

Structural Equation Modeling with Ordinal Variables using LISREL

Karl G Jöreskog

26 April 2002
Revised 14 April 2004

This document was originally published as a series of five notes on the analysis of ordinal variables at the website of Scientific Software International, Inc. (SSI), in *Karl's Corner*:

<http://www.ssicentral.com/lisrel/corner.htm>

The first note appeared in August 2001, shortly after the release of LISREL 8.50, a major improvement of LISREL 8.3, the previous version of the program. The fifth and final note appeared in June 2002. The development of this material gave rise to further improvements in the program, resulting in the release of LISREL 8.51 in October 2001 and the release of LISREL 8.52 in June 2002.

With one exception, all the examples discussed in this document can be run with the student edition of LISREL 8.52, which is available for free download at SSI's website: <http://www.ssicentral.com>.

After successful installation, both the full edition and the student edition of LISREL 8.52 feature a folder (ORDINAL) with all the files needed to run the examples. Those example files as well as this document (in PDF format) can be downloaded from SSI's website.

Contents

1	Preliminary Analysis	1
1.1	The Political Action Survey	1
1.2	Data Screening	2
1.3	Missing Values	3
1.3.1	Listwise Deletion	4
1.3.2	Pairwise Deletion	5
1.3.3	Imputation	6
1.4	Using Interactive LISREL	8
1.4.1	Get Yourself a PSF File	9
1.4.2	Imputation and Data Screening	9
2	Cross-Sectional Data	10
2.1	Ordinal Variables and Underlying Variables	10
2.2	Polychoric Correlations	11
2.3	Testing Underlying Bivariate Normality	13
2.4	Estimating the Polychoric Correlations and their Asymptotic Covariance Matrix	14
2.5	Fitting and Testing the Model	21
2.6	How to Become a Chi-square Collector	22
2.7	Alternative Parameterization	23
3	Longitudinal Data	26
3.1	Equal Thresholds	27
3.2	A Two Variables Example of Equal Thresholds	28
3.3	Estimating the Mean Vector and Covariance Matrix under Equal Thresholds	29
3.4	A Panel Model for Efficacy	31
3.4.1	Input	32
3.4.2	Output	34
3.4.3	Testing Sequence	36
3.4.4	Testing Equality of Variances	37
3.4.5	Error Variances and Reliabilities	38
3.4.6	LISREL Notation	38
3.4.7	LISREL Syntax	39
3.5	Four-Wave Models	40
3.5.1	A Four-Wave Model with a Single Latent Variable	40
3.5.2	A Four-Wave Model with Two Latent Variables	44
4	Multiple Groups	45
4.1	Data Screening	46
4.2	PRELIS Step	49
4.3	LISREL Step	50
4.4	Conclusion	57

5 Covariates	57
5.1 Univariate Probit Regression	58
5.2 Univariate Logit Regression	61
5.3 Testing the Model	62
5.4 Bivariate Probit Regression	62
5.5 Multivariate Probit Regression	64
5.6 PRELIS Implementation	65
5.7 A Small Example	65
5.8 Data Screening	67
5.9 Probit Regression of NOSAY	68
5.10 Probit and Logit Regression of All Efficacy Variables	70
5.11 Estimating the Joint Covariance Matrix	72
5.12 A MIMIC Model for Efficacy and Respons	73
References	75
Appendix 1: Derivation of the RMSEA Measure	76
Appendix 2: Questions and Answers	77

1 Preliminary Analysis

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 each cell of a multiway contingency table. To use ordinal variables in structural equation models requires other techniques than those that are traditionally employed with continuous variables.

In this paper I will illustrate how one can analyze ordinal variables with PRELIS and LISREL. Some of the features I will describe have been available for some time but have not been properly explained and illustrated. Other features are new in PRELIS 2.5 and LISREL 8.5.

1.1 The Political Action Survey

To illustrate the data analysis in this paper I use 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 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.¹

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.

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:²

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

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 I 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

¹The 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.

²These 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

- 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.

1.2 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.

PRELIS 2 automatically does such a data screening by determining for each variable the distinct data values present in the data and the number of each. If a variable has more than 15 distinct values, PRELIS 2 will group them in intervals and determine the number in each interval.

The data on the efficacy variables for the USA are available in the file `EFFICACY.RAW` in the `lis850ex` subdirectory. The data values have been recorded with one space before each number, *i.e.*, with two columns per variable.

There are two ways of running LISREL 8.5:

- From syntax files (in this case PRELIS syntax files)
- Using the point and click Windows interface as described in the new *Interactive LISREL: User's Guide* (Du Toit & Du Toit, 2001)

In the following I illustrate the examples using syntax files. In Section 1.4 I describe briefly how one can do the same thing using Interactive LISREL.

A simple PRELIS input file for screening the data is as follows (`ORD11.PR2`):

```
EFFICACY: PRELIS Run 1
!Data Screening of Political Action Data for the USA:
!Cross-Section Data - Variables: 136 - 141 (Political Efficacy Variables)

Data Ninputvariables = 6
Labels
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata=EFFICACY.RAW
Output
```

Note that the number of cases (records of data) need not be specified; PRELIS determines the sample size, all distinct data values for each variable and the absolute and relative frequency of occurrence of each value. PRELIS also gives a bar chart showing the percentage of each data value. The output file shows that PRELIS has correctly determined

- that there are 1719 cases in the data,
- that there are six distinct values on each variable, coded 1, 2, 3, 4, 8, 9,

Table 1: 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

- the distribution of these values.

The results are presented in compact form in Table 1.

These results agree exactly with those given in the Political Action Codebook, pp. 174–180. Thus, we have a solid base to continue our analysis. If something had been wrong in the data, it would have been detected by this kind of data screening.

In the Political Action Survey the responses *Agree Strongly*, *Agree*, *Disagree*, *Disagree Strongly*, *Don't Know*, *No Answer* to the political efficacy items were coded 1, 2, 3, 4, 8, 9, respectively. One can assign category labels to category codes by including the following line in the input file (see file ORD11.PR2):

```
CLabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS 8=DK 9=NA
```

1.3 Missing Values

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*. The usual way to deal with such responses is to declare them as missing values and include some treatment of missing values in the analysis. A lengthy discussion of this issue is beyond the scope of this paper. LISREL 8.50 provides *multiple imputation* for continuous normally distributed variables with data missing at random (MAR), see pp. 165–170 and 387–388 in the *Interactive LISREL: User's Guide* (Du Toit & Du Toit, 2001). One can also estimate a LISREL model directly from raw data with full information maximum likelihood (FIML) under multivariate normality and MAR, see pp. 234–250 and 388–389 in the *Interactive LISREL: User's Guide* (Du Toit & Du Toit, 2001). Although very powerful under the assumptions made, these procedures should not be used routinely to solve missing data problems. They cannot (should not) be used with categorical variables.

One should try to take the mechanism that generates the missing data into account. Why are data missing on a particular variable? Does the probability of a missing value on this variable depend on the values of the variable itself? If so, MAR does not hold. Does the probability of a missing value on this variable depend on other variables? If so, MAR may hold. But if one can find other variables that can be used to predict the missing values these can be taken into account.

One general procedure to do this is the matching procedure described in Section 1.3.3. This can be used with ordinal or continuous variables with any distribution and does not require MAR. For continuous variables and data missing completely at random (MCAR), this procedure, also called *similar response pattern imputation*, was evaluated by Brown (1994) and was found to work well as compared with several other procedures, including listwise and pairwise deletion.

In the following, I consider these three procedures in turn. I do not recommend that listwise and pairwise deletion be used to solve the missing data problem unless you have MCAR, and even with MCAR, I do not recommend computing covariance or correlation matrices with pairwise deletion. *I include the listwise and pairwise procedures here as descriptive devices as they provide further insight*

into the missing data problem. Computation of covariance and correlation matrices will be considered in Section 2.

1.3.1 Listwise Deletion

Missing values do not seem to be a serious problem in this example. As seen in Table 1, the percentage of *Don't Know* answers varies from 0.5% for COMPLEX to 3.6% for INTEREST. More people answer *Don't Know* for TOUCH and INTEREST than for the other items. The percentage of *No Answer* responses varies from 0.6% for NOSAY and VOTING to 1.1% for INTEREST. Listwise deletion means that all cases with *Don't Know* and *No Answer* responses will be excluded in the analysis. This is illustrated in the following input file (ORD12.PR2):

```
EFFICACY: PRELIS Run 2
!Data Screening of Political Action Data for the USA:
!Cross-Section Data - Variables: 136 - 141 (Political Efficacy Variables)
!Listwise Deletion of Missing Values
Data Ninputvariables = 6 Missing = 8,9
Labels
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata=EFFICACY.RAW
CLabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS
Output
```

This gives the following results.

Number of Missing Values per Variable

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

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

The distribution of missing values over variables are given first. 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.

Further down in the output is the distribution of missing values over cases. 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 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. It is seen that most people answer either *agree* or *disagree*. Fewer people answer with the *stronger* alternatives.

1.3.2 Pairwise Deletion

A more comprehensive data screening can be done by pairwise deletion. To do so, add the specification `Treatment = Pairwise` to the `Data` line and put `MP` (for *missing patterns*) on the `Output` line (see file `ORD12.PR2`).

This gives the following results:

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	Pattern
1554	0 0 0 0 0 0
16	1 0 0 0 0 0
12	0 1 0 0 0 0
1	1 1 0 0 0 0
4	0 0 1 0 0 0
11	0 0 0 1 0 0
31	0 0 0 0 1 0
1	0 1 0 0 1 0
2	1 1 0 0 1 0
1	0 1 1 0 1 0

```

4  0 0 0 1 1 0
1  0 0 1 1 1 0
32 0 0 0 0 0 1
1  0 1 0 0 0 1
1  1 1 0 0 0 1
1  1 0 1 0 0 1
5  0 0 0 1 0 1
2  1 0 0 1 0 1
1  0 0 1 1 0 1
1  1 0 1 1 0 1
14 0 0 0 0 1 1
4  1 0 0 0 1 1
4  0 1 0 0 1 1
2  1 1 0 0 1 1
1  0 1 1 0 1 1
1  0 0 0 1 1 1
2  0 1 1 1 1 1
9  1 1 1 1 1 1

```

Thus, there are 1554 cases with no missing data, there are 16 cases with missing values on variable 1 only, and 1 case with missing values on both variable 1 and 2, 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.

1.3.3 Imputation

Another way to deal with the problem of missing values is by imputation, *i.e.*, by substituting a real scale value 1, 2, 3, or 4 for the missing values 8 and 9. PRELIS has a procedure for imputing missing values on a variable by matching on other variables. This procedure is based on the idea that if person *a* has a missing value on variable *i* and has the same response pattern as person *b* on a set of matching variables, it is likely that he/she should have the same value on variable *i* as person *b*. Therefore, *b*'s value on variable *i* is substituted for *a*'s missing value on variable *i*. If there are several persons with the same response patterns on the matching variables and with the same values on variable *i*, there is an even stronger case for substituting this value for *a*'s value on variable *i*. For further details, see the PRELIS 2 User's Reference Guide pp. 155–160.

As most of the missing values are on the variables TOUCH and INTEREST, one idea to increase the listwise sample size is to impute missing values on TOUCH and INTEREST by matching on the other variables. This is merely an illustration. To do imputation the following PRELIS input file can be used (see file ORD13.PR2):

```

EFFICACY: PRELIS Run 3
Cross-Section Data - Variables: 136 - 141 (Political Efficacy Variables)
Imputation of Missing Values on TOUCH and INTEREST
Listwise Deletion of Missing Values after Imputation
Saving Imputed Data in EFFICACY.IMP
Data Ninputvariables = 6 Missing = 8,9
Labels
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata=EFFICACY.RAW
Impute (TOUCH INTEREST) (NOSAY - NOCARE) XN

```

CLabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS
 Output Rawdata = EFFICACY.IMP Width = 2 Ndecimals = 0

Missing values on TOUCH and INTEREST are only imputed if matching cases are found. Therefore, there may still be missing values on TOUCH and INTEREST after imputation. Note that PRELIS will only impute "legal" values, *i.e.*, values 1, 2, 3, and 4, not values such as 2.5 or 3.4. All the missing values on NOSAY - NOCARE are still there after imputation. In this example, all missing values remaining after imputation are eliminated by listwise deletion and the listwise sample is saved in the file EFFICACY.IMP. The specification Width = 2 Ndecimals = 0 means that the data will be written in two-column fields without any decimals. Hence, the file EFFICACY.IMP can be read in free format. This file has no missing values in it.

The output lists all imputed cases as follows:

Imputations for TOUCH			
Case 28	imputed with value	2 (Variance Ratio = 0.000),	NM= 1
Case 29	imputed with value	2 (Variance Ratio = 0.000),	NM= 1
Case 238	imputed with value	2 (Variance Ratio = 0.454),	NM= 10
Case 336	imputed with value	1 (Variance Ratio = 0.277),	NM= 26
Case 418	imputed with value	2 (Variance Ratio = 0.380),	NM= 76
Case 530	imputed with value	2 (Variance Ratio = 0.360),	NM= 143
Case 578	imputed with value	3 (Variance Ratio = 0.000),	NM= 1
Case 600	imputed with value	2 (Variance Ratio = 0.405),	NM= 31
Case 604	imputed with value	2 (Variance Ratio = 0.493),	NM= 23
Case 684	imputed with value	2 (Variance Ratio = 0.396),	NM= 32
Case 951	imputed with value	1 (Variance Ratio = 0.000),	NM= 1
Case 963	imputed with value	2 (Variance Ratio = 0.358),	NM= 144
Case 985	imputed with value	2 (Variance Ratio = 0.463),	NM= 28
Case 1189	imputed with value	2 (Variance Ratio = 0.356),	NM= 145
Case 1268	imputed with value	2 (Variance Ratio = 0.410),	NM= 5
Case 1311	imputed with value	2 (Variance Ratio = 0.410),	NM= 5
Case 1541	imputed with value	1 (Variance Ratio = 0.269),	NM= 27
Case 1578	imputed with value	2 (Variance Ratio = 0.399),	NM= 9
Case 1584	imputed with value	2 (Variance Ratio = 0.389),	NM= 33
Case 1704	imputed with value	2 (Variance Ratio = 0.476),	NM= 24

Imputations for INTEREST			
Case 12	imputed with value	1 (Variance Ratio = 0.292),	NM= 31
Case 28	imputed with value	1 (Variance Ratio = 0.000),	NM= 1
Case 48	imputed with value	2 (Variance Ratio = 0.337),	NM= 138
Case 67	imputed with value	3 (Variance Ratio = 0.309),	NM= 6
Case 332	imputed with value	2 (Variance Ratio = 0.335),	NM= 139
Case 418	imputed with value	2 (Variance Ratio = 0.429),	NM= 73
Case 469	imputed with value	2 (Variance Ratio = 0.333),	NM= 140
Case 510	imputed with value	2 (Variance Ratio = 0.287),	NM= 17
Case 530	imputed with value	2 (Variance Ratio = 0.331),	NM= 141
Case 578	imputed with value	3 (Variance Ratio = 0.000),	NM= 1
Case 598	imputed with value	2 (Variance Ratio = 0.329),	NM= 142
Case 653	imputed with value	2 (Variance Ratio = 0.424),	NM= 74
Case 680	imputed with value	3 (Variance Ratio = 0.375),	NM= 84
Case 796	imputed with value	3 (Variance Ratio = 0.465),	NM= 4
Case 951	imputed with value	1 (Variance Ratio = 0.000),	NM= 1
Case 1156	imputed with value	2 (Variance Ratio = 0.419),	NM= 75
Case 1227	imputed with value	2 (Variance Ratio = 0.327),	NM= 143
Case 1232	imputed with value	2 (Variance Ratio = 0.325),	NM= 144

```

Case 1249 imputed with value      3 (Variance Ratio = 0.435), NM= 27
Case 1368 imputed with value      2 (Variance Ratio = 0.414), NM= 76
Case 1477 imputed with value      3 (Variance Ratio = 0.266), NM= 7
Case 1516 imputed with value      2 (Variance Ratio = 0.323), NM= 145
Case 1556 imputed with value      3 (Variance Ratio = 0.497), NM= 6
Case 1704 imputed with value      2 (Variance Ratio = 0.240), NM= 24

```

Here NM is the number of matching cases and the Variance Ratio is the ratio between the variance of variable i for the matching cases and the total variance of variable i for all cases without missing values.

After imputation, the distribution of missing values over variables is:

Number of Missing Values per Variable After Imputation

```

      NOSAY      VOTING      COMPLEX      NOCARE      TOUCH      INTEREST
      39          37          21          37          57          57

```

Compared with the previous output we have gained 20 cases on TOUCH and 24 cases on INTEREST. These are the 44 cases just listed.

The distribution of missing values over cases is now:

Distribution of Missing Values

Total Sample Size = 1719

```

Number of Missing Values    0    1    2    3    4    5    6
      Number of Cases 1589   80   17   18   4    2    9

```

Listwise Deletion

Total Effective Sample Size = 1589

The listwise sample size after imputation is 1589. This was 1554 in the previous output. Altogether we have gained 35 cases with complete data. Of these 35 cases, 26 had 1 missing value and 9 had 2 missing values. All these missing values were imputed.

The analysis that follows in Section 2 will be based on the data on these 1589 cases stored in the file EFFICACY.IMP. The univariate distributions of the six political efficacy items estimated from these data are given in Table 2.

Table 2: Univariate Marginal Distributions

	Frequency				Percentage			
	AS	A	D	DS	AS	A	D	DS
NOSAY	163	492	811	123	10.3	31.0	51.0	7.7
VOTING	275	658	580	76	17.3	41.4	36.5	4.8
COMPLEX	314	905	309	61	19.8	57.0	19.4	3.8
NOCARE	241	656	636	56	15.2	41.3	40.0	3.5
TOUCH	261	857	446	25	16.4	53.9	28.1	1.6
INTEREST	260	740	560	29	16.4	46.6	35.2	1.8

1.4 Using Interactive LISREL

Here I describe briefly how to do data screening and imputation using Interactive LISREL. For further information see the *Interactive LISREL: User's Guide* (Du Toit & Du Toit, 2001).

1.4.1 Get Yourself a PSF File

To use Interactive LISREL, the first order of business is to get a PRELIS system file (PSF file) for the raw data in file EFFICACY.RAW. To do this, go through the following steps:

1. Select **Import Data in Free Format** in the **File Menu**.
2. Go to the `lis850ex` subdirectory where the file EFFICACY.RAW is and choose **Free Format Data (*.raw)** under **Files of Type**. Open EFFICACY.RAW.
3. In the **Enter Number of Variables** Dialog Box, specify 6 variables and click OK.
4. LISREL then shows the data in a spread sheet. This is the way the PSF file is displayed.
5. To complete the specification of the PSF file, go to the **Data Menu**.
 - (a) Select **Define Variables** and type the names of the variables: NOSAY, VOTING, etc.
 - (b) Still in the **Define Variables** Dialog Box, select **Variables Type**. Select all variables and define them as 'Ordinal.' Click OK.
 - (c) Still in the **Define Variables** Dialog Box, select **Category Labels**. Select all variables and define 1 = AS, 2 = A, etc. Click OK.
 - (d) Still in the **Define Variables** Dialog Box, select **Missing Values**. Specify 8 and 9 as global missing values. As deletion method choose listwise or pairwise. Click OK.
 - (e) Still in the **Define Variables** Dialog Box, click OK.

1.4.2 Imputation and Data Screening

Once the PSF has been defined, it can be displayed by selecting **Open** in the **File Menu** and choosing **PRELIS Data (*.psf)** under **Files of Type**. Then double clicking the filename will display the PSF file in spreadsheet form.

