

$$\mathsf{F}[\mathbf{S}, \boldsymbol{\Sigma}(\boldsymbol{\theta})] = \ln \mid \boldsymbol{\Sigma} \mid + \mathsf{tr}(\mathbf{S}\boldsymbol{\Sigma}^{-1}) - \ln \mid \mathbf{S} \mid - k \tag{37}$$

but Browne (1974) showed that minimizing F with $\mathbf{V}$ in (28) and minimizing (37) are equivalent. Minimizing F with $\mathbf{V}$ in (28) can be interpreted as ML estimated by means of iteratively reweighted least squares in which $\hat{\boldsymbol{\Sigma}}$ is updated in each iteration. Both of these fit functions have a minimum at the same point in the parameter space, namely at the ML estimates. However, the minimum value of the functions are not the same.

All fit functions are non-negative and equal to zero only for a saturated model, where $\hat{\boldsymbol{\Sigma}} = \mathbf{S}$.

[11] The ULS fit function was orginally defined in Jöreskog & Sörbom (1988) as $\sum_{i=1}^{k}\sum_{j=1}^{k}(s_{ij} - \sigma_{ij})^2$, which translate the matrix $\mathbf{V}$ in (26).

Under multivariate normality of the observed variables, ML and GLS estimates are asymptotically equivalent in the sense that they have the same asymptotic covariance matrix. There is no advantage in using GLS except when $\boldsymbol{\sigma}(\boldsymbol{\theta})$ is a linear function. In practice ML is most often used.

Under non-normality, WLS, also called ADF, see Browne (1984), is in principle the best method since it is valid for any non-normal distribution for continuous variables. But in practice this method does not work well because it is difficult to determine $\mathbf{W}$ (and hence $\mathbf{W}^{-1}$) accurately unless N is huge.

Let $\hat{\boldsymbol{\theta}}$ be the minimizer of $F(\mathbf{s}, \boldsymbol{\theta})$ and let $\boldsymbol{\theta}_0$ be a unique minimizer of $F(\boldsymbol{\sigma}, \boldsymbol{\theta})$. We assume here that the model holds so that $F(\boldsymbol{\sigma}, \boldsymbol{\theta}_0) = 0$. See Browne (1984), Satorra (1989), and Foss, Jöreskog & Olsson (2011) for the case where the model does not hold.

Let

$$\boldsymbol{\Delta} = \left[\frac{\partial \boldsymbol{\Sigma}}{\partial \boldsymbol{\theta}'}\right]_{\boldsymbol{\theta}_0} . \tag{38}$$

Then

$$NACov(\hat{\boldsymbol{\theta}}) = (\boldsymbol{\Delta}'\overline{\mathbf{V}}\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}'\overline{\mathbf{V}}\boldsymbol{\Omega}\overline{\mathbf{V}}\boldsymbol{\Delta}(\boldsymbol{\Delta}'\overline{\mathbf{V}}\boldsymbol{\Delta})^{-1} , \tag{39}$$

which can be estimated as

$$NEst[ACov(\hat{\boldsymbol{\theta}})] = (\hat{\boldsymbol{\Delta}}'\hat{\mathbf{V}}\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\mathbf{V}}\mathbf{W}\hat{\mathbf{V}}\hat{\boldsymbol{\Delta}}(\hat{\boldsymbol{\Delta}}'\hat{\mathbf{V}}\hat{\boldsymbol{\Delta}})^{-1} , \tag{40}$$

where $\hat{\boldsymbol{\Delta}}$ is $\boldsymbol{\Delta}$ evaluated at $\hat{\boldsymbol{\theta}}$. The standard errors reported by LISREL for each parameter are obtained from the diagonal elements of (40).

Two special cases are of particular interest:

- Under normality and with methods GLS or ML, $\hat{\mathbf{V}} = \mathbf{W}_{\mathrm{NT}}^{-1}$ and $\mathbf{W} = \mathbf{W}_{\mathrm{NT}}$, so that (40) reduces to

  $$NEst[ACov(\hat{\boldsymbol{\theta}}) = (\hat{\boldsymbol{\Delta}}'\hat{\mathbf{V}}\hat{\boldsymbol{\Delta}})^{-1} , \tag{41}$$

  which is the estimated Fisher Information Matrix.

