Analysis of ordinal variables: panel model for political efficacy

In many cases, especially when data are collected through interviews or questionnaires, the observed variables are ordinal, i.e., responses are classified into different ordered categories.

An ordinal *z* (*z* may be either a *y*- or *x*-variable in the LISREL sense may be regarded as a crude measurement of an underlying unobserved or unobservable continuous variable z^*. For example, a four-point ordinal scale may be conceived as:

- $z^* \le \tau_1$, z is scored 1
- $\tau_1 \le z^* \le \tau_2$, z is scored 2
- $\tau_2 \le z^* \le \tau_3$, z is scored 3
- $\tau_3 \le z^*$, z is scored 4

Where $\tau_1 \le \tau_2 \le \tau_3$ are *threshold* values for z^*. It is often assumed that z^* has a standard normal distribution, in which case the thresholds can be estimated from the inverse of the normal distribution function.

Suppose z_1 and z_2 are two ordinal variables with underlying continuous variables z_1^* and z_2^* respectively. Assuming that z_1^* and z_2^* have a bivariate normal distribution, their correlation is called the *polychoric correlation coefficient*. A special case of this is the *tetrachoric correlation coefficient* when both z_1^* and z_2^* are dichotomous. Now suppose further that z_3 is a continuous variable measured on an interval scale. The correlation between z_1^* and z_3 is called the *polyserial correlation coefficient* assuming that z_1^* and z_3 have a bivariate normal distribution. A special case of this is the *biserial correlation* when z_1 is dichotomous.

An ordinal variable z does not have a metric scale. To use such variable in a linear relationship we use the corresponding underlying variable z^* instead. The polychoric and polyserial correlations are not correlations computed from actual scores but are rather theoretical correlations of the underlying z^* variables. These correlations are estimated from the observed pairwise contingency tables of the ordinal variables. See Jöreskog & Sörbom (1988), Jöreskog (1990), and references therein for the theory on which the polychoric and polyserial correlations are based. Jöreskog & Aish (1994) give further examples of analysis with ordinal variables.

When the observed variables in LISREL are all ordinal or are of mixed scale types (ordinal and interval), the use or ordinary product-moment correlations based on raw scores is not recommended. Instead it is suggested that estimates of polychoric and polyserial correlations be computed and that the matrix of such correlations be analyzed by the WLS method.

The weight matrix required for such an analysis is the inverse of the estimated asymptotic covariance matrix $\mathbf{W}$ of the polychoric and polyserial correlations. The asymptotic covariance matrix as well as the matrix of polychoric correlations are obtained by PRELIS.

The steps involved in this analysis will be described in the following example. This example involves ordinal variables and polychoric correlations only. For other examples involving also continuous and/or censored variables and other types of correlations, see the PRELIS examples.

Aish & Jöreskog (1990) analyze data on political attitudes. Their data consist of 16 ordinal variables measured on the same people at two occasions. Six of the 16 variables were considered to be indicators of Political Efficacy and System Responsiveness. The attitude questions corresponding to these six variables are:

- People like me have no say in what the government does (NOSAY)
- Voting is the only way that people like me can have any say about how the government runs things (VOTING)
- Sometimes politics and government seem so complicated that a person like me cannot really understand what is going on (COMPLEX)
- I don't think that public officials care much about what people like me think (NOCARE)
- Generally speaking, those we elect to Parliament lose touch with the people pretty quickly (TOUCH)
- Parties are only interested in people's votes but not in their opinions (INTEREST)

Permitted responses to these questions were agree strongly, agree, disagree, disagree strongly, don't know and no answer.

Aish & Jöreskog (1990) considered many different models. Only one of them will be used here. The model considered here includes the idea of correlated measurement errors. Aish & Jöreskog (1990) split the raw data into two random subsamples, an exploration sample of size 410 and a confirmation (validation) sample of 395. Details are given in Aish & Jöreskog (1990).

This model involves two kinds of correlated error terms. The structural residuals are correlated, indicated by the small two-way arrows in the model. The measurement errors in VOTING and COMPLEX are

supposed to be correlated over time, indicated by the two long two-way arrows. The errors in VOTING and COMPLEX are said to be *autocorrelated* and the two-way arrows represent the *autocovariance*.

Jöreskog and Sörbom (1989) illustrate how to use PRELIS to obtain the matrix of polychoric correlations for the exploration sample and how to estimate the asymptotic covariance matrix of the polychoric correlation from the total sample. For our purposes, we assume that the matrix of polychoric correlations, estimated from the exploration sample, has been saved into the file **PANELUSA.PME** and that the corresponding asymptotic covariance matrix, estimated from the total sample, has been saved in the file **PANELUSA.ACP**. The names of the variables are in the file **PANELUSA.LAB**. These files can be found in the **SIMPLIS examples** folder.