To do imputation and data screening follow the following steps:

1. With the PSF file displayed, select the **Statistics Menu** and then **Impute Missing Values**.
2. In the **Impute Missing Values** dialog box, select TOUCH and INTEREST and click the first **Add** button. This defines the variables to be imputed. Then select NOSAY – NOCARE and click the third **Add** button. This defines the matching variables. Select **List only successful imputations** and then click the **Output Options** button.
3. To save the imputed data as a PSF file, select the **Data** Dialog Box within the **Output** Dialog Box and check **Save the transformed data to file**. Then type a file name with suffix PSF, for example EFFICACY.PSF. Then click OK.
4. You will be returned to the **Impute Missing Values** dialog box. To generate a PRELIS syntax file click the **Syntax** button. Otherwise, click the **Run** button.

A much easier way to generate the file EFFICACY.PSF after imputation is to add the specification `RA=EFFICACY.PSF` on the **Output** line in PRELIS Run 3, see file ORD13A.PR2.

The generated file EFFICACY.PSF corresponds exactly to the file EFFICACY.IMP and the results in the output file should be the same as for the second PRELIS run described in Section 1.3.3. As will be illustrated in the next section, the file EFFICACY.PSF can be used for further analysis using Interactive LISREL.

2 Cross-Sectional Data

In Section 1 I introduced a small data set of six variables from the cross-sectional USA sample of the Political Action Survey. After imputation of some missing values and listwise deletion after imputation I ended up with 1589 cases with complete data on all six variables. See page 1 for a description of the variables.

Political scientists assume that these items measure a uni-dimensional trait called *political efficacy* which has been defined as *the feeling that individual political action does have, or can have, an impact upon the political process* (Campbell, *et al.*, 1954). In this section, I will demonstrate how the assumption of uni-dimensionality can be tested. I will also describe the statistical model used in PRELIS and LISREL for this purpose and discuss the assumptions of this model.

2.1 Ordinal Variables and Underlying Variables

Observations on an ordinal variable represent responses to a set of ordered categories, such as a four-category Likert scale. It is only assumed that a person who selected a specific 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 each cell of a multiway contingency table. To use ordinal variables in structural equation models requires other techniques than those that are traditionally employed with continuous variables.

For each ordinal variable z (which may be a y - or a x -variable in LISREL sense), it is assumed that there is an *underlying* continuous variable z^* . This continuous variable z^* represent the attitude underlying the ordered responses to z and is assumed to have a range from $-\infty$ to $+\infty$. The underlying variable z^* can be used in structural equation modeling, not the observed variable z . The underlying variable assigns a metric to the ordinal variable.

If z has m categories labeled $1, 2, \dots, m$, the connection between z and z^* is

$$z = i \iff \tau_{i-1} < z^* < \tau_i, \quad i = 1, 2, \dots, m,$$

where

$$-\infty = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_{m-1} < \tau_m = +\infty,$$

are parameters called threshold values. With m categories, there are $m - 1$ threshold parameters $\tau_1, \tau_2, \dots, \tau_{m-1}$.

Because we have only ordinal information, the distribution of z^* is determined only up to a monotonic transformation. In principle, one can choose any continuous distribution for z^* . However, any continuous variable with a density and a distribution function can be transformed by a monotonic transformation to a normal distribution. It is therefore convenient to choose the standard normal distribution with density function $\phi(u)$ and distribution function $\Phi(u)$ for z^* . Then the probability of a response in category i is

$$\pi_i = Pr[z = i] = Pr[\tau_{i-1} < z^* < \tau_i] = \int_{\tau_{i-1}}^{\tau_i} \phi(u) du = \Phi(\tau_i) - \Phi(\tau_{i-1}), \quad (1)$$

so that

$$\tau_i = \Phi^{-1}(\pi_1 + \pi_2 + \dots + \pi_i), \quad i = 1, \dots, m - 1, \quad (2)$$

where Φ^{-1} is the inverse of the standard normal distribution function. The quantity $(\pi_1 + \pi_2 + \dots + \pi_i)$ is the probability of a response in category i or lower.

The probabilities π_i are unknown population quantities. In practice, π_i can be estimated consistently by the corresponding percentage p_i of responses in category i which is given in the PRELIS output. Then, estimates of the thresholds can be obtained as

$$\hat{\tau}_i = \Phi^{-1}(p_1 + p_2 + \dots + p_i), \quad i = 1, \dots, m - 1. \quad (3)$$

The quantity $(p_1 + p_2 + \dots + p_i)$ is the proportion of cases in the sample responding in category i or lower. Equation (3) is in fact the maximum likelihood estimator of τ_i based on the univariate marginal sample data. But the model is saturated; there are $m - 1$ parameters τ_i and there are $m - 1$ independent sample proportions p_i . The fit is perfect since

$$\hat{\pi}_i = \Phi(\hat{\tau}_i) - \Phi(\hat{\tau}_{i-1}) = p_i .$$

The estimation of thresholds is illustrated in Fig. 1. Suppose there are 8, 41, 39, and 12% responding in category 1, 2, 3, and 4, respectively. Cumulatively this is 8, 49, 88, 100%. The first threshold is located where the area under the normal to the left of the threshold is 8%. The second threshold is located where the area under the normal to the left of the threshold is 49%. The third threshold is located where the area under the normal to the left of the threshold is 88%. This gives approximately

$$\hat{\tau}_1 = -1.404 , \quad \hat{\tau}_2 = -0.025 , \quad \hat{\tau}_3 = 1.075 .$$

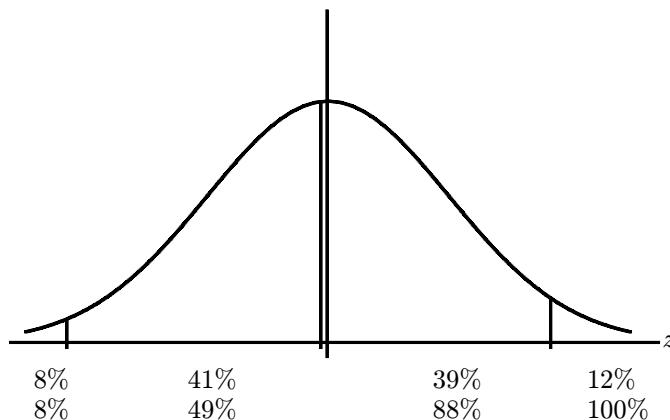


Figure 1: Illustrating Thresholds

2.2 Polychoric Correlations

Let z_1 and z_2 be two ordinal variables with m_1 and m_2 categories, respectively. Their marginal distribution in the sample is represented by a contingency table

$$\begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1m_2} \\ n_{21} & n_{22} & \cdots & n_{2m_2} \\ \vdots & \vdots & \ddots & \vdots \\ n_{m_1 1} & n_{m_1 2} & \cdots & n_{m_1 m_2} \end{pmatrix} ,$$

where n_{ij} is the number of cases in category i on variable 1 and in category j on variable 2. Since the underlying variables z_1^* and z_2^* are normal with zero means, unit variances, it is natural to assume that z_1^* and z_2^* are standard bivariate normal with correlation ρ . However, unlike the univariate marginal normality of z_1^* and z_2^* , which cannot be falsified using univariate marginal data, the bivariate normality of z_1^* and z_2^* is an assumption which is testable using bivariate marginal data. Normality of z_1^* and z_2^* does not imply bivariate normality of z_1^* and z_2^* .

The *polychoric correlation* is the correlation ρ in the bivariate normal distribution of the underlying variables z_1^* and z_2^* . If $m_1 = m_2 = 2$ this is called a *tetrachoric correlation*.

Let $\tau_1^{(1)}, \tau_2^{(1)}, \dots, \tau_{m_1-1}^{(1)}$ be the thresholds for variable z_1^* and let $\tau_1^{(2)}, \tau_2^{(2)}, \dots, \tau_{m_2-1}^{(2)}$ be the thresholds for variable z_2^* . Then the polychoric correlation can be estimated by maximizing the log-

likelihood of the multinomial distribution,

$$\ln L = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} n_{ij} \log \pi_{ij}(\boldsymbol{\theta}),$$

where

$$\pi_{ij}(\boldsymbol{\theta}) = Pr[z_1 = i, z_2 = j] = \int_{\tau_{i-1}^{(1)}}^{\tau_i^{(1)}} \int_{\tau_{j-1}^{(2)}}^{\tau_j^{(2)}} \phi_2(u, v) du dv, \quad (4)$$

and

$$\phi_2(u, v) = \frac{1}{2\pi\sqrt{(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^2)}(u^2 - 2\rho uv + v^2)},$$

is the standard bivariate normal density with correlation ρ .

The model defined by (4) expresses the $m_1 m_2$ probabilities $\pi_{ij}(\boldsymbol{\theta})$ as functions of the parameter vector

$$\boldsymbol{\theta} = (\tau_1^{(1)}, \tau_2^{(1)}, \dots, \tau_{m_1-1}^{(1)}, \tau_1^{(2)}, \tau_2^{(2)}, \dots, \tau_{m_2-1}^{(2)}, \rho), \quad (5)$$

consisting of the thresholds for the two variables and the polychoric correlation ρ . Maximizing $\ln L$ is equivalent to minimizing the fit function

$$F(\boldsymbol{\theta}) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} p_{ij} [\ln p_{ij} - \ln \pi_{ij}(\boldsymbol{\theta})] = \sum_{ij} p_{ij} \ln [p_{ij} / \pi_{ij}(\boldsymbol{\theta})], \quad (6)$$

where $p_{ij} = n_{ij}/N$ are the sample proportions.

The estimation of the polychoric correlation is illustrated in Fig. 2. Here the numbers are the sample proportions and the vertical and horizontal lines represent the thresholds. The ellipse represents the standard bivariate normal distribution with correlation ρ . The value of ρ determines the orientation and concentration of the ellipse. This can be fitted to the sample proportions.

	10.6	4.5	1.2	0.1
5.2	34.5	13.9	0.3	
0.4	7.2	19.5	0.9	
0.1	0.4	0.6	0.5	

Figure 2: Estimating the Polychoric Correlation

PRELIS estimates the parameters by a two-step procedure, see Olsson (1979). In the first step, the thresholds are estimated from the univariate marginal distributions by (3). In the second step, the

polychoric correlations are estimated from the bivariate marginal distributions by minimizing (6) *for given thresholds*. The parameters can also be estimated by a one-step procedure. It is no problem to estimate the thresholds and the polychoric correlations simultaneously, but this is not necessary and it is not practical. It is not necessary because the estimates are almost the same as with the two-step procedure. It is not practical because it would yield different thresholds for the same variable when paired with different variables. For an example, see Section 2.3.

2.3 Testing Underlying Bivariate Normality

To test the model, one can use the likelihood ratio (LR) test statistic

$$\chi_{\text{LR}}^2 = 2 \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} n_{ij} \ln(p_{ij}/\hat{\pi}_{ij}) = 2N \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} p_{ij} \ln(p_{ij}/\hat{\pi}_{ij}) = 2NF(\hat{\boldsymbol{\theta}}), \quad (7)$$

where $\hat{\boldsymbol{\theta}}$ is the estimated parameter vector and $\hat{\pi}_{ij} = \pi_{ij}(\hat{\boldsymbol{\theta}})$. Hence, this χ^2 is just $2N$ times the minimum value of the fit function (6). If the model holds, this is distributed approximately as χ^2 with degrees of freedom

$$d = (m_1 m_2 - 1) - (m_1 - 1) - (m_2 - 1) - 1 = m_1 m_2 - m_1 - m_2. \quad (8)$$

Alternatively, one can use the goodness-of-fit (GF) test statistic

$$\chi_{\text{GF}}^2 = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} [(n_{ij} - N\hat{\pi}_{ij})^2 / (N\hat{\pi}_{ij})] = N \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (p_{ij} - \hat{\pi}_{ij})^2 / \hat{\pi}_{ij}. \quad (9)$$

If the model holds, both statistics (7) and (9) have the same asymptotic distribution. In practice, when the fit is good the LR and GF statistic are quite close but when the model does not fit well they can be quite different.³

If $m_1 = m_2 = 2$, $d = 0$. That is, if both variables are dichotomous, the degrees of freedom is zero, the model is saturated and it is not possible to test underlying bivariate normality.

Because of the two-step procedure used in PRELIS, the χ_{LR}^2 reported by PRELIS slightly over-estimates the correct asymptotic chi-square. However, this is of no practical importance since the parameter estimates are essentially the same, whether the two-step or the one-step procedure is used. Table 3 gives an example based on the data for NOSAY vs VOTING.

Table 3: Two-Step vs One-Step Procedure

	Two-Step Method			One-Step Method		
Thresholds for NOSAY	-1.267	-0.222	1.423	-1.267	-0.222	1.419
Thresholds for VOTING	-0.942	0.220	1.666	-0.942	0.218	1.666
Polychoric Correlation	0.329			0.330		
LR Chi-Square with 8 df	222.381			222.365		
GF Chi-Square with 8 df	325.120			324.804		

This example shows that the parameter estimates and chi-squares are essentially the same for the two methods. It also shows that the LR-chi-square and the GF-chi-square can be quite different when the model does not fit. Lack of fit will be discussed in the next section.

³The LR and GF statistics are sometimes denoted G^2 and X^2 , respectively.

2.4 Estimating the Polychoric Correlations and their Asymptotic Covariance Matrix

We now return to the analysis of the data on the six political efficacy items. Do the six items measure one uni-dimensional latent variable? For this to work there must be a clear correspondence between the values of the latent variable and the categories of the ordinal variables, see Costner (1969). In this case, 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 these 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 I am high on *Efficacy* and I believe that voting is the only way I can influence politics, then I 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. However, to begin with, I shall keep it to demonstrate what consequences this has in the data analysis.

To estimate the one-factor model there are two steps:

- Use PRELIS to estimate the polychoric correlations and their asymptotic covariance matrix.
- Use these matrices in LISREL to estimate the one-factor model with weighted least squares (WLS).

The PRELIS step is described in this section and the LISREL step in Section 2.5. For the PRELIS step, use the following syntax file (file `ORD21.PR2`):

```
EFFICACY: PRELIS Run 4
Computing Polychoric Correlations and Asymptotic Covariance Matrix
Data Ninputvariables = 6
Labels
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata=EFFICACY.IMP
CLabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS
Output MA=PM PM=EFFICACY.PM AC=EFFICACY.ACP
```

The data is in the file `EFFICACY.IMP`, see p. 6. `MA=PM` on the `Output` line means that the matrix of polychoric correlations should be estimated, `PM=EFFICACY.PM` means that this matrix will be saved in the file `EFFICACY.PM` and `AC=EFFICACY.ACP` means that the asymptotic covariance matrix will be saved in the file `EFFICACY.ACP`. The first file is in text (ASCII, readable) format and the second is in binary (unreadable) format.

If you have saved the imputed data as a PSF file, `EFFICACY.PSF`, say, instead of as the text file `EFFICACY.IMP`, you can use a much shorter PRELIS input file⁴ (file `ORD21A.PR2`):

```
EFFICACY: PRELIS Run 4
Computing Polychoric Correlations and Asymptotic Covariance Matrix
SY=EFFICACY.PSF
Output MA=PM PM=EFFICACY.PM AC=EFFICACY.ACP
```

A new feature in PRELIS 2.51 not available in PRELIS 2.50 is that the program will write a separate output file listing all response patterns occurring in the sample and the frequency of occurrence of each pattern. This file has the name *inputfilename*.`FREQ`. For example, if the input file for PRELIS Run 4 is `ORD21.PR2`, the `FREQ` file will be `ORD21.FREQ`. The 20 most common response patterns will also be listed in the ordinary output file. The `FREQ` file should be regarded as a data file. It gives the data in the most concise form. The `FREQ` file may be read by PRELIS by specifying the frequency

⁴An easy way to obtain the file `EFFICACY.PSF` is to add `RA=EFFICACY.PSF` on the `Output` line in PRELIS Run 3, see p. 9 and file `ORD13A.PSF`.

variable as a weight variable. The reason for the `FREQ` file having the name of the input file rather than the name of the data file is that it depends on the `PRELIS` commands in the input file. If you have `PRELIS 2.51`, I suggest that you rerun the three `PRELIS` input files given in Section 1 and take a look at the corresponding `FREQ` files.

In the `Efficacy` example, each variable has four categories. Thus, there are 4096 possible response patterns but since we have data on only 1589 cases, every response pattern cannot be present in the data. In fact, there are only 476 different response patterns occurring in the sample. All these are given in the `FREQ` file. The output file lists the 20 most common response patterns as

There are 476 distinct response patterns, see `FREQ`-file.

The 20 most common patterns are :

```

106      2  2  2  2  2  2
  70      3  3  2  3  3  3
  51      3  2  2  2  2  2
  45      3  3  3  3  3  3
  45      3  3  2  2  2  2
  40      3  2  2  3  3  3
  32      3  3  2  3  2  2
  31      3  3  2  3  2  3
  25      2  2  1  2  2  2
  23      1  1  1  1  1  1
  20      3  2  2  3  2  2
  19      2  2  1  1  1  1
  19      2  3  2  2  2  2
  18      3  3  2  2  2  3
  16      3  3  3  3  2  3
  16      3  2  2  3  2  3
  15      3  3  2  3  3  2
  15      3  2  3  3  3  3
  15      2  1  2  2  2  2
  14      2  2  2  3  2  2

```

It is seen that the most common response pattern is to answer *Agree* to all six items (106 cases). The second most common response pattern is to answer *Agree* to `COMPLEX` and *Disagree* to all the other five items (70 cases). To investigate whether there is a tendency to give the same response to all six items, one can note that there are 45 cases who *Disagree* to all six items and 23 cases who *Agree Strongly* to all six items. Further screening of the `FREQ` file reveals that there are only 3 cases who *Disagree Strongly* to all six items. Note that among the 20 most common response patterns there are no response *Disagree Strongly* on *any* of the six variables.

The output gives the following table of univariate marginal parameters. This refers to the underlying variables. An alternative parameterization is presented in Section 2.7.

Univariate Marginal Parameters

Variable	Mean	St. Dev.	Thresholds		
-----	----	-----	-----	-----	-----
NOSAY	0.000	1.000	-1.267	-0.222	1.423
VOTING	0.000	1.000	-0.942	0.220	1.666
COMPLEX	0.000	1.000	-0.850	0.729	1.770
NOCARE	0.000	1.000	-1.029	0.162	1.809
TOUCH	0.000	1.000	-0.977	0.535	2.151
INTEREST	0.000	1.000	-0.980	0.330	2.091

With 6 variables there are 15 different pairs of variables. The contingency tables for each of these is given in the output as⁵

Bivariate Distributions for Ordinal Variables (Frequencies)

	VOTING				COMPLEX				NOCARE			
NOSAY	AS	A	D	DS	AS	A	D	DS	AS	A	D	DS
AS	70	51	28	14	84	52	17	10	81	60	19	3
A	82	313	88	9	123	303	57	9	106	292	93	1
D	93	280	413	25	90	496	202	23	44	286	457	24
DS	30	14	51	28	17	54	33	19	10	18	67	28

	TOUCH				INTEREST			
NOSAY	AS	A	D	DS	AS	A	D	DS
AS	76	69	16	2	86	51	24	2
A	101	329	61	1	101	296	95	0
D	63	415	318	15	57	362	380	12
DS	21	44	51	7	16	31	61	15

	COMPLEX				NOCARE				TOUCH			
VOTING	AS	A	D	DS	AS	A	D	DS	AS	A	D	DS
AS	104	127	37	7	74	110	79	12	82	137	50	6
A	134	407	101	16	99	342	208	9	103	415	136	4
D	63	343	153	21	55	193	312	20	58	279	233	10
DS	13	28	18	17	13	11	37	15	18	26	27	5

	INTEREST			
VOTING	AS	A	D	DS
AS	86	105	75	9
A	107	368	181	2
D	50	249	270	11
DS	17	18	34	7

	NOCARE				TOUCH				INTEREST			
COMPLEX	AS	A	D	DS	AS	A	D	DS	AS	A	D	DS
AS	136	117	56	5	133	145	31	5	125	143	41	5
A	84	435	370	16	91	538	269	7	101	462	332	10
D	16	89	187	17	23	156	124	6	25	119	162	3
DS	5	15	23	18	14	18	22	7	9	16	25	11

⁵PRELIS 2.50 automatically gives these tables and the corresponding tables of sample proportions in the output; they can be omitted by putting XB on the Output line. Several users have suggested that it should be the other way around, *i.e.*, they should only be given if requested. In PRELIS 2.51, we have therefore changed it to this effect. In PRELIS 2.51, these tables will only appear in the output if requested by BT on the Output line.