- Under non-normality and with method WLS, $\hat{\mathbf{V}} = \mathbf{W}_{\mathrm{NNT}}^{-1}$ and $\mathbf{W} = \mathbf{W}_{\mathrm{NNT}}$, so that (40) reduces also to (41).

To test the model under multivariate normality, one can use

$$C_1 = N[\log \mid \hat{\boldsymbol{\Sigma}} \mid + \mathsf{tr}(\mathbf{S}\hat{\boldsymbol{\Sigma}}^{-1}) - \log \mid \mathbf{S} \mid - k] . \tag{42}$$

Although C_1 can be computed for any $\mathbf{V}$, *i.e.*, any fit function in (26)–(30), its natural use is with ML. Then C_1 is the likelihood ratio χ^2 statistic. Under multivariate normality C_1 has an asymptotic χ^2 distribution with

$$d = s - t \tag{43}$$

degrees of freedom if the model holds. Recall that $s = (1/2)k(k+1)$ and t is the number of independent parameters in the model.

The matrix $\mathbf{\Delta}$ is of order $s \times t$. If $t < s$, there exists an an orthogonal complement $\mathbf{\Delta}_c$ of order $s \times d$ to $\mathbf{\Delta}$ such that $\mathbf{\Delta}'\mathbf{\Delta}_c = \mathbf{0}$. Let $\hat{\mathbf{\Delta}}_c$ be $\mathbf{\Delta}_c$ evaluated at $\hat{\boldsymbol{\theta}}$. To test the model for any fit function, Browne (1984) developed a general formula:

$$C_2(\mathbf{W}) = N(\mathbf{s} - \hat{\boldsymbol{\sigma}})'[\hat{\mathbf{\Delta}}_c(\hat{\mathbf{\Delta}}_c'\mathbf{W}\hat{\mathbf{\Delta}}_c)^{-1}\hat{\mathbf{\Delta}}_c'](\mathbf{s} - \hat{\boldsymbol{\sigma}}) \tag{44}$$

This is Browne's ADF chi-square statistic (Browne, 1984, equation 2.20a). If $\mathbf{W}_{\text{NNT}}$ is available, LISREL 9 computes both $C_2(\mathbf{W}_{\text{NT}})$ and $C_2(\mathbf{W}_{\text{NNT}})$. Otherwise, LISREL 9 gives only $C_2(\mathbf{W}_{\text{NT}})$. In the first case, an advantage is that one can compare $C_2(\mathbf{W}_{\text{NNT}})$ and $C_2(\mathbf{W}_{\text{NT}})$. The difference can be interpreted as an effect of non-normality. $C_2(\mathbf{W}_{\text{NT}})$ is valid for all methods with $\hat{\mathbf{V}}$ defined in (31)–(34) under normality. $C_2(\mathbf{W}_{\text{NNT}})$ is valid for the same methods under non-normality. Under general assumptions $C_2(\mathbf{W}_{\text{NNT}})$ has an asymptotic χ^2 distribution with d degrees of freedom if the model holds. With the ML method, C_1 and $C_2(\mathbf{W}_{\text{NT}})$ are asymptotically equivalent under multivariate normality. It has been found in simulation studies that $C_2(\mathbf{W}_{\text{NNT}})$ does not work well under non-normality, see *e.g*, Curran, West, & Finch (1996). This is probably because the matrix $\mathbf{W}_{\text{NNT}}$ is often unstable unless the sample size N is huge

As a remedy for the poor behaviour of C_1 and $C_2(\mathbf{W}_{\text{NT}})$ under non-normality, Satorra & Bentler (1988) developed an alternative test procedure. Let

$$\mathbf{U} = \hat{\mathbf{V}} - \hat{\mathbf{V}}\hat{\mathbf{\Delta}}(\hat{\mathbf{\Delta}}'\mathbf{V}\hat{\mathbf{\Delta}})^{-1}\hat{\mathbf{\Delta}}'\hat{\mathbf{V}} \ . \tag{45}$$

By Khatri's (1966) lemma, $\mathbf{U}$ can also be written as

$$\mathbf{U} = \hat{\mathbf{\Delta}}_c(\hat{\mathbf{\Delta}}_c'\hat{\mathbf{V}}^{-1}\hat{\mathbf{\Delta}}_c)^{-1}\hat{\mathbf{\Delta}}_c' \tag{46}$$