First we run the model without correlated measurement errors. The input file is **EX9A.SPL**:

```
Panel Model for Political Efficacy
See Aish and Joreskog, Quality and Quantity (1990)
Observed Variables from File PANEL.LAB
Correlation Matrix from File PANELUSA.PME
Asymptotic Covariance Matrix from File PANELUSA.ACP
Sample size: 410
Latent Variables: Efficac1 Respons1 Efficac2 Respons2
Relationships:
   NOSAY1 - NOCARE1 = Efficac1
   NOCARE1 - INTERES1 = Respons1

   NOSAY2 - NOCARE2 = Efficac2
   NOCARE2 - INTERES2 = Respons2

   Efficac2 = Efficac1
   Respons2 = Respons1

Let the Errors for Efficac2 and Respons2 correlate
Method of Estimation: Weighted Least Squares
Path Diagram
End of Problem
```

The new element here is the line

```
Asymptotic Covariance Matrix from File PANELUSA.ACP
```

The file PANELUSA.ACP produced by PRELIS is necessary for the analysis of polychoric correlations. The line above tells LISREL to read the asymptotic covariance matrix from this file and to use WLS (weighted Least Squares) to estimate the parameters of the model. If no asymptotic covariance matrix is read, ML will use ML (Maximum Likelihood) by default to estimate the parameters, as in all previous examples. Using ML to fit the model to a matrix of polychoric correlations violates the statistical theory on which the estimation of chi-square and standard errors are based.

Since the observed variables are ordinal and have no units of measurement, it is not possible to use any of these as reference variables to assign units of measurement to the latent variables. The most reasonable thing to do is to assume that all variables, observed as well as latent, are standardized.

The first two lines of relationships define the measurement model at time 1, and the next two lines define a similar measurement model at time 2. The following two lines of relationships define the structural relationships among the latent variables. Finally, the last line of relationships specifies that the two structural disturbance terms should be correlated as indicated by the small two-way arrow in the middle of the path diagram.

The output file reveals that the model does not fit the data. Chi-square is

```
Degrees of Freedom                                    48
 Weighted Least Squares Chi-Square (C1)                114.104 (P = 0.0000
```

The *P*-value for this is 0.00000019. If the model is "true", and all assumptions of the analysis hold, the odds of obtaining such a chi-square are less than 3 in 1,000,000.

The output file gives the following table of modification indices for error covariances:

```
The Modification Indices Suggest to Add an Error Covariance
  Between    and     Decrease in Chi-Square    New Estimate
 NOSAY1    VOTING2            13.1                -0.15
 VOTING1   VOTING2            17.9                 0.23
 VOTING1   NOSAY1              9.4                 0.16
 COMPLEX1  COMPLEX2           43.9                 0.36
```

The two largest modification indices are for the covariance of the measurement errors in VOTING1 and VOTING2 and the covariance for the measurement error in COMPLEX1 and COMPLEX2. This suggest that there are large specific factors in VOTING and COMPLEX.

In the next run, we therefore specify the model with the two sets of correlated error terms. This is done by adding the two lines

```
Let the errors of VOTING1 and VOTING2 correlate
Let the errors of COMPLEX1 and COMPLEX2 correlate
```

in the input file (see **EX9B.SPL**).

```
Degrees of Freedom                                    46
 Weighted Least Squares Chi-Square (C1)                50.640 (P = 0.2955)
```

The chi-square for this model is 50.52 with 46 degrees of freedom. Hence, the model fits the data well. The autocovariances in VOTING and COMPLEX are given in the output as

```
Error Covariance for VOTING1 and VOTING2 = 0.234
                                          (0.0514)
                                            4.545

 Error Covariance for COMPLEX1 and COMPLEX2 = 0.353
                                            (0.0504)
                                              7.010
```

Both autocovariances are highly significant. Hence, there is strong evidence that there are large specific factors in VOTING and COMPLEX. The results of model 9B are summarized in the tables that follows. For further analysis of this data and interpretation of the results, see Jöreskog & Aish (1994).

Table 1: Loadings and their standard errors

	Time 1		Time 2	
	Efficacy	*Respons*	*Efficacy*	*Respons*
NOSAY	0.85 (0.05)		0.79 (0.08)	
VOTING	0.58 (0.06)		0.71 (0.08)	
COMPLEX	0.56 (0.05)		0.47 (0.06)	
NOCARE	0.37 (0.11)	0.50 (0.11)	0.42 (0.12)	0.49 (0.12)
TOUCH		0.80 (0.04)		0.85 (0.05)
INTEREST		0.92 (0.03)		0.89 (0.06)

Table 2: Error variances, reliabilities and autocovariances

	Error variances		*Reliabilities*		*Autocovariances*
	Time 1	*Time 2*	*Time 1*	*Time 2*	*Specific Variances*
NOSAY	0.28	0.37	0.72	0.63	
VOTING	0.43	0.26	0.57	0.74	0.23 (.05)
COMPLEX	0.34	0.43	0.66	0.57	0.35 (.05)
NOCARE	0.32	0.27	0.68	0.73	
TOUCH	0.36	0.29	0.64	0.71	
INTEREST	0.15	0.20	0.85	0.80	

Table 3: Correlations between Efficacy and Responsiveness

Time 1	*Time 2*
0.78	0.81

Table 2: Stability coefficients and residual covariance matrix

	Stability Coefficients	Residual Covariance Matrix	
Efficacy	0.72 (0.09)	.48	
Respons	0.64 (0.07)	0.45	0.59