	TOUCH				INTEREST			
	AS	A	D	DS	AS	A	D	DS
NOCARE	152	74	13	2	156	71	14	0
AS	86	483	85	2	81	456	117	2
A	20	286	319	11	17	201	407	11
D	3	14	29	10	6	12	22	16

	INTEREST			
	AS	A	D	DS
TOUCH	169	72	19	1
AS	83	548	221	5
A	7	114	310	15
D	1	6	10	8

Note that there are two zero cells in these tables, one in the table for NOSAY and INTEREST and one in the table for NOCARE and INTEREST. The fit function (6) is not defined if p_{ij} is zero. In the estimation of the polychoric correlations, PRELIS skips such zero cells. But too many zero cells in a table can be problematic and give estimates that are imprecise and unreliable. If there is only one nonzero cell in a row or a column, the estimation procedure breaks down. A zero cell is particularly problematic in the case when both variables are dichotomous because then the tetrachoric correlation is undefined. In this case PRELIS replaces the zero by a half. But the estimate of the tetrachoric correlation is very sensitive to changes in the value used to substitute the zero cell. Other programs use different values and therefore get different estimates of the tetrachoric correlation.

From the table of NOSAY vs VOTING it is seen that of the 163 cases who agree strongly with NOSAY, 14 disagree strongly with VOTING. It is also seen that of the 275 cases who agree strongly with VOTING, 30 disagree strongly with NOSAY. Are these numbers consistent with an underlying bivariate normal distribution? The answer is given in the next table in the output.

Correlations and Test Statistics

		(PE=Pearson Product Moment, PC=Polychoric, PS=Polyserial)					
				Test of Model		Test of Close Fit	
Variable vs.	Variable	Correlation	Chi-Squ.	D.F.	P-Value	RMSEA	P-Value
VOTING vs.	NOSAY	0.329 (PC)	222.381	8	0.000	0.130	0.000
W_A_R_N_I_N_G: Underlying bivariate normality may not hold, see BTS-file							
COMPLEX vs.	NOSAY	0.330 (PC)	81.237	8	0.000	0.076	0.995
COMPLEX vs.	VOTING	0.291 (PC)	42.658	8	0.000	0.052	1.000
NOCARE vs.	NOSAY	0.559 (PC)	77.651	8	0.000	0.074	0.998
NOCARE vs.	VOTING	0.276 (PC)	76.577	8	0.000	0.073	0.998
NOCARE vs.	COMPLEX	0.462 (PC)	66.600	8	0.000	0.068	1.000
TOUCH vs.	NOSAY	0.402 (PC)	85.511	8	0.000	0.078	0.991
TOUCH vs.	VOTING	0.243 (PC)	61.711	8	0.000	0.065	1.000
TOUCH vs.	COMPLEX	0.353 (PC)	95.351	8	0.000	0.083	0.967
TOUCH vs.	NOCARE	0.645 (PC)	92.483	8	0.000	0.082	0.977
INTEREST vs.	NOSAY	0.456 (PC)	83.077	8	0.000	0.077	0.994
INTEREST vs.	VOTING	0.242 (PC)	82.070	8	0.000	0.076	0.995
INTEREST vs.	COMPLEX	0.379 (PC)	70.733	8	0.000	0.070	0.999
INTEREST vs.	NOCARE	0.681 (PC)	132.135	8	0.000	0.099	0.535
INTEREST vs.	TOUCH	0.685 (PC)	89.920	8	0.000	0.080	0.983

For each pair of variables this gives the estimated polychoric correlation, the LR-chi-square, the degrees of freedom, and the P -value. It is seen that the hypothesis of underlying bivariate normality is rejected for all pairs of variables.

Models do not have to be true to be useful. It is sufficient that the model holds approximately. If the model holds approximately, it will be rejected by the LR-chi-square in large samples. We are not interested in this test as such. We are only interested in obtaining a suitable correlation for ordinal variables. The assumption of underlying bivariate normality is needed to calculate the polychoric correlation. How robust is the polychoric correlation to violations of underlying bivariate normality? This was studied in the dissertation of Ana Maria Quiroga (1992) who found that the polychoric correlation is very robust to violations of underlying bivariate normality but the LR-chi-square is very sensitive. This has led me to consider an alternative test based on the non-central, rather than the central, chi-square distribution.

I have developed an RMSEA measure of population discrepancy similar to Steiger's (1990) RMSEA measure for structural equation models. The theory for this is outlined in Appendix 1. By further simulation studies based on varying degree of underlying bivariate non-normality I have found that there are no serious effects of non-normality unless RMSEA is larger than 0.1. The last two columns of the table give the value of RMSEA and the P -value for the test of the hypothesis that the population value of RMSEA is less than 0.1. By these criteria it is seen that the hypothesis of approximate underlying bivariate normality is rejected only for the pair NOSAY vs VOTING. I will therefore take a closer look at the data for this pair.

If the P -value for the test of approximate underlying normality is less than 0.05, PRELIS writes some further information to a file *inputfilename.BTS*, where *inputfilename* is the name of the input file. For example, if the input file for PRELIS Run 4 is *ORD21.PR2*, the BTS file will be *ORD21.BTS*. In this case, the BTS file gives four bivariate tables for NOSAY vs VOTING. These will be explained in turn.

The first table is the contingency table for NOSAY and VOTING.

Observed Frequencies

	AS	A	D	DS	Rowsum
AS	70	51	28	14	163
A	82	313	88	9	492
D	93	280	413	25	811
DS	30	14	51	28	123
Colsum	275	658	580	76	1589

The second table gives the expected frequencies under underlying bivariate normality. Note that the univariate margins fit exactly. Note also that the observed frequencies overestimate the expected frequencies in the (AS,DS) and (DS,AS) cells.

Expected Frequencies

	AS	A	D	DS	Rowsum
AS	57.0	73.0	31.5	1.5	163.0
A	110.5	224.1	146.0	11.4	492.0
D	101.3	325.5	337.8	46.4	811.0
DS	6.3	35.4	64.7	16.7	123.0
Colsum	275.0	658.0	580.0	76.0	1589.0

The next table gives the cell contributions to the LR chi-square statistic. The LR cell contribution in row i and column j is

$$2Np_{ij} \ln(p_{ij}/\hat{\pi}_{ij}) .$$

The sum of these cell contributions is the LR-chi-square statistic (7) previously given in the output as 222.381. Note that the LR-contributions can be positive or negative. A positive value indicates that the observed value overestimates the corresponding expected value, whereas a negative value indicates that the observed value underestimates the corresponding expected value.

LR Contributions

	AS	A	D	DS	Rowsum
AS	28.8	-36.6	-6.6	62.1	47.7
A	-48.9	209.2	-89.1	-4.3	66.9
D	-15.9	-84.3	165.9	-30.9	34.9
DS	94.0	-26.0	-24.2	29.1	72.9
Colsum	58.1	62.3	46.0	55.9	222.4

The next table gives the cell contributions to the GF-chi-square statistic (9). The GF cell contribution in row i and column j is

$$N(p_{ij} - \hat{\pi}_{ij})^2 / \hat{\pi}_{ij} .$$

These are always non-negative, so they do not provide any information about over- or underestimation. The sum of these cell contributions is the GF-statistic which in this case is 325.4, quite different from the LR-statistic. Here it can be seen more clearly that most of the lack of fit is associated with the (AS,DS) and (DS,AS) cells. It is also seen that the bad fit is associated with the AS and DS categories.

GF Contributions

	AS	A	D	DS	Rowsum
AS	3.0	6.6	0.4	102.0	112.0
A	7.3	35.3	23.0	0.5	66.2
D	0.7	6.4	16.7	9.9	33.6
DS	90.1	13.0	2.9	7.7	113.7
Colsum	101.1	61.2	43.0	120.1	325.4

The best way to locate the source of bad fit is to examine the standardized residuals. A residual for a cell can be defined as the square root of a GF contribution, *i.e.*,

$$\frac{n_{ij} - N\hat{\pi}_{ij}}{\sqrt{N\hat{\pi}_{ij}}} = \sqrt{N} \frac{p_{ij} - \hat{\pi}_{ij}}{\sqrt{\hat{\pi}_{ij}}} . \quad (10)$$

It turns out that the asymptotic variance of (10) is not 1 but smaller than 1. The asymptotic variance of (10) is (see *e.g.*, Rao (1965), eq. (6b.3.2))

$$v_{ij} = 1 - \pi_{ij} - \sum_s \sum_t (1/\pi_{ij})(\partial\pi_{ij}/\partial\theta_s)(\partial\pi_{ij}/\partial\theta_t)e^{st} ,$$

where e^{st} is an element of the inverse of the information matrix

$$\mathbf{E} = \sum_{ij} (1/\pi_{ij})(\partial\pi_{ij}/\partial\theta)(\partial\pi_{ij}/\partial\theta') .$$

The residual in (10) divided by the square root of v_{ij} gives the standardized residual. The standardized residuals are given in the last table of the BTS file. It is seen that several standardized residuals are too large or too small. The two largest standardized residuals occur for the (AS,DS) and (DS,AS) cells.

Standardized Residuals

	AS	A	D	DS	Rowsum
AS	2.7	-3.7	-0.8	10.6	8.7
A	-4.0	9.9	-7.4	-0.8	-2.4
D	-1.3	-4.7	8.6	-5.2	-2.7
DS	10.6	-4.8	-2.6	3.5	6.6
Colsum	7.9	-3.3	-2.3	8.0	10.3

All the tables in the BTS file suggest that there are too many cases in the (AS,DS) and (DS,AS) cells. This is in line with the previous argument that VOTING is an ambiguous item.

What should one do when underlying bivariate normality does not even hold approximately? From a practical point of view there are three things that can be done.

- Reduce the number of categories. In this case one could collapse the two categories *agree strongly* and *agree* into one category and the two categories *disagree strongly* and *disagree* into one category. This can be done for any of the variables. If it is done for all variables, all variables will be dichotomous. The model of underlying bivariate normality then becomes saturated as the total number of parameters (3) equals the number of independent cells in the contingency table.
- Eliminate the most offending variables thereby obtaining more homogeneity for the retained variables. In this case I would recommend deleting the variable VOTING since this has an ambiguous question wording.
- If the probability of the various response patterns depends on covariates such as gender, age, income, and education, one can replace the assumption of underlying bivariate normality with the assumption of underlying bivariate normality conditional on the covariates. This is a much more flexible assumption. I will consider this case in Section 5.

It is better to be theory-driven than data-driven. So, I choose the second alternative. The VOTING variable can be eliminated in PRELIS or in LISREL. I choose to eliminate it in LISREL. To eliminate it in PRELIS, just add the line

Sdelete VOTING

in PRELIS Run 4. Alternatively, one can select the variables to be included with

Select NOSAY COMPLEX - INTEREST

At the end of the output file for PRELIS Run 4, the matrix of polychoric correlations is given as

Correlation Matrix

	NOSAY	VOTING	COMPLEX	NOCARE	TOUCH	INTEREST
NOSAY	1.000					
VOTING	0.329	1.000				
COMPLEX	0.330	0.291	1.000			
NOCARE	0.559	0.276	0.462	1.000		
TOUCH	0.402	0.243	0.353	0.645	1.000	
INTEREST	0.456	0.242	0.379	0.681	0.685	1.000

Note that the correlations for VOTING are all smaller than all the other correlations. This is another indication that VOTING does not belong with the other items.

2.5 Fitting and Testing the Model

Do the five efficacy items (after exclusion of VOTING) measure a unidimensional trait? This can be tested by running the following SIMPLIS syntax file (file ORD22.SPL):

```
EFFICACY: LISREL Run 1
Testing Measurement Model 1
Observed Variables: NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Correlation Matrix from File EFFICACY.PM
Asymptotic Covariance Matrix from File EFFICACY.ACP
Sample Size: 1589
Latent Variable: Efficacy
Relationships:
    NOSAY COMPLEX - INTEREST = Efficacy
Path Diagram
End of Problem
```

Note that the selection of variables is automatic, *i.e.*, no selection line is needed. VOTING is included in the correlation matrix and in the asymptotic covariance matrix but since it is not included in the model, it will not be used.

Since PRELIS Run 4 generates a data system file or DSF file for short, see Jöreskog, *et al.* (2001), p. 169. By reading the data from this file, here assumed to have the name ORD21.DSF, one can use the following shorter SIMPLIS command file (file ORD22A.SPL):

```
EFFICACY: LISREL Run 1
Testing Measurement Model 1
System File from File ORD21.DSF
Latent Variable: Efficacy
Relationships:
    NOSAY COMPLEX - INTEREST = Efficacy
Path Diagram
End of Problem
```

Some selected lines from the fit statistics are

```
Degrees of Freedom = 5
Minimum Fit Function Chi-Square = 28.10 (P = 0.00)
Root Mean Square Error of Approximation (RMSEA) = 0.054
90 Percent Confidence Interval for RMSEA = (0.036 ; 0.074)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.33
```

In terms of exact fit, the model is rejected. However, the last three lines indicate that the model might hold approximately in the population, see Browne & Cudeck (1993). But since I want a model that has a better chance of being invariant over time and across countries, I will consider an alternative model. 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. To estimate this model, use the following SIMPLIS input file (file ORD23A.SPL) or, alternatively, ORD23.SPL:

```
EFFICACY: LISREL Run 2
Testing Measurement Model 2
System File from File ORD21.DSF
Latent Variables: Efficacy Respons
Relationships:
    NOSAY COMPLEX NOCARE = Efficacy
    NOCARE TOUCH INTEREST = Respons
Path Diagram
End of Problem
```

This gives the following test statistics

```
Degrees of Freedom = 3
Minimum Fit Function Chi-Square = 1.38 (P = 0.71)
Root Mean Square Error of Approximation (RMSEA) = 0.0
90 Percent Confidence Interval for RMSEA = (0.0 ; 0.031)
P-Value for Test of Close Fit (RMSEA < 0.05) = 1.00
```

indicating that the model fits very well. The path diagram with parameter estimates is shown in Fig. 3. The path from Respons to NOCARE is not significant. Its t -value is 1.10. This does not mean that this path should be eliminated if one believes it exists.

I will investigate this model and its invariance over time in Section 3 and its invariance across countries in Section 4.

2.6 How to Become a Chi-square Collector

It is easy to become a chi-square collector. Each output from LISREL gives you at least two chi-squares, sometimes four or five. I have many “wrong” ones and some “right” ones in my collection. To get some of each kind, add the line

Method: Maximum Likelihood

before the End of Problem line. This means that the ML method will be used to fit the model but standard errors and chi-squares will be corrected for non-normality using the asymptotic covariance matrix provided. This has the advantage that the asymptotic covariance matrix need not be inverted. This method works better with smaller sample sizes. Simulation studies suggest that this may work with samples as small as 200.

The output gives four different chi-squares as

```
Degrees of Freedom = 3
Minimum Fit Function Chi-Square = 3.85 (P = 0.28)
Normal Theory Weighted Least Squares Chi-Square = 3.83 (P = 0.28)
Satorra-Bentler Scaled Chi-Square = 1.44 (P = 0.70)
Chi-Square Corrected for Non-Normality = 1.38 (P = 0.71)
```

The first two are the “wrong” ones and the last two are the “right” ones.⁶ The first two (3.85 and 3.83) assume normality but normality does not hold here. The last two chi-squares (1.44 and 1.38)

⁶You might say that it does not matter if you choose the wrong chi-square since it does not change the conclusion about the fit of the model. However, the “wrong” ones are in fact about 80% larger than the “right” ones. In many cases these differences are large enough to make a real difference in conclusion. I suggest that you do the same analysis with LISREL Run 1. Then the “right” ones suggest that fit is almost OK but the “wrong” ones suggest that the model does not fit at all.

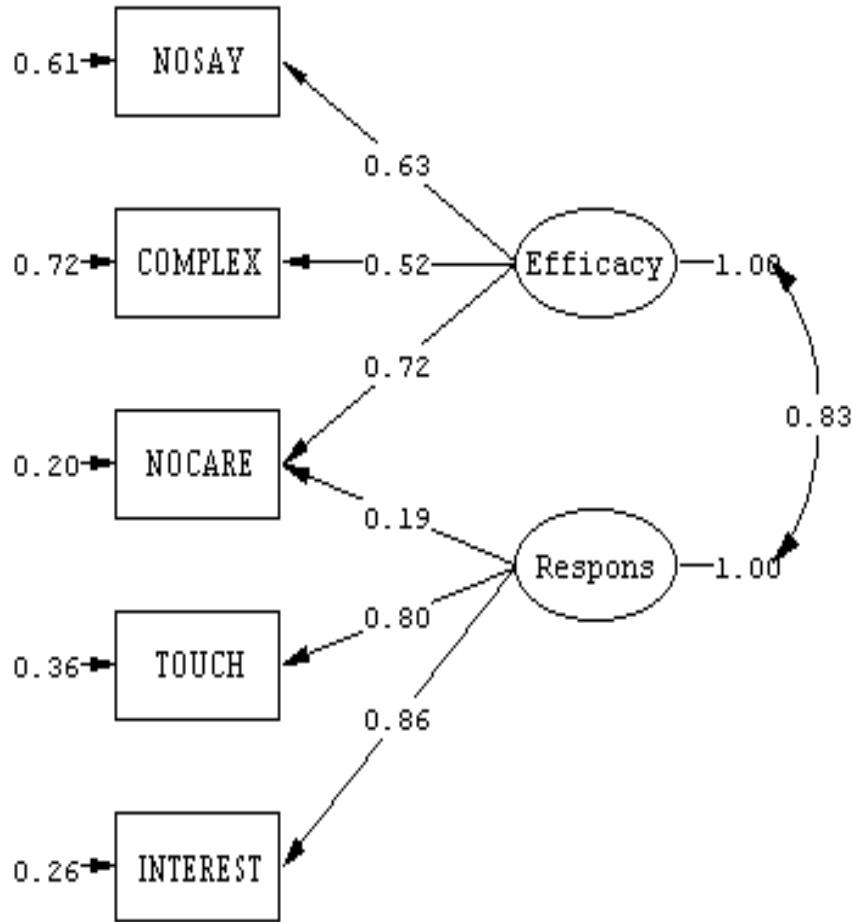


Figure 3: Path Diagram with Parameter Estimates

take non-normality into account by using asymptotic covariance matrix. The third chi-square (1.44) does not require this matrix to be inverted whereas the fourth one (1.38) does. It is the third one that seems to work best in small samples. For further explanation see Jöreskog, *et al.* (2001), Chapter 4 and Appendix A.

2.7 Alternative Parameterization

A weakness of the approach outlined in Sections 2.1 and 2.2 is that all underlying variables are standardized to zero means and unit standard deviations. But when the response alternatives are the same for several variables, differences in distributions of these variables may reflect differences in means and/or variances of the underlying variables.

To use such differences, another parameterization of the underlying variables is necessary. This is obtained by putting AP on the OU line in PRELIS 2.50. In PRELIS 2.51, this parameterization will automatically be used if MA=CM (or MA=MM or MA=AM). We refer to the default parameterization as the *Standard Parameterization* and the AP parameterization as the *Alternative Parameterization*.