Under non-normality the asymptotic distribution of C_1 and $C_2(\mathbf{W}_{\text{NT}})$ is not known but Satorra & Bentler (1988) invokes Theorem 2.1 of Box (1954) to conclude that C_1 and $C_2(\mathbf{W}_{\text{NT}})$ are asymptotically distributed as a linear combination of χ^2's with one degree of freedom where the coefficients of the linear combination are the non-zero eigenvalues of $\mathbf{UW}$. Based on this result Satorra & Bentler (1988) suggest a scale factor for C_1 and $C_2(\mathbf{W}_{\text{NT}})$ such that this statistic has the correct asymptotic mean d. Let

$$h_1 = \mathsf{tr}(\mathbf{U}\mathbf{W}_{\text{NNT}}) = \mathsf{tr}[(\hat{\mathbf{\Delta}}_c'\hat{\mathbf{V}}^{-1}\hat{\mathbf{\Delta}}_c)^{-1}(\hat{\mathbf{\Delta}}_c'\mathbf{W}_{\text{NNT}}\hat{\mathbf{\Delta}}_c)] \ . \tag{47}$$

Note that for GLS and ML, this is

$$h_1 = \mathsf{tr}(\mathbf{U}\mathbf{W}_{\text{NNT}}) = \mathsf{tr}[(\hat{\mathbf{\Delta}}_c'\mathbf{W}_{\text{NT}}\hat{\mathbf{\Delta}}_c)^{-1}(\hat{\mathbf{\Delta}}_c'\mathbf{W}_{\text{NNT}}\hat{\mathbf{\Delta}}_c)] \ , \tag{48}$$

but (47) is valid also for ULS and DWLS. The scale factor to be multiplied is d/h_1. In LISREL 9 we apply this scaling to C_1 if the ML method is used and define[12]

$$C_3 = (d/h_1)C_1 \ . \tag{49}$$

For all other methods we apply the scale factor to $C_2(\mathbf{W}_{\text{NT}})$ and define

$$C_3 = (d/h_1)C_2(\mathbf{W}_{\text{NT}}) \ . \tag{50}$$

[12]In previous versions of LISREL we applied this scaling to $C_2(\mathbf{W}_{\text{NT}})$ for all methods.

Although C_3 does not have an asymptotic χ^2 distribution under non-normality, it is often used as an approximate chi-square. C_3 is called the Satorra-Bentler scaled chi-square or the Satorra-Bentler mean adjusted chi-square. It has been found to work well in practice as it outperforms $C_2(\mathbf{W}_{\mathrm{NNT}})$ under non-normality, see *e.g*, Curran, West, & Finch (1996).

To get an even better approximation to a χ^2 distribution Satorra & Bentler (1988) suggest a mean and variance adjusted statistic. This is a Satterthwaite (1941) type of adjustment. Let

$$h_2 = \mathsf{tr}(\mathbf{U}\mathbf{W}_{\mathrm{NNT}}\mathbf{U}\mathbf{W}_{\mathrm{NNT}}) \ . \tag{51}$$

Using (46), h_2 may be written

$$h_2 = \mathsf{tr}[(\hat{\mathbf{\Delta}}_c'\mathbf{V}^{-1}\hat{\mathbf{\Delta}}_c)^{-1}(\hat{\mathbf{\Delta}}_c'\mathbf{W}_{\mathrm{NNT}}\hat{\mathbf{\Delta}}_c)(\hat{\mathbf{\Delta}}_c'\mathbf{V}^{-1}\hat{\mathbf{\Delta}}_c)^{-1}(\hat{\mathbf{\Delta}}_c'\mathbf{W}_{\mathrm{NNT}}\hat{\mathbf{\Delta}}_c)] \ , \tag{52}$$

so that only matrices of order $d \times d$ are involved. Let

$$d' = h_1^2/h_2 \ , \tag{53}$$

and define

$$C_4 = (d'/h_1)C = (h_1/h_2)C \ , \tag{54}$$

where $C = C_1$ for ML and $C = C_2(\mathbf{W}_{\mathrm{NT}})$ for all other methods. Regard C_4 as a an approximate chi-square with d' degrees of freedom. In LISREL 9 we give d' with three decimals and use this fractional degrees of freedom to compute the P-value for C_4.