The two parameterizations can be explained as follows. The variable z^* underlying the ordinal variable z is determined only up to a monotonic transformation. If we want to retain normality of the underlying variable, the transformation must be linear. In principle, one can make an arbitrary

linear transformation of the underlying variable. If the number of categories $m \geq 3$, one such transformation is obtained by specifying that $\tau_1 = 0$ and $\tau_2 = 1$. Then the mean μ and standard deviation σ of z^* can be defined instead.

The parameters of the two parameterizations are given in the following table.

Parameterization	Mean	St.Dev.	Thresholds				
Standard	0	1	τ_1	τ_2	τ_3	\dots	τ_{m-1}
Alternative	μ^*	σ^*	0	1	τ_3^*	\dots	τ_{m-1}^*

where

$$\begin{aligned} \mu^* &= -\tau_1/(\tau_2 - \tau_1), & \sigma^* &= 1/(\tau_2 - \tau_1), \\ \tau_i^* &= (\tau_i - \tau_1)/(\tau_2 - \tau_1), & i &= 3, 4, \dots, m - 1. \end{aligned}$$

It should be emphasized that the two parameterizations are equivalent in the sense that there is a one-to-one correspondence between the two sets of parameters. This correspondence carries further as will be illustrated with the Efficacy example.

For the Alternative Parameterization to be meaningful there must be at least three categories. If there are only two categories there is only one threshold and it is impossible to estimate both the mean and the standard deviation of the underlying variable. In this case, PRELIS will fix the threshold at 0, the standard deviation at 1 and estimate the mean.

Before I return to the Efficacy example, I will illustrate the Alternative Parameterization using a small data set consisting of two ordinal variables on a 5-point scale.⁷ The contingency table and the marginal distributions are

	X2					
X1	1	2	3	4	5	
1	85	307	315	291	56	1054
2	61	353	520	687	236	1857
3	22	172	248	478	227	1147
4	8	53	138	393	232	824
5	0	6	15	44	53	118
	176	891	1236	1893	804	5000

If these two variables are on the same response scale, there is a shift in the distribution: For X1 most people are in the low end of the scale whereas for X2 most people are in the high end of the scale. How can this fact be represented?

To read the data in PRELIS, one can use a frequency data file `TWOVARS.FRQ` consisting of 24 lines as follows

```

85 1 1
307 1 2
315 1 3
291 1 4
56 1 5
61 2 1
353 2 2
520 2 3
687 2 4
236 2 5
22 3 1
172 3 2
```

⁷Rolf Steyer generated 5000 observations on these two variables, but for convenience I have converted the data to a 5×5 contingency table.

```

248 3 3
478 3 4
227 3 5
  8 4 1
 53 4 2
138 4 3
393 4 4
232 4 5
  6 5 2
 15 5 3
 44 5 4
 53 5 5

```

Reading this is much faster than reading the data file with 5000 lines.

Now run the following PRELIS syntax file (file TWOVARS.PR2):

```

Illustration of Alternative Parameterization
DA NI=3
RA=TWOVARS.FRQ
LA
FREQ X1 X2
WEIGHT FREQ
OU MA=CM AP

```

With PRELIS 2.51 AP on the OU line is not needed.

The output reveals the following

Univariate Marginal Parameters

Variable	Mean	St. Dev.	Thresholds			
-----	----	-----	-----	-----	-----	-----
X1	0.795	0.989	0.000	1.000	1.669	2.757
X2	1.783	0.986	0.000	1.000	1.686	2.760

from which it is seen that there is a shift in the mean of the underlying variables of about one unit, whereas the standard deviations and the thresholds are roughly the same.

I now return to the Efficacy example. Run the following PRELIS input (file ORD24.PR2):

```

EFFICACY: PRELIS Run 5
Computing Covariance Matrix and Asymptotic Covariance Matrix
Data Ninputvariables = 6
Labels
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
RA=EFFICACY.IMP
CLabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS
Output MA=CM AP CM=EFFICACY.CM AC=EFFICACY.ACC

```

This is the same as PRELIS Run 4 except for the last line. Here we are computing the covariance matrix and its asymptotic covariance matrix instead of the polychoric correlation matrix and its asymptotic covariance matrix. With PRELIS 2.51 AP on the OU line is not needed. The output gives the covariance matrix, the mean vector and the standard deviations as

Covariance Matrix

	NOSAY	VOTING	COMPLEX	NOCARE	TOUCH	INTEREST
NOSAY	0.916					
VOTING	0.271	0.740				
COMPLEX	0.200	0.158	0.401			
NOCARE	0.449	0.199	0.246	0.704		
TOUCH	0.254	0.138	0.148	0.358	0.437	
INTEREST	0.333	0.159	0.183	0.436	0.346	0.583

Means

	NOSAY	VOTING	COMPLEX	NOCARE	TOUCH	INTEREST
	1.212	0.811	0.538	0.864	0.646	0.748

Standard Deviations

	NOSAY	VOTING	COMPLEX	NOCARE	TOUCH	INTEREST
	0.957	0.860	0.633	0.839	0.661	0.764

The covariance matrix obtained in PRELIS Run 5 is just a scaling of the matrix of polychoric correlations obtained in PRELIS Run 4 using the standard deviations from Run 5 as scale factors. These scale factors are also used in the computation of the asymptotic covariance matrix. The mean vector and covariance matrix can be used in LISREL in the same way as for continuous variables. The alternative parameterization is most useful in combination with equal thresholds or fixed thresholds. This will be the topics in Sections 3 and 4.

For the moment it is sufficient to do the following. Replace the two lines in ORD23.SPL, see file ORD25.SPL

```
Correlation Matrix from File EFFICACY.PM
Asymptotic Covariance Matrix from File EFFICACY.ACP
```

with

```
Covariance Matrix from File EFFICACY.CM
Asymptotic Covariance Matrix from File EFFICACY.ACC
```

Alternatively, you can replace ORD21.DSF in ORD23A.SPL with ORD24.DSF, see ORD25A.SPL.

Verify that the factor loadings and error variances are different but the standardized solution and all fit statistics are the same.

3 Longitudinal Data

In the two previous sections I explained how one can analyze ordinal variables in cross-sectional studies.

This section considers models for analyzing data from longitudinal studies where the same individuals are observed or measured at two or more occasions. Such studies are very common in psychology, education, sociology and other fields. Many papers have considered the specification of models incorporating causation and measurement errors in the analysis of data from panel studies, and statistical models and methods for analysis of longitudinal data.

The characteristic feature of a longitudinal research design is that the same measurement instruments are used on the same people at two or more occasions. The purpose of a longitudinal or panel study is to assess the changes that occur between the occasions, and to attribute these changes to certain background characteristics and events existing or occurring before the first occasion and/or to various treatments and developments that occur after the first occasion. Often, when the same variables are used repeatedly, there is a tendency for the measurement errors in these variables to correlate over time because of specific factors, memory or other retest effects. Hence there is a need to consider models with correlated measurement errors.

The analysis of *ordinal variables* in longitudinal studies requires special techniques and procedures which are different from those used with continuous longitudinal variables. This section illustrates these techniques and procedures using the Political Action Panel Study for the USA which is a two-wave panel study. The model considered here is an extension of a model developed by Aish & Jöreskog (1990). The original USA sample consisted of 1719 cases interviewed in 1974. Five years later 933 of these cases were reinterviewed using the same six political efficacy items that were analyzed in Section 2.

In order to estimate differences in means and variances of latent variables over time one must ascertain that the latent variables are on the same scale at different occasions. Both the origin and the unit of measurement must be the same over time. If the observed indicators are continuous, this can be achieved by anchoring each latent variable in one of its observed indicators, a so called reference variable, and by assuming that the mean of the latent variable is zero at one occasion, *e.g.*, the first. By choosing the same reference variable at all occasions one can ascertain that this latent variable is on the same scale over time. However, if the observed indicators are ordinal this is not sufficient, for ordinal variables do not have metric scales so it is meaningless to say that they are on the same scale over time, see Section 2. Again we must use the underlying variables instead of the observed ordinal variables. The underlying variables can be put on the same scale by assuming *equal thresholds* for the underlying variables of the *same* ordinal variable across time.

3.1 Equal Thresholds

Consider k ordinal variables z_1, z_2, \dots, z_k with m_1, m_2, \dots, m_k categories, respectively. Assume that these ordinal variables have been measured on N individuals at T occasions, as in a longitudinal study. Denote by $\pi_{it,c}$ the probability of a response in category c of variable i at time t , $c = 1, 2, \dots, m_i$, $i = 1, 2, \dots, k$, $t = 1, 2, \dots, T$. Let

$$\tau_{i,1} < \tau_{i,2} < \dots < \tau_{i,m_i-1}$$

be a set of thresholds for variable i assumed to be the same at all occasions. Assuming that the underlying variable z_{it}^* is normally distributed with mean μ_{it} and standard deviation σ_{it} at time t , the probability of a response in category c of variable i at time t is

$$\pi_{it,c} = \int_{(\tau_{i,c-1}-\mu_{it})/\sigma_{it}}^{(\tau_{i,c}-\mu_{it})/\sigma_{it}} \phi(u) du, \quad (11)$$

where ϕ is the standard normal density function. Consider a given variable i with m categories. Omitting index i , equation (11) becomes

$$\pi_{t,c} = \int_{(\tau_{c-1}-\mu_t)/\sigma_t}^{(\tau_c-\mu_t)/\sigma_t} \phi(u) du. \quad (12)$$

The probability of a response in category a or lower is

$$\pi_{t,a}^* = \sum_{c=1}^a \pi_{t,c} = \int_{-\infty}^{(\tau_a-\mu_t)/\sigma_t} \phi(u) du = \Phi\left(\frac{\tau_a - \mu_t}{\sigma_t}\right), \quad (13)$$

where Φ is the distribution function of the standard normal distribution. There are two fundamental indeterminacies in (13). One can add a constant to all the τ 's and to μ and one can multiply all the τ 's, μ , and σ by a non-zero constant without altering the right hand side of the equation. In other words, a linear transformation of z^* changes the τ 's, μ , and σ while retaining normality and leaving (13) unchanged. This is a reflection of the fact that although the scale of z_t^* is the same over time, the origin and unit of measurement are still arbitrary. Solving for $(\tau_a - \mu_t)/\sigma_t$, equation (13) can be written

$$\frac{\tau_a - \mu_t}{\sigma_t} = \Phi^{-1}(\pi_{a,t}^*), \quad (14)$$

or equivalently,

$$\tau_a = \mu_t + \sigma_t \Phi^{-1}(\pi_{a,t}^*) = \mu_t + \sigma_t \tau_{a,t}^*, \quad (15)$$

where Φ^{-1} is the inverse function of Φ . The quantity $\Phi^{-1}(\pi_{a,t}^*)$ on the right side of (14) and in the middle of (15) is the unconstrained threshold $\tau_{a,t}^*$ determined for each variable from the univariate marginal probabilities at each time. It is convenient to refer to the τ_a on the left side of equation (15) as the constrained threshold. Equation (15) represents a set of constraints on μ_t and σ_t because the right hand side varies with t whereas the left hand side does not. If $m \geq 3$, the common thresholds, μ_t and σ_t can be estimated from the univariate marginal data of those variables whose thresholds are supposed to be equal. If $m = 2$, *i.e.*, if a variable is dichotomous, only one of μ_t or σ_t can be estimated. In this case, PRELIS sets $\sigma_t = 1$ and estimates μ_t .

To identify the parameters, the origin and the unit of measurement of the common scale must be fixed. In the standard parameterization, this is done such that $\sum_{t=1}^T \hat{\mu}_t = 0$ and $\sum_{t=1}^T \hat{\sigma}_t^2 = T$, *i.e.*, the average mean is 0 and the average variance is 1. In the alternative parameterization, the scale is fixed by $\tau_1 = 0$ and $\tau_2 = 1$.

For $T = 2$ and the standard parameterization, equal thresholds is illustrated in the following table

Variable	Mean	St.Dev.	Thresholds				
1	μ_1	σ_1	τ_1	τ_2	τ_3	\dots	τ_{m-1}
2	μ_2	σ_2	τ_1	τ_2	τ_3	\dots	τ_{m-1}

where $\mu_1 + \mu_2 = 0$ and $\sigma_1^2 + \sigma_2^2 = 2$.

For $T = 2$ and the alternative parameterization, equal thresholds is illustrated in the following table

Variable	Mean	St.Dev.	Thresholds				
1	μ_1^*	σ_1^*	0	1	τ_3^*	\dots	τ_{m-1}^*
2	μ_2^*	σ_2^*	0	1	τ_3^*	\dots	τ_{m-1}^*

The alternative parameterization is obtained from the standard parameterization as follows.

1. For each variable, subtract τ_1 from all the thresholds and from μ .
2. For each variable, divide all thresholds and μ and σ by $\tau_2 - \tau_1$.

The estimated means and standard deviations of the underlying variables can be used to form an estimated mean vector and covariance matrix for all variables. These matrices can be used in LISREL in the usual way. The procedures are illustrated in the following sections.

3.2 A Two Variables Example of Equal Thresholds

First I will illustrate equal thresholds by means of the two variables data that I introduced in Section 2. It consists of 5000 cases on two variables on a five point scale. To estimate the thresholds and the polychoric correlation under the condition of equal thresholds and the standard parameterization, use the following PRELIS command file (file `TWOVARS31.PR2`):

Illustration of Equal Thresholds

```
DA NI=3
RA=TWOVARS.FRQ
LA
FREQ X1 X2
WEIGHT FREQ
ET X1 X2
OU MA=PM
```

This gives the following marginal parameters

Variable	Mean	St. Dev.	Thresholds			
-----	----	-----	-----	-----	-----	-----
X1	-0.498	1.003	-1.304	-0.290	0.388	1.492
X2	0.498	0.997	-1.304	-0.290	0.388	1.492

Note that

- The thresholds are equal
- The sum of the means is 0
- The sum of squares of the standard deviations is 2

To estimate the marginal parameters under the alternative parameterization one can either put AP on the OU line or replace MA=PM by MA=CM, see file TWOVARS32.PR2. This gives the following univariate marginal parameters

Variable	Mean	St. Dev.	Thresholds			
-----	----	-----	-----	-----	-----	-----
X1	0.795	0.989	0.000	1.000	1.669	2.757
X2	1.777	0.983	0.000	1.000	1.669	2.757

The thresholds are still equal but the second and third property no longer holds. The underlying variables are now on a different scale.

To obtain a test of the hypothesis of equal thresholds, one can run TWOVARS31.PR2 with and without the ET line and compute the difference in chi-square for test of underlying bivariate normality, see Section 2. With ET this gives a chi-square of 18.411 with 19 degrees of freedom. Without ET this gives a chi-square of 17.249 with 15 degrees of freedom. The test of equal thresholds gives a chi-square of $18.411 - 17.249 = 1.162$ with $19 - 15 = 4$ degrees of freedom. Hence, the hypothesis of equal thresholds cannot be rejected.

3.3 Estimating the Mean Vector and Covariance Matrix under Equal Thresholds

Most of the remaining part of this section will be devoted to the analysis of the USA panel data of the political efficacy items. The six political efficacy items NOSAY, VOTING, COMPLEX, NOCARE, TOUCH, and INTEREST and their response category codings were introduced in Section 2. The difference now is that I will be analyzing the panel data consisting of the 933 cases answering the same items at two points in time with a five year time lag. The panel data in free format is available in the file PANUSA.RAW in the PR2EX subdirectory. There are 12 variables. The first six are responses to the efficacy items at time 1 and the second are responses to the same items at time 2.

The PRELIS command file below (ORD31.PR2) will do the following

- Eliminate the VOTING item since this will not be used, see Section 2 for the reason for this.

- Impute missing values of the time 2 variables using time 1 variables as matching variables (for information on imputation, see Section 1).
- Estimate the mean vector and the covariance matrix of the underlying variables under the condition of equal threshold for each item over time.
- The asymptotic covariance matrix of the covariance matrix will also be estimated.
- The mean vector, the covariance matrix, and the asymptotic covariance matrix are saved in files.

```
PANELUSA: PRELIS Run 1
Estimating Mean Vector and Covariance Matrix under Equal Thresholds
DA NI=12 MI=8,9
LA
NOSAY1 VOTING1 COMPLEX1 NOCARE1 TOUCH1 INTERES1
NOSAY2 VOTING2 COMPLEX2 NOCARE2 TOUCH2 INTERES2
RA=PANUSA.RAW
SD VOTING1 VOTING2
IM (NOSAY2 - INTERES2) (NOSAY1 - INTERES1) XN
ET NOSAY1 NOSAY2
ET COMPLEX1 COMPLEX2
ET NOCARE1 NOCARE2
ET TOUCH1 TOUCH2
ET INTERES1 INTERES2
Output MA=CM ME=PANUSA.ME CM=PANUSA.CM AC=PANUSA.ACC
```

After imputation there are 849 cases with complete data on all 10 variables. The univariate marginal parameters are estimated under the alternative parameterization as

Variable	Mean	St. Dev.	Thresholds		
-----	-----	-----	-----	-----	-----
NOSAY1	1.294	0.960	0.000	1.000	2.680
COMPLEX1	0.579	0.631	0.000	1.000	1.677
NOCARE1	0.926	0.793	0.000	1.000	2.351
TOUCH1	0.709	0.674	0.000	1.000	2.141
INTERES1	0.821	0.754	0.000	1.000	2.383
NOSAY2	1.326	0.780	0.000	1.000	2.486
COMPLEX2	0.643	0.545	0.000	1.000	1.667
NOCARE2	0.954	0.638	0.000	1.000	2.214
TOUCH2	0.657	0.574	0.000	1.000	2.144
INTERES2	0.838	0.632	0.000	1.000	2.350

Note that the thresholds are equal for each item over time. The covariance matrix of the 10 variables is estimated as

NOSAY1	.922									
COMPLEX1	.242	.398								
NOCARE1	.451	.213	.628							
TOUCH1	.296	.141	.328	.455						
INTERES1	.363	.175	.403	.357	.569					
NOSAY2	.326	.111	.227	.131	.187	.608				
COMPLEX2	.142	.157	.096	.056	.091	.155	.297			
NOCARE2	.260	.106	.245	.164	.220	.295	.130	.406		
TOUCH2	.158	.073	.145	.142	.185	.172	.072	.235	.330	
INTERES2	.218	.103	.202	.150	.234	.226	.100	.285	.262	.400

By this procedure the mean vector and covariance matrix have been estimated such that the underlying variables of each item are on the same scale across time, at least to a sufficient degree of approximation. One can therefore proceed as if these underlying variables had been measured on the same scale over time. This assumption is fundamental for the rest of the analysis.

3.4 A Panel Model for Efficacy

In this section I develop a panel model for efficacy. The measurement model involves two components of efficacy called **Efficacy** and **Respons**. It is the same as considered in Section 2 and estimated from cross-sectional data. This measurement model is applied at each time point.

The objective of the panel model is to answer such questions as: Has the level of efficacy increased or decreased over time? Has the variance of efficacy increased or decreased over time?

A conceptual path diagram of the panel model for efficacy is shown in Fig. 4.

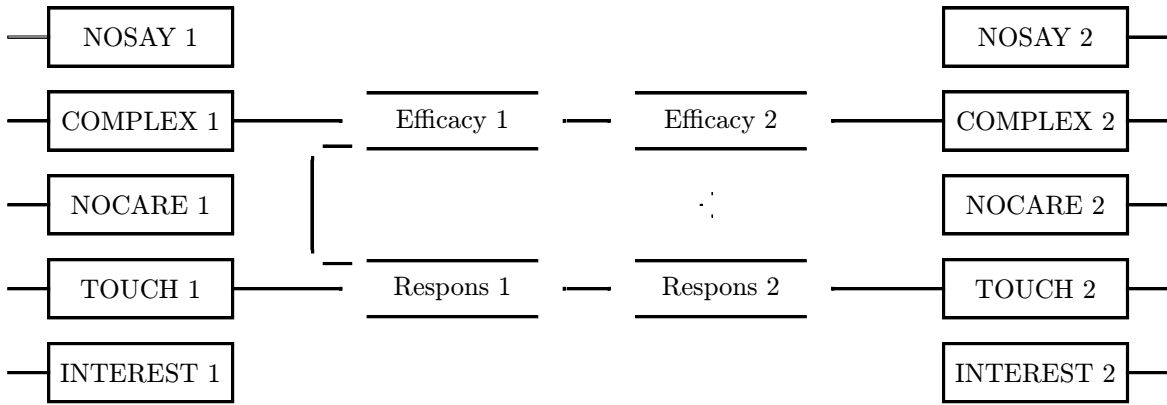


Figure 4: Panel Model for Efficacy and Respons

In the following I refer to the observed variables in the sense of their underlying variables. Thus, when I refer to **COMPLEX1** and **COMPLEX2**, for example, I mean the variables underlying **COMPLEX1** and **COMPLEX2**.

The model also involves a structural model in the middle of the path diagram in which Efficacy at time 2 is predicted by Efficacy at time 1 without the use of Respons at time 1, and Respons at time 2 is predicted by Respons at time 1 without the use of Efficacy at time 1. In addition to these features, the model includes several features not visible in Fig. 4.

- The measurement error in each variable at time 1 correlates with the measurement error in the corresponding variable at time 2 due to a specific factor in each item. To explain this further, I take **COMPLEX** as an example. Let x be **COMPLEX1** and y be **COMPLEX2**. Then the measurement equations for **COMPLEX1** and **COMPLEX2** can be written

$$\text{Time 1 : } x = \lambda_1 \xi + \delta = \lambda_1 \xi + s + d$$

$$\text{Time 2 : } y = \lambda_2 \eta + \epsilon = \lambda_2 \eta + s + e ,$$

where ξ is **Efficac1** and η is **Efficac2**, δ and ϵ are the so called measurement errors in the LISREL model. Each of these error terms are the sum of two components, one specific factor s unique to the item **COMPLEX**, and one pure random error component, d and e , respectively, where d and e are uncorrelated. It follows that δ and ϵ are correlated and that

$$Cov(\delta, \epsilon) = Var(s) .$$

Thus, the specific error variance can be estimated as the autocovariance between the LISREL measurement errors.

- The loading of NOSAY1 on Efficac1 and of NOSAY2 on Efficac2 are fixed to 1 to fix the unit of measurement for Efficac1 and Efficac2. Since NOSAY1 and NOSAY2 have the same unit of measurement (by the construction of equal thresholds), Efficac1 and Efficac2 will also have the same unit of measurement. Similarly, the loadings of INTERES1 on Respons1 and of INTERES2 on Respons2 are fixed to 1 to fix the unit of measurement for Respons1 and Respons2. Since INTERES1 and INTERES2 have the same unit of measurement, Respons1 and Respons2 will also have the same unit of measurement.
- The other four loadings on the latent variables are constrained to be the same across time.
- There is also an intercept term (not visible in the path diagram) in each measurement equation. These intercept terms are also constrained to be equal across time.
- The equality of intercepts and factor loadings across time is necessary in order to compare the latent variables over time on the same scale, *i.e.*, with the same origin and unit of measurement.

3.4.1 Input

A SIMPLIS command file for estimating this panel model is (file ORD32.SPL):

```

SIMPLIS File for Estimating the Panel Model
Observed Variables:                !   1
NOSAY1 COMPLEX1 NOCARE1 TOUCH1 INTERES1      !   2
NOSAY2 COMPLEX2 NOCARE2 TOUCH2 INTERES2      !   3
Means from File PANUSA.ME              !   4
Covariance Matrix from File PANUSA.CM        !   5
Asymptotic Covariance Matrix from File PANUSA.ACC !   6
Sample Size: 832                       !   7
Latent Variables: Efficac1 Respons1 Efficac2 Respons2 !   8
Relationships                          !   9

      NOSAY1 - NOCARE1 = CONST Efficac1      !  10
      NOCARE1 - INTERES1 = CONST Respons1    !  11
      NOSAY1 = 1*Efficac1                    !  12
      INTERES1 = 1*Respons1                   !  13

      NOSAY2 - NOCARE2 = CONST Efficac2      !  14
      NOCARE2 - INTERES2 = CONST Respons2    !  15
      NOSAY2 = 1*Efficac2                    !  16
      INTERES2 = 1*Respons2                   !  17

Let the errors of NOSAY1 and NOSAY2 correlate !  18
Let the errors of COMPLEX1 and COMPLEX2 correlate !  19
Let the errors of NOCARE1 and NOCARE2 correlate !  20
Let the errors of TOUCH1 and TOUCH2 correlate !  21
Let the errors of INTERES1 and INTERES2 correlate !  22

Set Efficac1 -> COMPLEX1 = Efficac2 -> COMPLEX2 !  23
Set Efficac1 -> NOCARE1 = Efficac2 -> NOCARE2 !  24
Set Respons1 -> NOCARE1 = Respons2 -> NOCARE2 !  25
Set Respons1 -> TOUCH1 = Respons2 -> TOUCH2 !  26

```

```

Set CONST -> NOSAY1 = CONST -> NOSAY2          ! 27
Set CONST -> COMPLEX1 = CONST -> COMPLEX2       ! 28
Set CONST -> NOCARE1 = CONST -> NOCARE2        ! 29
Set CONST -> TOUCH1 = CONST -> TOUCH2         ! 30
Set CONST -> INTERES1 = CONST -> INTERES2     ! 31

      Efficac2 = CONST Efficac1                ! 32
      Respons2 = CONST Respons1               ! 33

Let the errors of Efficac2 and Respons2 correlate ! 34

Path Diagram
End of Problem

```

To refer to different lines in this input file, I have numbered the lines in the right margin.

Lines 1–7 specify the names of variables and the data. One can replace all these lines by the single line (see file ORD32A.SPL):

System File from File ORD31.DSF

The system file ORD31.DSF, obtained by running ORD31.PR2, has all the information about the variables and the data, even the location of the asymptotic covariance matrix.

The variable CONST is a variable which is equal to 1 for every case. This variable is always available in SIMPLIS; it need not be in the data. It is used to estimate an intercept term or a mean. For example,

$Y = \text{CONST } X$

is used to specify the regression of Y on X:

$$Y = \alpha + \gamma X .$$

α is the coefficient of CONST just like γ is the coefficient of X. One can also use CONST to estimate a mean. For example,

$Y = \text{CONST}$

will estimate the mean of Y as the coefficient of CONST. Y and X can be any variables, observed or latent. For further examples of CONST, see Jöreskog & Sörbom (1999b), Chapter 2.

Lines 10–11 give the measurement model at time 1 and lines 12–13 specify NOSAY1 and INTERES1 as reference variables for Efficac1 and Respons1, respectively. Similarly lines 14–15 give the measurement model at time 2 and lines 16–17 specify the corresponding reference variables at time 2. Note that all the measurement equations include intercept terms. Also note that NOCARE loads on both Efficacy and Respons.

Lines 18–22 specify the autocorrelated measurement errors needed to estimate the specific factor in each item.

Lines 23–26 constrain the factor loadings to be the same at time 1 and time 2. Here Efficac1 -> COMPLEX1, for example, is short for “the path from Efficac1 to COMPLEX1.” There are 4 factor loadings that are supposed to be equal. The other two loadings at each time point are fixed at 1.

Lines 27–31 constrain the intercepts in the measurement equations to be equal at time 1 and time 2.

Lines 32–33 give the structural equations. The idea is that Efficacy at time 2 is predicted by Efficacy at time 1 without the use of Respons at time 1, and Respons at time 2 is predicted by Respons at time 1 without the use of Efficacy at time 1. These structural equations also include

intercept terms. As will be seen later, these can be interpreted as the mean difference in the latent variables.

Line 34 specifies that the two error terms in the structural equations are allowed to be correlated. This is my way of saying that I do not think that the correlation between **Efficacy** and **Respons** at time 2 can be entirely explained by **Efficacy** and **Respons** at time 1. There may be many variables out there in the world that influences **Efficacy** and **Respons**. One such variable might be *interest in politics*, for example. If there is such a variable that affects all four latent variables but is not included in the model it implies that **Efficacy** and **Respons** will be correlated at time 2 even after controlling for the effects of **Efficacy** and **Respons** at time 1.

3.4.2 Output

The output from ORD32.SPL or ORD32A.SPL gives the following fit statistics

Degrees of Freedom = 31
Minimum Fit Function Chi-Square = 21.85 (P = 0.89)

indicating that the model fits quite well. The model was fitted by WLS using the inverse of the asymptotic covariance matrix as a weight matrix. The chi-square is a C1 in the sense of Jöreskog, *et al.* (2001), Chapter 4 and Appendix A. As explained in Section 2, one can also use ML to fit the model and the asymptotic covariance matrix to correct the chi-square for non-normality, see file ORD32B.SPL. This gives a C3 as

Degrees of Freedom = 31
Satorra-Bentler Scaled Chi-Square = 24.99 (P = 0.77)

Note that these two ways of fitting the model gives approximately the same chi-square.

In the following I present results from the output of ORD32.SPL or ORD32A.SPL. The measurement equations are estimated as

```

NOSAY2 = 1.29 + 1.00*Efficac2, Errorvar.= 0.25
          (0.045)                               (0.043)
          28.66                                5.77

COMPLEX2 = 0.60 + 0.44*Efficac2, Errorvar.= 0.23
           (0.042) (0.034)                               (0.014)
           14.35  12.87                                15.81

NOCARE2 = 0.93 + 0.44*Efficac2 + 0.58*Respons2, Errorvar.= 0.11
          (0.051) (0.086) (0.083)                               (0.020)
          18.34  5.03  7.00                                5.66

TOUCH2 = 0.69 + 0.80*Respons2, Errorvar.= 0.12
         (0.058) (0.028)                               (0.016)
         11.89  29.06                                7.54

INTERES2 = 0.83 + 1.00*Respons2, Errorvar.= 0.069
          (0.060)                               (0.021)
          13.83                                3.27

NOSAY1 = 1.29 + 1.00*Efficac1, Errorvar.= 0.37
          (0.045)                               (0.063)
          28.66                                5.82

```

COMPLEX1 = 0.60 + 0.44*Efficac1, Errorvar.= 0.29
 (0.042) (0.034) (0.020)
 14.35 12.87 14.51

NOCARE1 = 0.93 + 0.44*Efficac1 + 0.58*Respons1, Errorvar.= 0.17
 (0.051) (0.086) (0.083) (0.030)
 18.34 5.03 7.00 5.84

TOUCH1 = 0.69 + 0.80*Respons1, Errorvar.= 0.17
 (0.058) (0.028) (0.023)
 11.89 29.06 7.60

INTERES1 = 0.83 + 1.00*Respons1, Errorvar.= 0.13
 (0.060) (0.029)
 13.83 4.32

It is seen the measurement model is the same at both time points and that all parameters are statistically significant.

The error covariances (specific factors) are estimated as

Error Covariance for NOSAY1 and NOSAY2 = 0.043
 (0.031)
 1.38

Error Covariance for COMPLEX1 and COMPLEX2 = 0.10
 (0.013)
 7.96

Error Covariance for NOCARE1 and NOCARE2 = 0.021
 (0.012)
 1.66

Error Covariance for TOUCH1 and TOUCH2 = 0.013
 (0.010)
 1.21

Error Covariance for INTERES1 and INTERES2 = 0.015
 (0.012)
 1.23

All the estimates of error covariances are positive which is in line with the interpretation of them as variances of the specific factors. However, only the specific error variance of COMPLEX is statistically significant. This does not mean that the other specific factors do not exist, only that they are smaller than that of COMPLEX and that the sample is not large enough to make them significant.

The structural equations are estimated as

Efficac2 = 0.050 + 0.51*Efficac1, Errorvar.= 0.22
 (0.069) (0.038) (0.030)
 0.72 13.54 7.09

Respons2 = - 0.011 + 0.50*Respons1, Errorvar.= 0.22
 (0.059) (0.032) (0.017)
 -0.19 15.94 12.51

This shows that the stability coefficients are statistically significant suggesting that the latent variables at time 2 can be predicted from those of time 1, to some extent. But as indicated by the R^2 's the predictions are not very accurate. Other variables outside of the model may be needed to make these predictions more accurate. The covariance between the two error terms is estimated as

$$\begin{aligned} \text{Error Covariance for Respons2 and Efficac2} &= 0.14 \\ &(0.018) \\ &7.98 \end{aligned}$$

indicating that this is highly significant.

The means of the latent variables cannot be determined on an absolute scale. Although **Efficac1** and **Efficac2** are on the same scale, the origin of the scale is undetermined. We can only estimate the mean difference between **Efficac2** and **Efficac1** and between **Respons2** and **Respons1**. We can fix the origin of the **Efficacy** scale at the mean of **Efficac1** and then estimate the mean of **Efficac2**. Similarly, we can fix the origin of the **Respons** scale at the mean of **Respons1** and estimate the mean of **Respons2**. By this convention (or identification condition) the mean differences equal the intercept terms in the structural equations. It is seen that both of these are non-significant indicating no change in level of the two latent variables over time. A larger sample is needed to be able to tell whether there is a change in level over time. If the intercept terms had been significant there is probably an increase rather than a decrease over time.

Further information about the four latent variables in the output is summarized in Tables 4 and 5.

Table 4: Estimated Means and Covariance Matrix for Efficacy

	Efficacy1	Efficacy2	Means
Efficacy1	0.55		0.00
Efficacy2	0.28	0.36	0.05

Table 5: Estimated Means and Covariance Matrix for Respons

	Respons1	Respons2	Means
Respons1	0.04		0.00
Respons2	0.22	0.33	-0.01

The variance of both **Efficacy** and **Respons** has decreased over time indicating that the population is more homogenous at time 2 than at time 1. The question is whether this decrease is statistically significant. It is not possible to test the hypothesis of equal variances of the latent variables over time using this model. One needs to use another parameterization for this. I will consider this in Section 3.4.4.

3.4.3 Testing Sequence

Some readers may say: How can I put up this model directly without first testing all the assumptions it is based on? Obviously, since the model fits the data well and all results make good sense, there is no need to test all the intermediate steps. However, in general this is a good idea. For instructional purposes, I will therefore go over the steps here. I leave it to the reader to actually do these steps.

There are several ways to do the tests. I recommend the following sequential testing procedure.

1. Test the measurement model at each time point separately. To test the measurement model at time 1, include only the lines 10–13. To test the measurement model at time 2, include only the

lines 14–17. Selection of variables is automatic in SIMPLIS; only the variables included in the model will be used. If the measurement model fits at both occasions one can continue with the next step. If the measurement model fits at one occasion but not at the other occasion it means that the measurement models are not functionally equivalent, *i.e.*, the observed variables do not measure the same latent variables at each occasion. Then further analysis is meaningless. If the measurement models fail to fit at both occasions, both models should be modified and their functional equivalence should be tested again.

2. Test the two measurement models jointly. This is done by including only the lines 10–17. The covariance matrix of the four latent variables is left unconstrained. If this model does not fit (which is the case here), one must introduce the specific factors by adding the lines 18–22. If, after adding the specific factors, the model still does not fit, there must be something fundamentally wrong. This should not happen if Step 1 has been satisfactorily resolved.
3. Test the equality of factor loadings over time. This is done by adding the lines 23–26. The test of equality of factor loadings is obtained by computing the difference in chi-squares between this model and the previous one. If the hypothesis of equal factor loadings is rejected it will not be possible to compare the latent variables over time with the same unit of measurement.⁸
4. Test the equality of intercept terms over time. This is done by adding lines 27–31. The test of equality of intercept terms is obtained by computing the difference in chi-squares between this model and the previous one. If the hypothesis of equal intercept terms is rejected it will not be possible to compare the latent variables over time with the same origin, but see footnote 8.
5. Test the structural model. This is done by adding lines 32–34. This test is obtained by computing the difference in chi-squares between this model and the previous one. If the structural model does not fit, replace the two lines

```

Efficac2 = CONST Efficac1           ! 32
Respons2 = CONST Respons1          ! 33

```

with

```

Efficac2 = CONST Efficac1 Respons1
Respons2 = CONST Efficac1 Respons1

```

This model is equivalent to the model in Step 4. So it will fit if that model fits.

3.4.4 Testing Equality of Variances

In Section 3.4.2 I found that the factor variances decreased over time but I was unable to test whether this decrease is statistically significant. In this section I show how this can be done by a slight modification of the input.

Replace the lines

```

Efficac2 = CONST Efficac1           ! 32
Respons2 = CONST Respons1          ! 33

```

```

Let the errors of Efficac2 and Respons2 correlate          ! 34

```

with (see file ORD33.SPL)

⁸Under partial invariance of factor loadings, *i.e.*, if some factor loadings are equal while others are different, such a comparison may still be possible, but this is very tricky.

```

Efficac1 = 0*CONST
Respons1 = 0*CONST
Efficac2 = CONST
Respons2 = CONST

```

This will leave the covariance matrix of the four latent variables unconstrained but the model is otherwise the same as before. In fact, the model is the same as in Step 4 of the previous section. This model has a chi-square of 22.60 with 29 degrees of freedom.

A test of equality of factor variances can now be obtained by adding the lines (see `ORD33A.SPL`):

```

Equal Variances: Efficac1 Efficac2
Equal Variances: Respons1 Respons2

```

This model has a chi-square of 64.83 with 31 degrees of freedom. The difference in chi-squares is 42.23 with 2 degrees of freedom which is highly significant. The hypothesis of equal variances must be rejected. Thus, we can conclude that the factor variances have decreased over time. This could be tested separately for `Efficacy` and `Respons`.

3.4.5 Error Variances and Reliabilities

In Steps 3 and 4 of Section 3.4.3 I discussed functional equivalence of measurement models across time. I did not include equal error variances in this concept. Equal error variances is not necessary to compare latent variables over time. However, if the goal is to have the most parsimonious model, it may be of interest to test the hypothesis of equal error variances over time. To do so, add the lines

```

Equal Error Variances: NOSAY1 NOSAY2
Equal Error Variances: COMPLEX1 COMPLEX2
Equal Error Variances: NOCARE1 NOCARE2
Equal Error Variances: TOUCH1 TOUCH2
Equal Error Variances: INTERES1 INTERES2

```

in `ORD32.SPL` or `ORD32A.SPL`, see file `ORD34.SPL`. This gives a chi-square of 45.93 with 36 degrees of freedom. This should be compared with 27.47 with 31 degrees of freedom for the original panel model. The difference in chi-squares is 18.46 with 5 degrees of freedom. This difference is significant at the 1% level. So the hypothesis of equal error variances is rejected. However, this hypothesis can be tested for each item separately. The test for `NOSAY` gives (see file `ORD34A.SPL`) a chi-square of 29.92 with 32 degrees of freedom. The chi-square difference is 2.45 with 1 degree of freedom. Hence, the hypothesis of equal error variances for `NOSAY` cannot be rejected.

The fact that the error variances of `NOSAY1` and `NOSAY2` are equal does not imply that their reliabilities are equal. I found in Section 3.4.4 that the variance of `Efficacy` decreased over time. As a consequence of this, the fitted variance of `NOSAY1` and `NOSAY2` will be different.

To estimate the reliabilities of `NOSAY1` and `NOSAY2` proceed as follows.

- From the output of `ORD34A.SPL` find the error variance of `NOSAY` and the specific variance of `NOSAY`. These are 0.31 and 0.031, respectively. Thus, the pure random measurement error variance of `NOSAY` is $0.31 - 0.03 = 0.28$. This is the same for `NOSAY1` and `NOSAY2`.
- Next find the fitted variances of `NOSAY1` and `NOSAY2` (these are obtained if the line `Print Residuals` is included in the input). These are 0.87 and 0.65, respectively.
- The reliability of `NOSAY1` is now computed as $1 - 0.28/0.87 = 0.68$. For `NOSAY2` I get the reliability $1 - 0.28/0.65 = 0.57$.