6 References

Aish, A.M., & Jöreskog, K.G. (1990) A panel model for political efficacy and responsiveness: An application of LISREL 7 with weighted least squares. *Quality and Quantity*, 24,405-426.

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Uman Decision Processes*, 50, 179–211.

Akaike, H. (1974) A new look at statistical model identification. *IEEE transactions on Automatic Control*, 19, 716–723.

Barnes S.H., & Kaase M.(Eds.) (1979). *Political action: Mass participation in five western democracies.* Beverly Hills and London: Sage Publications.

Bartholomew, D.J., Steele, F., Moustaki, I., & Galbraith, J.I. (2002) *The analysis and interpretation og multivariate data for social scientists*. Chapman & Hall.

Bollen, K.A. & Curran, P.J. (2006) *Latent Curve Models: A Structural Equation Approach*. Wiley.

Browne, M.W. (1977) Generalized least squares estimators in the analysis of covariance structures. Pp 205-226 in D.J. Aigner & A.S. Goldberger (Eds.): *Latent variables in socio-economic models*. Amsterdam: North-Holland.

Browne, M.W. (1984) Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 62–83.

Browne, M. W. & Cudeck, R. (1993) Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Editors): *Testing Structural Equation Models*, Sage Publications.

Campbell, A., Gurin, G., & Miller, W.A. (1954). *The voter decides.* New York: Wiley.

Craig, S.C., and Maggiotto, M.A. (1981) Measuring political efficacy. *Political Methodology, 8,* 85-109.

Cudeck, R. (1989) Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, 105, 317–327.

Curran, P.J., West S.G. & Finch, J.F. (1996) The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, **1**, 16–29.

Forero, C.G., Maydeu-Olivares A., & Gallardo-Pujol, D. Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling*, 17, 392–423.

Foss, T, Jöreskog, K.G., & Olsson, U.H. (2011) Testing structural equation models: The effect of kurtosis. *Computational Statistics and Data Analysis*, **55**, 2263–2275.

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

Holzinger, K., & Swineford, F. (1939) *A study in factor analysis: The stability of a bifactor solution.* Supplementary Educational Monograph no. 48. Chicago: University of Chicago Press.

Hu, L., Bentler, P.M., & Kano, Y. (1992) Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin,* Vol. **112**, 351–362.

Jöreskog, K.G. (1990) New developments in LISREL: analysis of ordinal variables using polychoric correlations and weighted least squares, *Quality and Quantity*, 24, 4, 387-404.

Jöreskog, K. G., & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*, **36**, 347–387.

Jöreskog, K. G., & Moustaki, I. (2006). Factor analysis of ordinal variables with full information maximum likelihood. *www.ssicentral.com/lisrel/techdocs.orfiml.pdf*.

Jöreskog, K.G. & Sörbom, D. (2003a) LISREL*: Structural Equation Modeling with the SIMPLIS Command Language*. Lincolnwood: Scientific Software International.

Jöreskog, K.G. & Sörbom, D. (2003b) LISREL *User's Reference Guide*. Lincolnwood: Scientific Software International.

Jöreskog, K.G., Sörbom, D., Du Toit, S., & Du Toit, M. (2003) LISREL*: New Statistical Features*. Third printing with revisions. Chicago: Scientific Software International.

Kaiser, H.F. (1958) The varimax criterion for analytical rotation in factor analysis. *Psychometrika*, **23**, 187–200.

Satorra, A. (1989) Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika*, **54**, 131–151.

Satorra, A., & Bentler, P.M. (1988) Scaling corrections for chi-square statistics in covariance structure analysis. Proceedings of the Business and Economic Statistics Section of the American Statistical Association, 308–313.

Satterthwaite, F.E. (1941) Synthesis of variance. *Psychometrika*, 6, 309–316.

Schafer, J. L. (1997) *Analysis of incomplete data*. Chapman & Hall/CRC.

Schwarz, G.E. (1978) Estimating the dimension of a model. *Annals if Statistics*, **6**, 461–464.

Yang-Wallentin, F, Jöreskog, K.G., & Luo, H. Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling*, 17, 392–423.