3.4.6 LISREL Notation

Some readers may be interested in doing the modeling using LISREL notation and syntax. Those who are not interested in this can skip this section.

Let

$$\begin{aligned}\mathbf{x} &= (\text{NOSAY1}, \text{COMPLEX1}, \text{NOCARE1}, \text{TOUCH1}, \text{INTERES1}), \\ \mathbf{y} &= (\text{NOSAY2}, \text{COMPLEX2}, \text{NOCARE2}, \text{TOUCH2}, \text{INTERES2}), \\ \boldsymbol{\xi} &= (\text{Efficac1}, \text{Respons1}), \\ \boldsymbol{\eta} &= (\text{Efficac2}, \text{Respons2}).\end{aligned}$$

Then the panel model in LISREL notation is (see Jöreskog & Sörbom, 1999b, Chapter 10 or Jöreskog, *et al.*, 2001, Chapter 1).

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & \lambda_{31} \\ 0 & \lambda_{41} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & \lambda_{31} \\ 0 & \lambda_{41} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{pmatrix}$$

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}$$

These correspond to the three general equations in LISREL.

$$\mathbf{x} = \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta},$$

$$\mathbf{y} = \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon},$$

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}.$$

3.4.7 LISREL Syntax

An input file in LISREL syntax corresponding to ORD32A.SPL is (file ORD35.LS8)

```
LISREL File for Estimating the Panel Model
SY=ORD31.DSF
SE
6 7 8 9 10 1 2 3 4 5
MO NY=5 NX=5 NE=2 NK=2 GA=FI PS=SY,FR TH=FI TX=FR TY=FR AL=FR
LK Efficac1 Respons1
LE Efficac2 Respons2
VA 1 LY(1,1) LY(5,2) LX(1,1) LX(5,2)
FR LX(2,1) LX(3,1) LX(3,2) LX(4,2)
FR GA(1,1) GA(2,2)
FR TH(1,1) TH(2,2) TH(3,3) TH(4,4) TH(5,5)
EQ LX(2,1) LY(2,1)
EQ LX(3,1) LY(3,1)
EQ LX(3,2) LY(3,2)
```

```
EQ LX(4,2) LY(4,2)
EQ TX(1) TY(1)
EQ TX(2) TY(2)
EQ TX(3) TY(3)
EQ TX(4) TY(4)
EQ TX(5) TY(5)
OU
```

The lines

```
SE
6 7 8 9 10 1 2 3 4 5
```

are needed to order the variables such that the y -variables come first, as required by LISREL. The rest of the output is self-explanatory for users who are familiar with LISREL syntax. It can be verified that the results from ORD35.LS8 are the same as for ORD32A.SPL.

3.5 Four-Wave Models

The previous two-wave model can be generalized to the multiwave situation when the same ordinal variables are used at more than two occasions. I do not have access to any multiwave ordinal data, so I will use generated data on four occasions. I will pretend that the variables are the six efficacy items (this time I include VOTING) measured on a four-category scale with categories having the same meaning, see Section 2.

3.5.1 A Four-Wave Model with a Single Latent Variable

The data for this example is in the file EFFI4WS.RAW in free format. There are $3 \times 4 = 12$ variables, where the first 3 correspond to Time 1, the next 3 to Time 2, etc.

PRELIS Step I begin by estimating the mean vector, the covariance matrix, and the asymptotic covariance matrix of the variables underlying the ordinal variables under equal thresholds for each item over time. The PRELIS command file for this is (file EFFI4WS1.PR2):

```
Computing ME, CM, and ACC for EFFI4WS.RAW
Under Equal Thresholds over Time
DA NI=12
LA
NOSAY1 VOTING1 COMPLEX1 NOSAY2 VOTING2 COMPLEX2
NOSAY3 VOTING3 COMPLEX3 NOSAY4 VOTING4 COMPLEX4
RA=EFFI4WS.RAW
CL NOSAY1 - COMPLEX4 1=AS 2=A 3=D 4=DS
ET NOSAY1 NOSAY2 NOSAY3 NOSAY4
ET VOTING1 VOTING2 VOTING3 VOTING4
ET COMPLEX1 COMPLEX2 COMPLEX3 COMPLEX4
OU MA=CM ME=EFFI4WS.ME CM=EFFI4WS.CM AC=EFFI4WS.ACC
```

The output reveals that there are 2357 cases and that the univariate marginal parameters are

Variable	Mean	St. Dev.	Thresholds		
-----	----	-----	-----	-----	-----
NOSAY1	-0.032	1.090	0.000	1.000	2.058
VOTING1	-0.003	0.935	0.000	1.000	1.850
COMPLEX1	-0.008	1.001	0.000	1.000	1.946

NOSAY2	0.818	1.192	0.000	1.000	2.058
VOTING2	0.582	1.019	0.000	1.000	1.850
COMPLEX2	0.599	1.097	0.000	1.000	1.946
NOSAY3	2.035	1.399	0.000	1.000	2.058
VOTING3	1.402	1.134	0.000	1.000	1.850
COMPLEX3	1.455	1.204	0.000	1.000	1.946
NOSAY4	3.800	1.666	0.000	1.000	2.058
VOTING4	2.470	1.199	0.000	1.000	1.850
COMPLEX4	2.736	1.430	0.000	1.000	1.946

Note that there is an increasing trend in the means of the underlying variables over time.

LISREL Step The measurement model assumes that NOSAY, VOTING, and COMPLEX are indicators of a single latent variable *Efficacy*. This measurement model is employed at each time point and is assumed to be invariant over time. Both the intercepts and the factor loadings are assumed to be the same across time. In addition, it is assumed that VOTING and COMPLEX, but not NOSAY, contain specific factors denoted *Voting* and *Complex*. These specific factors are constant over time and uncorrelated with *Efficacy* at all times.

First I focus on the estimation of the means and variances of *Efficacy* over time without assuming any structural model. A SIMPLIS command file to estimate such a model is (file EFFI4WS2.SPL):

```

SIMPLIS Input for Four-Wave Model
One Factor
Observed Variables:
NOSAY1 VOTING1 COMPLEX1 NOSAY2 VOTING2 COMPLEX2
NOSAY3 VOTING3 COMPLEX3 NOSAY4 VOTING4 COMPLEX4

Means from File EFFI4WS.ME
Covariance Matrix from File EFFI4WS.CM
Asymptotic Covariance Matrix from File EFFI4WS.ACC
Sample Size: 2357
Latent Variables: Efficac1 Efficac2 Efficac3 Efficac4
                  Voting Complex
Relationships

NOSAY1 - COMPLEX1 = CONST (1)*Efficac1
NOSAY1 = 1*Efficac1

NOSAY2 - COMPLEX2 = CONST (1)*Efficac2
NOSAY2 = 1*Efficac2

NOSAY3 - COMPLEX3 = CONST (1)*Efficac3
NOSAY3 = 1*Efficac3

NOSAY4 - COMPLEX4 = CONST (1)*Efficac4
NOSAY4 = 1*Efficac4

Set CONST -> NOSAY1 = CONST -> NOSAY2
Set CONST -> VOTING1 = CONST -> VOTING2
Set CONST -> COMPLEX1 = CONST -> COMPLEX2

Set CONST -> NOSAY1 = CONST -> NOSAY3
Set CONST -> VOTING1 = CONST -> VOTING3
Set CONST -> COMPLEX1 = CONST -> COMPLEX3

```

```

Set CONST -> NOSAY1 = CONST -> NOSAY4
Set CONST -> VOTING1 = CONST -> VOTING4
Set CONST -> COMPLEX1 = CONST -> COMPLEX4

Set Efficac1 -> VOTING1 = Efficac2 -> VOTING2
Set Efficac1 -> COMPLEX1 = Efficac2 -> COMPLEX2

Set Efficac1 -> VOTING1 = Efficac3 -> VOTING3
Set Efficac1 -> COMPLEX1 = Efficac3 -> COMPLEX3

Set Efficac1 -> VOTING1 = Efficac4 -> VOTING4
Set Efficac1 -> COMPLEX1 = Efficac4 -> COMPLEX4

VOTING1 VOTING2 VOTING3 VOTING4 = 1*Voting          ! A
COMPLEX1 COMPLEX2 COMPLEX3 COMPLEX4 = 1*Complex     ! B

Set the covariances of Voting and Complex to 0

Efficac1 = 0*CONST
Efficac2 = CONST
Efficac3 = CONST
Efficac4 = CONST

Set the covariances of Efficac1 - Efficac4 free
Set the covariances between Voting - Complex and Efficac1 - Efficac4 to 0

End of Problem

```

A difference between this example and the two-wave models considered previously is that the specific factors are now modeled directly as factors affecting the underlying variables rather than as error covariances, see the lines marked A and B. It is possible to specify them as error covariances also but this is tedious in the SIMPLIS command language because one has to define six error covariances for each of VOTING and COMPLEX and one has to specify the equality of these six error covariances as well.

Without (1)* in front of Efficacy in the measurement model, the LISREL iterations do not converge. But this is only a starting value problem. All one needs to do is to give LISREL some help to get iterations started. This is done by specifying 1 as a starting value for the factor loading of Efficacy. Note the difference between (1)* and 1* as in lines A and B. The latter is a fixed value.

I leave it to the reader to contemplate most of the information in the output. Here I will confine myself to giving only the covariance matrix and the mean vector of the latent variables

Covariance Matrix of Independent Variables

	<u>Efficac1</u>	<u>Efficac2</u>	<u>Efficac3</u>	<u>Efficac4</u>	<u>Voting</u>	<u>Complex</u>
Efficac1	0.79 (0.02) 37.14					
Efficac2	0.70 (0.02) 36.30	1.08 (0.02) 49.49				

Efficac3	0.75 (0.02) 30.79	1.14 (0.02) 54.58	1.61 (0.03) 63.22			
Efficac4	0.81 (0.03) 23.89	1.27 (0.03) 41.00	1.74 (0.03) 56.22	2.35 (0.04) 53.68		
Voting	- -	- -	- -	- -	0.21 (0.01) 19.55	
Complex	- -	- -	- -	- -	- -	0.30 (0.01) 21.18

Mean Vector of Independent Variables

Efficac1	Efficac2	Efficac3	Efficac4	Voting	Complex
-----	-----	-----	-----	-----	-----
- -	0.83 (0.03) 32.26	2.04 (0.03) 81.46	3.78 (0.03) 134.78	- -	- -

which shows that there is an increasing trend in both means and variances of the latent variables over time. It also shows that there are specific factors in VOTING and COMPLEX.

A Simplex Model Next I assume that Efficacy is generated by an autoregressive process:

$$\text{Efficacy}_t = \alpha_t^{(E)} + \beta_t^{(E)} \text{Efficacy}_{t-1}, \quad t = 2, 3, 4$$

This kind of model is sometimes called a Simplex model, see, *e.g.*, Jöreskog & Sörbom, 1999b, pp. 230–238.

A SIMPLIS command file to estimate this model is (file EFFI4WS3.SPL):

```
SIMPLIS Input for Four-Wave Model
Single Simplex
...
...
...
Set the covariance between Voting and Complex to 0

Efficac2 = CONST Efficac1
Efficac3 = CONST Efficac2
Efficac4 = CONST Efficac3

No x-variables
Admissibility Check Off
End of Problem
```

The lines down to and including

```
Set the covariance between Voting and Complex to 0
```

are the same as before.

This kind of model is best handled as a model with only y - and η -variables in LISREL (LISREL submodel 3B, see Jöreskog & Sörbom, 1999c, Chapter 6). The SIMPLIS command to specify this is

No x-variables

The line

Admissibility Check Off

is needed to make the iterations converge (LISREL checks the admissibility of parameters after 50 iterations⁹ and stops if this check is not set off, see Jöreskog & Sörbom, 1999c, pp. 322–323).

The chi-square for this model is 119.18 with 61 degrees of freedom, compared with the previous 115.71 with 58 degrees of freedom. The difference 3.47 with 3 degrees of freedom means that the Simplex model cannot be rejected.

3.5.2 A Four-Wave Model with Two Latent Variables

The previous four-wave model can be generalized to a model with six observed and two latent variables. This generalization is straightforward but the input files are very long and elaborate. Rather than listing these files here, I leave it for the interested reader to carry out the analysis.

The data is in the file `EFFI4W.RAW` in free format. There are $6 \times 4 = 24$ variables, where the first 6 correspond to Time 1, the next 6 to Time 2, etc.

PRELIS Step I begin by estimating the mean vector, the covariance matrix, and the asymptotic covariance matrix of the variables underlying the ordinal variables under equal thresholds for each item over time. The PRELIS command file for this is `EFFI4W1.PR2`.

LISREL Step The measurement model is shown in Fig. 5. This measurement model is employed

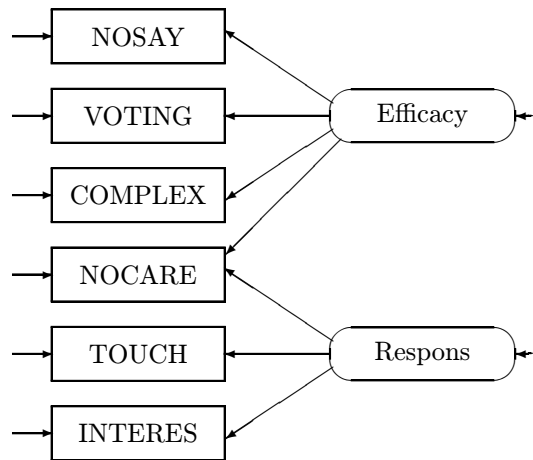


Figure 5: Six Variables Measurement Model for Efficacy and Respons

at each time point and is assumed to be invariant over time. In addition, I assume that `VOTING` and `COMPLEX` contain specific factors denoted `Voting` and `Complex`. These specific factors are constant over time and uncorrelated with the factors `Efficacy` and `Respons` at all times. `Efficacy` and `Respons` are themselves contemporaneously correlated.

⁹The book says 10 but the current version of LISREL uses 50.

The structural model assumes that **Efficacy** and **Respons** are generated by an autoregressive process:

$$\begin{aligned}\text{Efficacy}_t &= \alpha_t^{(E)} + \beta_t^{(E)}\text{Efficacy}_{t-1}, & t = 2, 3, 4 \\ \text{Respons}_t &= \alpha_t^{(R)} + \beta_t^{(R)}\text{Respons}_{t-1}, & t = 2, 3, 4\end{aligned}$$

A SIMPLIS command file to estimate this model is `EFFI4W2.SPL`. This command file is quite large and elaborate. For this kind of models, the LISREL command language is more convenient. Users who have learned to master the LISREL command language can use the shorter input given in `EFFI4W3.LS8`.

4 Multiple Groups

In Section 2 I explained how one can analyze ordinal variables in longitudinal studies. This section considers the situation where data on the same ordinal variables have been collected in several groups. These groups may be different nations, states or regions, culturally or socioeconomically different groups, groups of individuals selected on the basis of some known selection variables, groups receiving different treatments, etc. In fact, they may be any set of mutually exclusive groups of individuals which are clearly defined. It is assumed that the data is a random sample of individuals from each group. The objective is to compare different characteristics across groups. In particular, the procedure to be described can be used for testing factorial invariance and for estimating differences in factor means. For information and examples on how to do this with continuous variables see Chapters 9 and 10 in Jöreskog & Sörbom (1999c) and Chapter 2 in Jöreskog & Sörbom (1999b). Just as in the cases of analysis of cross-sectional data, described in Section 2, and analysis of longitudinal data, described in Section 3, the analysis of ordinal variables in multiple groups requires a PRELIS step before one can proceed to analyze models with LISREL. For further information about PRELIS, see Jöreskog & Sörbom (1999a).

In this section I continue the analysis of the efficacy variables in the Political Action Survey which was carried out in eight countries. For information about this survey and the efficacy variables, see Section 1. In Sections 2 and 3 I analyzed only data from the USA. Here I will analyze the data from all eight countries. The procedure to be described here makes it possible to answer questions like these:

- Do the efficacy items measure the same latent variables in all countries?
- If so, are the factor loadings invariant over countries?
- Are the intercepts invariant over countries?

If these conditions are satisfied one can estimate differences in means, variances, and covariances of the latent variables **Efficacy** and **Respons** between countries. Recall from Section 2 that **Efficacy** and **Respons** are two different components of Political Efficacy, where **Efficacy** indicates *individuals' self-perceptions that they are capable of understanding politics and competent enough to participate in political acts such as voting*, and **Respons** (short for *Responsiveness*) indicates *the belief that the public cannot influence political outcomes because government leaders and institutions are unresponsive*. People who are low on **Efficacy** or low on **Respons** are expected to agree or agree strongly with the items. Hence, the items measure these components from low to high.

Complete factorial invariance over all eight countries should not be expected to hold for the following reasons:

- The items are stated in different languages.
- Words may have different connotations in different languages.

- Other cultural differences between countries may lead to different response styles or response patterns in different countries.

These reasons may imply that the items are interpreted differently in different countries.

The procedure to study factorial invariance with ordinal variables is as follows:

PRELIS Step Define a set of thresholds for each variable to be the same in each country. Since the underlying variables are only determined up to a monotonic transformation, one can simply choose these as 0, 1, 2, ... for all variables. Alternatively, one can use PRELIS to estimate a set of thresholds from the total sample by pooling the data from all groups into one data file. Either way, these thresholds define a scale for the underlying variables common to all groups. Using the thresholds as *fixed thresholds*, PRELIS can estimate the mean vector, the covariance matrix, and the asymptotic covariance matrix of the underlying variables for each group.

LISREL Step These mean vectors, covariance matrices, and asymptotic covariance matrices can be used in a multigroup analysis in LISREL as if the underlying variables had been observed.

4.1 Data Screening

The data for all countries is in the datafile `EFFITOT.RAW` in free format. The first variable is `COUNTRY` coded as 1 = USA (USA), 2 = Germany (GER), 3 = The Netherlands (NET), 4 = Austria (AUS), 5 = Britain (BTN), 6 = Italy (ITY), 7 = Switzerland (SWI), and 8 = Finland (FIN). The other variables are the six efficacy variables described in Section 1. The item `VOTING` is included here but will be eliminated in the LISREL step. The response categories and their codings are those described in Section 1, but in Italy there was an additional response category *Don't Understand* (DU) coded as 6.

A data screening of all the data can be obtained by running the following PRELIS command file (file `ORD41.PR2`):

```
!Data Screening of EFFITOT.RAW
Data Ninputvariables = 7
Labels
COUNTRY NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata = EFFITOT.RAW
Clabels COUNTRY 1=USA 2=GER 3=NET 4=AUS 5=BTN 6=ITY 7=SWI 8=FIN
Clabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS 6=DU 8=DK 9=NA
Output
```

The output gives the numbers in the right and bottom margins of Tables 6 and 7. Just like in the USA sample, there are more people in the *Don't Know* than in the *No answer* categories.

To screen the data for one country, use the following PRELIS command file, here illustrated with the USA (file `ORD42.PR2`):

```
!Data Screening of EFFITOT.RAW for Country USA
Data Ninputvariables = 7
Labels
COUNTRY NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata = EFFITOT.RAW
Clabels COUNTRY 1=BTN 2=GER 3=NET 4=AUS 5=USA 6=ITY 7=SWI 8=FIN
Clabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS 6=DU 8=DK 9=NA
Sdelete COUNTRY = 1
Output
```

To repeat this for another country, just change the line

Table 6: Observed Frequency Distributions

NOSAY

Response Category	USA	GER	NET	AUS	BTN	ITY	SWI	FIN	Total
Agree Strongly	175	721	171	804	211	306	347	295	3030
Agree	518	907	479	432	693	797	463	413	4702
Disagree	857	464	460	215	488	365	314	425	3588
Disagree Strongly	130	133	49	50	33	57	137	50	639
Don't Understand	-	-	-	-	-	74	-	-	74
Don't Know	29	27	40	82	42	173	26	35	454
No Answer	10	3	2	2	16	7	3	6	49
All Responses	1719	2255	1201	1585	1483	1779	1290	1224	12536

VOTING

Response Category	USA	GER	NET	AUS	BTN	ITY	SWI	FIN	Total
Agree Strongly	283	790	195	903	218	341	371	429	3530
Agree	710	861	634	431	884	902	403	512	5337
Disagree	609	448	285	135	289	271	310	234	2581
Disagree Strongly	80	103	38	28	19	45	150	18	481
Don't Understand	-	-	-	-	-	57	-	-	57
Don't Know	26	49	43	85	57	159	53	26	498
No Answer	11	4	6	3	16	4	3	5	52
All Responses	1719	2255	1201	1585	1483	1779	1290	1224	12536

COMPLEX

Response Category	USA	GER	NET	AUS	BTN	ITY	SWI	FIN	Total
Agree Strongly	343	688	262	531	312	484	495	386	3483
Agree	969	801	592	557	777	858	394	588	5536
Disagree	323	516	273	283	310	225	267	198	2395
Disagree Strongly	63	214	45	125	40	55	112	37	691
Don't Understand	-	-	-	-	-	39	-	-	39
Don't Know	9	30	25	86	28	114	18	28	338
No Answer	12	6	4	3	16	4	4	5	54
All Responses	1719	2255	1201	1585	1483	1779	1290	1224	12536

Table 7: Observed Frequency Distributions

NOCARE

Response Category	USA	GER	NET	AUS	BTN	ITY	SWI	FIN	Total
Agree Strongly	250	569	156	590	205	402	354	226	2752
Agree	701	880	487	474	756	853	418	552	5121
Disagree	674	638	421	334	404	286	364	354	3475
Disagree Strongly	57	103	36	86	33	24	82	26	467
Don't Understand	-	-	-	-	-	45	-	-	25
Don't Know	20	57	95	98	68	184	48	60	630
No Answer	17	8	6	3	17	5	8	6	66
All Responses	1719	2255	1201	1585	1483	1779	1290	1224	12536

TOUCH

Response Category	USA	GER	NET	AUS	BTN	ITY	SWI	FIN	Total
Agree Strongly	273	697	197	602	276	481	325	319	3170
Agree	881	978	575	533	737	869	523	578	5674
Disagree	462	425	267	257	347	167	262	230	2417
Disagree Strongly	26	49	25	49	18	15	56	21	259
Don't Understand	-	-	-	-	-	45	-	-	45
Don't Know	60	101	133	140	88	198	116	68	904
No Answer	17	5	4	4	17	4	8	8	67
All Responses	1719	2255	1201	1585	1483	1779	1290	1224	12536

INTEREST

Response Category	USA	GER	NET	AUS	BTN	ITY	SWI	FIN	Total
Agree Strongly	264	541	147	629	280	469	343	331	3004
Agree	762	792	443	458	709	834	432	519	4949
Disagree	581	698	443	309	377	247	318	280	3253
Disagree Strongly	31	149	44	76	24	26	93	28	471
Don't Understand	-	-	-	-	-	24	-	-	24
Don't Know	62	66	120	110	76	172	96	58	760
No Answer	19	9	4	3	17	7	8	8	75
All Responses	1719	2255	1201	1585	1483	1779	1290	1224	12536

```
Sdelete Country = 1
```

The `Sdelete` line (short for Select and Delete) first selects all cases with `COUNTRY = 1` and then deletes the variable `COUNTRY` (it is no use to keep the variable `COUNTRY` after selection of cases since all cases have the same value 1 on this variable).

One can screen the data for each and all countries simultaneously by running a file with stacked input, see file `ORD42A.PR2`. Collecting the results from this output gives the results shown in Tables 6 and 7.

It is seen in Tables 6 and 7 that there are considerable differences between countries in the univariate marginal distribution of these variables but these distributions are rather similar across variables. Countries like the USA and Britain which use the same language (English) are rather similar. Germany and Austria, where the German language was used, are also rather similar. But there is a considerable difference between these two pairs of countries. Most notably is the distribution in Austria where many people respond in the *Agree Strongly* category. We shall see that these manifest differences may be viewed as reflections of differences in the means of the latent variables between countries.

4.2 PRELIS Step

Write a text file of thresholds called `EFFITOT.THR`, say, as follows

```
0 1 2
0 1 2
0 1 2
0 1 2
0 1 2
0 1 2
```

The first line gives the thresholds for `NOSAY`, the second for `VOTING`, etc. After listwise deletion of all responses in the *Don't Understand*, *Don't Know*, and *No Answer* categories, there are four categories on the ordinal scale. Thus there should be three thresholds for each variable. Any set of monotonically increasing thresholds will do, and they do not have to be the same for all variables. But they must be the same for all countries.

In the `PRELIS` step we compute the mean vector, the covariance matrix, and the asymptotic covariance matrix of the underlying variables for each country. For the USA this can be done with the following `PRELIS` command file (file `ORD43.PR2`):

```
Computing mean vector, covariance matrix and asymptotic covariance matrix
for country USA
Data Ninputvariables = 7
Labels
COUNTRY NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Rawdata=EFFITOT.RAW
Sdelete COUNTRY = 1
Missing 6,8,9 NOSAY - INTEREST
CLabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS
FT=EFFITOT.THR NOSAY
FT VOTING
FT COMPLEX
FT NOCARE
FT TOUCH
FT INTEREST
Output MA=CM ME=EFFUSA.ME CM=EFFUSA.CM AC=EFFUSA.ACC
```

In the same way one can obtain the three matrices for all the other countries by changing the country code on the `Sdelete` line and changing the three file names on the output line. File `ORD43A.PR2`, not listed here, shows how this can be done for all countries in one single run.

The listwise sample sizes are given in Table 8. The estimated means and standard deviations of the underlying variables are given in Table 9.

Table 8: Listwise Sample Sizes

USA	GER	NET	AUS	BTN	ITY	SWI	FIN
1554	2062	945	1318	1266	1237	1085	1066

Table 9: Means and Standard Deviations

Means

	USA	GER	NET	AUS	BTN	ITY	SWI	FIN
NOSAY	1.018	0.458	0.823	-0.024	0.716	0.609	0.677	0.659
VOTING	0.758	0.361	0.640	-0.292	0.587	0.494	0.640	0.301
COMPLEX	0.586	0.553	0.605	0.452	0.569	0.388	0.425	0.383
NOCARE	0.781	0.589	0.789	0.337	0.686	0.416	0.638	0.633
TOUCH	0.637	0.367	0.601	0.206	0.557	0.245	0.508	0.425
INTEREST	0.691	0.661	0.822	0.222	0.580	0.308	0.578	0.462

Standard Deviations

	USA	GER	NET	AUS	BTN	ITY	SWI	FIN
NOSAY	0.729	0.986	0.742	1.112	0.654	0.770	1.082	0.811
VOTING	0.760	0.989	0.731	1.081	0.628	0.744	1.187	0.788
COMPLEX	0.755	1.102	0.803	1.122	0.736	0.841	1.207	0.839
NOCARE	0.700	0.858	0.691	1.097	0.669	0.704	0.994	0.691
TOUCH	0.642	0.815	0.697	0.994	0.666	0.721	0.880	0.739
INTEREST	0.649	0.926	0.705	1.115	0.690	0.791	1.016	0.790

Table 9 shows that there are considerable differences in the means between countries. Note particularly the large differences between the USA and Austria on `NOSAY` and `VOTING`.

4.3 LISREL Step

The model is the same as the one considered in Section 2.¹⁰ This was found to fit the cross-sectional data for the USA very well. A path diagram is shown in Fig. 6.

In standard LISREL notation the model is

$$\mathbf{x}^{(g)} = \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x \boldsymbol{\xi}^{(g)} + \boldsymbol{\delta}^{(g)}, \quad (16)$$

where $\mathbf{x}^{(g)}$ is a vector of the underlying variables of `NOSAY`, `COMPLEX`, `NOCARE`, `TOUCH`, and `INTEREST`

¹⁰The `VOTING` item is not included for reasons explained in Section 2.

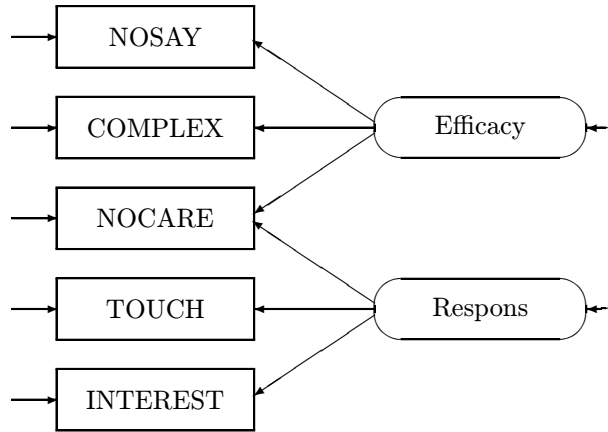


Figure 6: Measurement Model for Efficacy and Respons

in group g , τ_x is a vector of intercepts, Λ_x is the matrix

$$\begin{pmatrix} 1 & 0 \\ \lambda_{21}^{(x)} & 0 \\ \lambda_{31}^{(x)} & \lambda_{32}^{(x)} \\ 0 & \lambda_{42}^{(x)} \\ 0 & 1 \end{pmatrix},$$

$\xi^{(g)}$ is a vector of the latent variables Efficacy and Respons in group g , and $\delta^{(g)}$ is a vector of measurement errors in group g .

The parameter matrices τ_x and Λ_x are regarded as attributes of the variables and are therefore assumed to be invariant over groups. The unit of measurement in the latent variables are defined by the two elements 1 in Λ_x . This makes these units the same across groups which makes it possible to compare the variances and covariances of the latent variables across groups. The fact that τ_x is also invariant over groups makes it possible to estimate differences in the means of the latent variables between groups. To do so, it is convenient to fix the mean of ξ to $\mathbf{0}$ in the first group and estimate the mean of ξ in the other groups. Any group can be chosen as the first group.

USA vs Britain

I begin with the comparison of the two countries USA and Britain where the same language, English, was used in the survey.¹¹

A SIMPLIS command file for analysis of these two samples is (file ORD44.SPL)

```
Group USA
Observed Variables: NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Means from File EFFUSA.ME
Covariance Matrix from File EFFUSA.CM
Asymptotic Covariance Matrix from File EFFUSA.ACC
Sample Size: 1554
Latent Variables: Efficacy Respons
Relationships:
  NOSAY = CONST 1*Efficacy
```

¹¹The same wording was used in these countries, except the USA had *Congress in Washington* whereas Britain had *Parliament* in the TOUCH item, see Section 1.

```
COMPLEX = CONST Efficacy
NOCARE = CONST Efficacy Respons
TOUCH = CONST Respons
INTEREST = CONST 1*Respons
```

```
Group BTN
Observed Variables: NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
Means from File EFFBTN.ME
Covariance Matrix from File EFFBTN.CM
Asymptotic Covariance Matrix from File EFFBTN.ACC
Sample Size: 1266
Latent Variables: Efficacy Respons
Relationships:
  Efficacy Respons = CONST
Set the error variance of NOSAY free
Set the error variances of COMPLEX - INTEREST free
Set the variances of Efficacy - Respons free
Set the covariance between Efficacy and Respons free
Path Diagram
End of Problem
```

The general rule in SIMPLIS is that everything is the same as in the previous group unless otherwise stated. The model is specified as relationships. Note that VOTING is not included in the relationships. As a consequence, VOTING is automatically excluded in the model although it is included in the data, *i.e.*, it is included in the mean vector, the covariance matrix, and the asymptotic covariance matrix. Since the relationships are the same in group 2 they are not repeated in that group. The mean vector of ξ is zero in group 1 by default. The line

```
Efficacy Respons = CONST
```

in group 2 means that we want to estimate that mean vector in that group. The two lines

```
Set the error variance of NOSAY free
Set the error variances of COMPLEX - INTEREST free
```

specify that we want to estimate the measurement error variances as free parameters in group 2. Note that it is *not* possible to specify this with the single line

```
Set the error variances of NOSAY - INTEREST free
```

because this will include VOTING. The two lines

```
Set the variances of Efficacy - Respons free
Set the covariance between Efficacy and Respons free
```

specify that we do not want to constrain the covariance matrix of ξ to be the same in the two groups.

The output file gives the following fit statistics of overall fit (here I have selected the only fit statistics you need to consider).

```
Degrees of Freedom = 13
Minimum Fit Function Chi-Square = 19.47 (P = 0.11)
Root Mean Square Error of Approximation (RMSEA) = 0.019
90 Percent Confidence Interval for RMSEA = (0.0 ; 0.035)
```

As judged by the P -value for exact fit, the model fits very well. Some people might say that I should test whether τ_x and Λ_x are invariant over groups and not just assume that. However, since the model fits well and makes good sense, I do not need to do that.

The results concerning the distribution of the latent variables is summarized in Table 10. This gives the estimated means, variances, and covariances of the latent variables, with their standard errors and t -values (The standard error is given in parenthesis and the t -value follows after the standard errors).

Table 10: USA vs Britain: Estimated Means and Covariance Matrices

<i>USA</i>			
	Efficacy	Respons	<i>Means</i>
Efficacy	0.19(0.02)9.85		0.00
Respons	0.20(0.01)17.09	0.31(0.02)20.08	0.00

<i>Britain</i>			
	Efficacy	Respons	<i>Means</i>
Efficacy	0.15(0.01)10.09		-0.13(0.03)-4.00
Respons	0.18(0.01)14.57	0.38(0.02)21.42	-0.09(0.04)-2.14

The t -values for the means suggest that there is a significant mean difference between the two countries for both **Efficacy** and **Respons**. The means are larger in the USA than in Britain. One might say that “the USA is ahead of Britain on Efficacy” or that “people in the USA are more efficacious than in Britain”.

The t -values for the variances and covariances are not useful. The only hypothesis of any interest is whether the covariance matrix of the latent variables are the same in both countries. This can be tested formally by omitting the two lines

```
Set the variances of Efficacy - Respons free
Set the covariance between Efficacy and Respons free
```

in the second group. However, for most purposes, it does not matter if this hypothesis holds or not. It is more important to consider the interpretation of differences in variances and covariances. If the variance of **Efficacy**, say, is noticeably smaller in one country than in another country, it can be interpreted as “People in the first country are more homogeneous with regard to their feeling of efficacy” in the first country compared to the other country. If the differences between the variances are small but the difference between the covariances is large, such that the correlation between the latent variables are different in the two countries, this might suggest that “there is more confusion about the distinction between **Efficacy** and **Respons**” in the country with the smaller covariance than in the country with the larger covariance. In my example, the differences in variances and covariances are rather small, even if some difference may be statistically significant.

Two other countries with the same language are Germany and Austria. I leave it as an exercise for anyone interested to carry out the same analysis for these two countries and to verify that the same model also fits well in these two countries. I now turn to the analysis of the USA and Germany.

USA vs Germany

Fitting the model of factorial invariance (in the sense of equal intercepts and equal factor loadings) to the USA and Germany gives a chi-square of 140.17 with 13 degrees of freedom and a RMSEA of 0.074 indicating that the model does not fit well.¹² There are large modification indices for the

¹²Since I do not believe that the model holds exactly in the population, I use RMSEA and the guidelines of Browne & Cudeck (1993) to judge whether the model fits approximately in the population.

two loadings of `NOCARE` indicating that these are different in the two countries (they are larger in Germany). Allowing these two loadings to be free in each country gives a chi-square of 72.47 with 11 degrees of freedom and a RMSEA of 0.056. Note that it is still possible to compare the variances and covariances of the two latent variables across countries since they are still measured in the same units (because they are measured in the units of the underlying variables of `NOSAY` and `INTEREST` which themselves are in the same units by the construction in the `PRELIS` step). The fit is still not adequate. There are large modification indices for the intercepts of `NOSAY` and `COMPLEX`. This indicates that the large differences between the mean vectors of the underlying variables between the two countries cannot be entirely accounted for by the mean differences in the latent variables.

One must be careful in relaxing the assumption of equal intercepts, however. It is this assumption that makes it possible to estimate the means of the latent variables on a scale with the same origin. In this case, if we allow the intercepts of `NOSAY`, `COMPLEX`, and `NOCARE` to be different, one will not be able to estimate the mean difference in `Efficacy`. One can, however, allow one of these intercepts to be different. Relaxing the intercept of `NOCARE` will not improve the fit much. In choosing between `NOSAY` and `COMPLEX`, it is best to choose `COMPLEX`, because the mean difference of `Efficacy` is well defined by the mean difference in `NOSAY`. Thus, in the next model, I will allow the intercept of `COMPLEX` to be different in the two countries but assume that all the other intercepts are the same. This gives a chi-square of 20.21 with 10 degrees of freedom and a RMSEA of 0.024. Thus, this model fits well. Note that chi-square decreased from 72.47 to 20.21 only by adding one single parameter.

The `SIMPLIS` command file for the last model is not listed here but is given in file `ORD45.SPL`. The estimated means, variances, and covariances of the latent variables, with their standard errors and *t*-values, are given in Table 11.

Table 11: USA vs Germany: Estimated Means and Covariance Matrices

<i>USA</i>			
	<i>Efficacy</i>	<i>Respons</i>	<i>Means</i>
<i>Efficacy</i>	0.19(0.02)10.01		0.00
<i>Respons</i>	0.20(0.01)17.35	0.32(0.01)22.72	0.00

<i>Germany</i>			
	<i>Efficacy</i>	<i>Respons</i>	<i>Means</i>
<i>Efficacy</i>	0.62(0.02)25.47		-0.52(0.05)-11.30
<i>Respons</i>	0.44(0.01)37.05	0.61(0.02)39.16	-0.12(0.03)-3.92

All Countries

I will now analyze the data from all countries simultaneously. There are several ways to do this and several models that can be considered. In line with the previous sections my objective is to achieve maximum factorial invariance over countries in the sense that intercepts and factor loadings should be invariant over countries. But the model must also fit the data reasonably well and all estimated parameters should be statistically significant and meaningful. In particular, I want to estimate the mean vector (relative to the USA) and the covariance matrix of the latent variables in each country.

I know from the previous analysis of the USA and Germany that I will not be able to fit the model of complete factorial invariance over all eight countries. If I do I get a chi-square of 473.32 with 73 degrees of freedom and a RMSEA of 0.059. This is not a satisfactory fit.

My guess is that I can fit the model of complete factorial invariance over the USA, The Netherlands, Britain, Italy, Switzerland, and Finland but not with Germany and Austria. I can also fit complete factorial invariance over Germany and Austria, but the model in Germany and Austria differs slightly from that of the other countries. The only differences are that (i) the two loadings of `NOCARE` are different and (ii) the intercept of `COMPLEX` is different. To estimate this model with `SIMPLIS` it is

convenient to order the countries in the command file such that Germany and Austria are the last two countries. The SIMPLIS command file is not listed here but is given in file ORD46.SPL.

Note the following

- The measurement relations are not repeated in The Netherlands, Britain, Italy, Switzerland, and Finland. Hence the intercepts and factor loadings in these countries are the same as in the USA.
- The lines

```
Set the error variance of NOSAY free
Set the error variances of COMPLEX - INTEREST free
Set the variances of Efficacy - Respons free
Set the covariance between Efficacy and Respons free
```

are repeated in each country to allow the error variances and the covariance matrix of the latent variables to be free in each country.

- The line

```
Efficacy Respons = CONST
```

is included in each country except the USA to specify that the means of the latent variables are to be estimated in these countries. These means are zero in the USA.

- The two lines

```
COMPLEX = CONST
NOCARE = Efficacy Respons
```

are added in Germany to specify that the two loadings of NOCARE are different and that the intercept of COMPLEX is different in Germany. Note that these lines are not repeated in Austria, which makes these quantities equal in Germany and Austria.

The output gives the following set of fit statistics.

```
Degrees of Freedom = 70
Minimum Fit Function Chi-Square = 236.51 (P = 0.0)
Root Mean Square Error of Approximation (RMSEA) = 0.043
90 Percent Confidence Interval for RMSEA = (0.037 ; 0.049)
```

The hypothesis that the model holds exactly in the population is rejected. But, following the guidelines of Browne & Cudeck (1993), since RMSEA is below 0.05 and its upper confidence limit is below 0.08, I judge that the fit represents a reasonable degree of approximation in the population.

If one includes the line

LISREL Output

the set of factor loadings common to the USA, The Netherlands, Britain, Italy, Switzerland, and Finland is given in the output as

	Efficacy	Respons
	-----	-----
NOSAY	1.00	- -

COMPLEX	1.04 (0.02) 55.99	- -
NOCARE	0.97 (0.05) 18.90	0.26 (0.03) 7.63
TOUCH	- -	0.88 (0.01) 94.51
INTEREST	- -	1.00

The loadings for Germany and Austria differ from these only for NOCARE. For Germany and Austria these loadings are

NOCARE	0.43 (0.02) 17.45	0.61 (0.02) 27.40
--------	-------------------------	-------------------------

The loading on *Efficacy* is noticeably smaller and the loading on *Respons* is noticeably larger. Thus, the item NOCARE functions differently in these two groups of countries. In the USA, The Netherlands, Britain, Italy, Switzerland, and Finland, NOCARE is mainly a measure of *Efficacy*, whereas in Germany and Austria, NOCARE is more of a measure of *Respons* than of *Efficacy*.

In the USA, The Netherlands, Britain, Italy, Switzerland, and Finland the intercepts in the measurement equations are estimated as

NOSAY	COMPLEX	NOCARE	TOUCH	INTEREST
-----	-----	-----	-----	-----
0.91	0.66	0.89	0.55	0.69
(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
35.72	26.37	42.25	19.87	23.18

In Germany and Austria these intercepts are estimated as

NOSAY	COMPLEX	NOCARE	TOUCH	INTEREST
-----	-----	-----	-----	-----
0.91	1.21	0.89	0.55	0.69
(0.03)	(0.04)	(0.02)	(0.03)	(0.03)
35.72	28.07	42.25	19.87	23.18

The only difference is for COMPLEX but this difference is highly significant.

The estimated means of *Efficacy* and *Respons* are given in Table 12. All these means are statistically significant except for the mean of *Respons* in Britain. Note that all estimated means are negative except for *Respons* in The Netherlands. Hence, the USA is “ahead” of all countries on *Efficacy* and “ahead” of all countries except The Netherlands on *Respons*. Note also that Germany and Austria are “way below” the USA on *Efficacy* and that Austria is “way below” the USA on *Respons* as well. If we rank the countries in order of decreasing *Efficacy*, the order is the USA, The Netherlands, Britain, Switzerland, Finland, Italy, Germany and Austria.

A command file in LISREL syntax for doing exactly the same thing as in ORD46.SPL is given in ORD46A.LS8.

Table 12: Estimated Means of Efficacy and Respons

	USA	GER	NET	AUS	BTN	ITY	SWI	FIN
Efficacy	0.000	-0.55	-0.08	-0.82	-0.15	-0.30	-0.23	-0.25
Respons	0.000	-0.10	0.08	-0.41	-0.07	-0.39	-0.09	-0.18

4.4 Conclusion

In this section I have examined the factorial invariance of the five efficacy items **NOSAY**, **COMPLEX**, **NOCARE**, **TOUCH**, and **INTEREST** in the eight countries USA, The Netherlands, Britain, Italy, Switzerland, Finland, Germany, and Austria using data from the Political Action Survey. The conclusions are

- The five efficacy items measure the same two latent variables **Efficacy** and **Respons** in all countries.
- **NOSAY**, **TOUCH**, and **INTEREST** are functionally equivalent in all eight countries, in the sense that the intercepts and slopes (factor loadings) are the same in all eight countries. **COMPLEX** and **NOCARE** are functionally equivalent in the six countries USA, The Netherlands, Britain, Italy, Switzerland, and Finland. They are also functionally equivalent in Germany and Austria, but these countries differ from the other six countries in two respects
 - The loadings of **NOCARE** on **Efficacy** and **Respons** are different. In Germany and Austria the loading on **Efficacy** is much smaller and the loading on **Respons** is much larger than in the other six countries. Thus, the item **NOCARE** functions differently in these two groups of countries. In the USA, The Netherlands, Britain, Italy, Switzerland, and Finland, **NOCARE** is mainly a measure of **Efficacy**, whereas in Germany and Austria, **NOCARE** is more of a measure of **Respons** than of **Efficacy**.
 - **COMPLEX** measures **Efficacy** at a higher level in the sense that the intercept is much larger in Germany and Austria than in the other six countries.
- The means of **Efficacy** and **Respons** are considerably different among the eight countries, with the USA being the most efficacious and Germany and Austria the least efficacious.

5 Covariates

In the previous four sections I assumed that all observed variables were ordinal. Thus, in Section 2, I described the analysis of ordinal variables in cross-sectional studies, in Section 3, I described the analysis of ordinal variables in longitudinal studies, and in Section 4, I described the analysis of ordinal variables observed in several groups. In this section I consider the case when one or more ordinal variables are observed jointly with a set of possibly explanatory variables, so called *covariates*. These covariates can be dummy-coded categorical variables or measured variables on an interval scale. They are assumed not to contain measurement error. With PRELIS one can estimate the effect of the covariates on the probability of response in various categories of the ordinal variables using either the probit or the logit model. PRELIS can also estimate the joint covariance matrix of the covariates and the variables underlying the ordinal variables. This can be used for further modeling in LISREL.

Continuing my analysis of the Efficacy variables from the Political Action Survey, I illustrate the analysis of the six Efficacy variables using four covariates: Gender, Age, Education, and a Left-Right Scale. For information about the Political Action Survey and the Efficacy variables, see Section 1. As in Sections 2 and 3, I will only use the data from the USA sample.

Since probit and logit regression have not been well documented in the LISREL literature, I give a rather technical description in Sections 5.1–5.5. Readers who are merely interested in how to do it with LISREL can skip this and proceed to Section 5.6.

5.1 Univariate Probit Regression

Let y be a single ordinal variable¹³ with m categories and let $\mathbf{x}(q \times 1)$ be a vector of covariates. Corresponding to y there is an underlying continuous variable y^* . The connection between the ordinal variable y and the underlying variable y^* is

$$y = a \iff \tau_{a-1} < y^* \leq \tau_a, \quad a = 1, 2, \dots, m, \quad (17)$$

where

$$\tau_0 = -\infty, \tau_1 < \tau_2 < \dots < \tau_{m-1}, \tau_m = +\infty,$$

are threshold parameters. For variable y with m categories, there are $m - 1$ threshold parameters

$$\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_{m-1}).$$

The specification (17) is the same as in Section 2. In fact, the development in Section 2 is a special case of the more general case described here, namely when $q = 0$.

Consider the regression of y^* on \mathbf{x} :

$$y^* = \alpha + \boldsymbol{\gamma}'\mathbf{x} + z, \quad (18)$$

where α is an intercept term, $\boldsymbol{\gamma}$ is a vector of regression coefficients, and z is an error term. The underlying variable y^* is not observed; only the ordinal variables y and \mathbf{x} are.

The probit model assumes that z is normally distributed with mean 0 and variance ψ^2 , *i.e.*, y^* is normal conditional on \mathbf{x} :

$$y^* \sim N(\alpha + \boldsymbol{\gamma}'\mathbf{x}, \psi^2).$$

It follows that the probability $\pi_c(\mathbf{x})$ of a response in category c or lower, conditional on \mathbf{x} , where $c = 1, 2, \dots, m - 1$, is

$$\pi_c(\mathbf{x}) = \Phi\left(\frac{\tau_c - \alpha - \boldsymbol{\gamma}'\mathbf{x}}{\psi}\right), \quad (19)$$

where Φ is the standard normal distribution function.

Equation (19) can be viewed as a special case of a generalized linear model, see *e.g.*, McCullagh & Nelder (1983). In this tradition there are no concepts of underlying variables and thresholds. Instead, (19) is written

$$\pi_c(\mathbf{x}) = \Phi(\alpha_c^* - \boldsymbol{\gamma}^*\mathbf{x}), \quad (20)$$

where

$$\alpha_c^* = \psi^{-1}(\tau_c - \alpha),$$

is interpreted as an intercept term and

$$\boldsymbol{\gamma}^* = \psi^{-1}\boldsymbol{\gamma},$$

is a vector of regression coefficients.

To explain the term *probit regression*, take the inverse of (20):

$$\Phi^{-1}[\pi_c(\mathbf{x})] = \alpha_c^* - \boldsymbol{\gamma}^*\mathbf{x}, \quad (21)$$

where Φ^{-1} is the inverse function of Φ . The quantity $\Phi^{-1}(\pi)$ is called the probit of π . If π goes from 0 to 1, $\Phi^{-1}(\pi)$ goes from $-\infty$ to $+\infty$. Equation (21) shows that the probit of $\pi_c(\mathbf{x})$ is linear in \mathbf{x} , hence the term probit regression. Note that the sign of $\boldsymbol{\gamma}$ is negative in (20) but positive in (18). In (20), for example, if γ_1 is positive, the probability of a response in category c or lower decreases as x_1 increases. Which says the same thing as the probability of a response in a category higher than c increases with x_1 . In (18), however, y^* increases if x_1 increases which increases the probability of a higher response. Thus, the two models are equivalent.

¹³The term univariate is used here in the sense of one variable at a time. One can very well have several ordinal variables but they are analyzed one at a time.

One can regard (21) as $m - 1$ parallel regression lines. Note that the intercepts vary with c but the regression coefficients are the same. For ordinal variables the intercepts must satisfy the order condition

$$\alpha_1^* < \alpha_2^* < \dots < \alpha_{m-1}^* .$$

To illustrate the function $\pi_c(\mathbf{x})$ in (19), consider the case of a single covariate x . Let $\alpha = 0$, $\psi = 1$ and denote

$$\pi_{\tau\gamma}(x) = \Phi(\tau - \gamma x) .$$

Fig. 7 shows four curves $\pi_{\tau\gamma}(x)$ for $-10 < x < 10$ using the parameter values

Curve 1 $\tau = -0.5$ and $\gamma = 1.0$

Curve 2 $\tau = 1.5$ and $\gamma = 1.0$

Curve 3 $\tau = -0.5$ and $\gamma = 0.4$

Curve 4 $\tau = 1.5$ and $\gamma = 0.4$.

It is seen that the probability of a response in category c or lower, *i.e.*, the probability that $y^* \leq \tau$, decreases with x . The larger γ is, the faster is the rate of decrease. As τ increases or decreases, the curves are just shifted vertically.

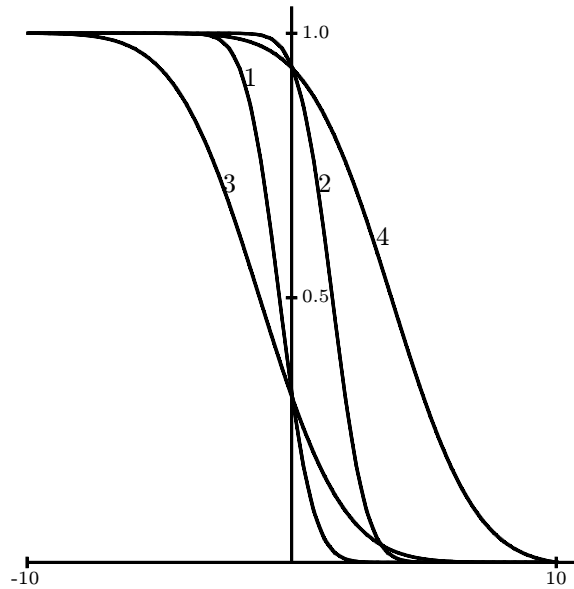


Figure 7: Four Cumulative Response Functions

I now return to the development of equation (19). There are two fundamental indeterminacies in (19):

- One can add a constant to all τ 's and to α .
- One can multiply all τ 's, α , and γ by a constant and multiply ψ by the same constant.

Neither of these changes has any effect on the right hand side of (19). This is a reflection of the fact that since only ordinal information is available about y , y^* is only determined up to a linear transformation¹⁴.

PRELIS has two ways of resolving these indeterminacies:

¹⁴Actually, y^* is only determined up to a monotonic transformation, but under normality the transformation must be linear.

Standard Parameterization: $\alpha = 0$ and $\psi = 1$

Alternative Parameterization: $\tau_1 = 0$ and $\tau_2 = 1$

These parameterizations fix the origin and unit of measurement of y^* in two different ways. The Standard Parameterization is the same as used in Generalized Linear Models. The Alternative Parameterization requires that $m \geq 3$. If $m = 2$ under this parameterization, PRELIS will set $\tau_1 = 0$ and $\psi = 1$.

For $m \geq 3$, the parameters of the two parameterizations are given in the following table.

Parameterization	Intercept	Error Var.	Thresholds					Regr. Coeff.			
Standard	0	1	τ_1	τ_2	τ_3	\dots	τ_{m-1}	γ_1	γ_2	\dots	γ_q
Alternative	α	ψ^2	0	1	τ_3^*	\dots	τ_{m-1}^*	γ_1^*	γ_2^*	\dots	γ_q^*

where

$$\begin{aligned} \alpha &= -\tau_1/(\tau_2 - \tau_1), \quad \psi = 1/(\tau_2 - \tau_1), \\ \tau_i^* &= (\tau_i - \tau_1)/(\tau_2 - \tau_1), \quad i = 3, 4, \dots, m-1, \\ \gamma_i^* &= \gamma_i/(\tau_2 - \tau_1), \quad i = 1, 2, \dots, q. \end{aligned}$$

It should be emphasized that the two parameterizations are equivalent in the sense that there is a one-to-one correspondence between the two sets of parameters.

For estimation, the probability of a response in category a is needed, where $a = 1, 2, \dots, m$. This is

$$Pr\{y = a \mid \mathbf{x}\} = \pi_a(\mathbf{x}) - \pi_{a-1}(\mathbf{x}) = \Phi\left(\frac{\tau_a - \alpha - \boldsymbol{\gamma}'\mathbf{x}}{\psi}\right) - \Phi\left(\frac{\tau_{a-1} - \alpha - \boldsymbol{\gamma}'\mathbf{x}}{\psi}\right). \quad (22)$$

It is convenient to refer to (22) as the category probability function and to (19) as the cumulative probability function.

For a single x , the category probability functions in (22) are shown in Fig. 8 for $\alpha = 0$, $\psi = 1$, $\tau_1 = -0.5$, $\tau_2 = 1.5$ and $\gamma = 1$ (Curve 1) and $\gamma = 0.4$ (Curve 2). As x increases the category probability increases up to a maximum

$$2\Phi\left[\frac{1}{2}(\tau_2 - \tau_1)\right] - 1 \quad \text{at} \quad x = \frac{\tau_1 + \tau_2}{2\gamma},$$

and then decreases. The rate of increase and decrease is larger for larger γ than for smaller γ . Note that the maximum is independent of γ .

Suppose we have a random sample of N independent observations of y and \mathbf{x} :

$$(y_i, \mathbf{x}_i), \quad i = 1, 2, \dots, N.$$

Let $k_{ia} = 1$, if $y_i = a$, and $k_{ia} = 0$, otherwise. Then

$$E(k_{ia} \mid \mathbf{x}_i) = \Phi\left(\frac{\tau_a - \alpha - \boldsymbol{\gamma}'\mathbf{x}_i}{\psi}\right) - \Phi\left(\frac{\tau_{a-1} - \alpha - \boldsymbol{\gamma}'\mathbf{x}_i}{\psi}\right) = \pi_{ia}^*(\mathbf{x}_i), \quad (23)$$

say. The likelihood of the sample is

$$L = \prod_{i=1}^N \left\{ \prod_{a=1}^m [\pi_{ia}^*(\mathbf{x}_i)]^{k_{ia}} \right\} p(\mathbf{x}_i), \quad (24)$$

where $p(\mathbf{x})$ is the density function of \mathbf{x} . The latter is unspecified and assumed to have no parameters of interest. The parameter vector is

$$\boldsymbol{\theta} = (\boldsymbol{\tau}, \alpha, \boldsymbol{\gamma}, \psi).$$

This can be estimated by maximizing the likelihood L of either of the two parameterizations.

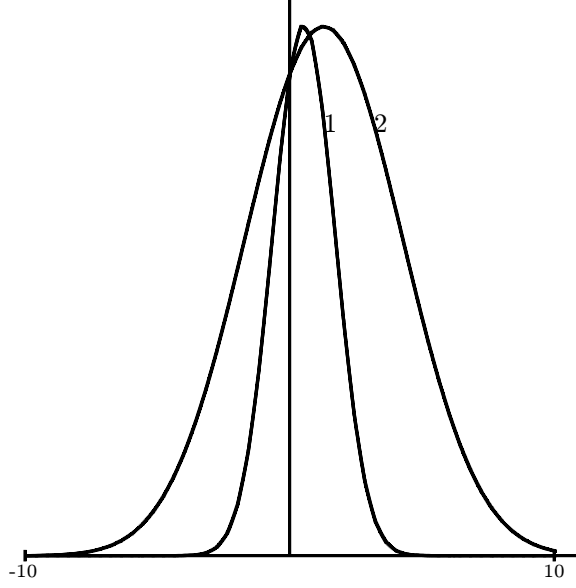


Figure 8: Two Category Response Functions (probit)

5.2 Univariate Logit Regression

One can obtain *logit regression*, sometimes called *logistic regression*, in the same way simply by replacing the normal distribution function $\Phi(x)$ by the logistic distribution function

$$\Psi(u) = \frac{e^u}{1 + e^u}.$$

The inverse function of Ψ is

$$\Psi^{-1}(\pi) = \ln \frac{\pi}{1 - \pi}.$$

The quantity $\ln \frac{\pi}{1 - \pi}$ is called the *logit* of π . If π goes from 0 to 1, *logit*(π) goes from $-\infty$ to $+\infty$.

The logit model is

$$\ln \frac{\pi_c(\mathbf{x})}{1 - \pi_c(\mathbf{x})} = \alpha_c^* - \gamma^{*'} \mathbf{x}. \quad (25)$$

This is also a special case of a Generalized Linear Model, see McCullagh & Nelder (1983)¹⁵.

PRELIS estimates the logit model in the form

$$\pi_c(\mathbf{x}) = \Psi\left(\frac{\tau_c - \alpha - \gamma' \mathbf{x}}{\psi}\right), \quad (26)$$

using either the Standard Parameterization or the Alternative Parameterization as defined in the previous Section.

The probability of a response in category a is

$$Pr\{y = a \mid \mathbf{x}\} = \pi_a(\mathbf{x}) - \pi_{a-1}(\mathbf{x}) = \Psi\left(\frac{\tau_a - \alpha - \gamma' \mathbf{x}}{\psi}\right) - \Psi\left(\frac{\tau_{a-1} - \alpha - \gamma' \mathbf{x}}{\psi}\right), \quad (27)$$

This probability as a function of a single x is shown in Fig. 9 for the same parameters as in Fig. 8.

It is seen that the logit model gives less probability to the category corresponding to $\tau_1 < y^* \leq \tau_2$ and more probability to the other categories than the probit model. In general, the logit model gives

¹⁵The logit model seems to be more often used in practice than the probit model. This is probably because Ψ^{-1} has an explicit form. However, with computers it is almost as easy to compute Φ^{-1} as it is to compute Ψ^{-1} .

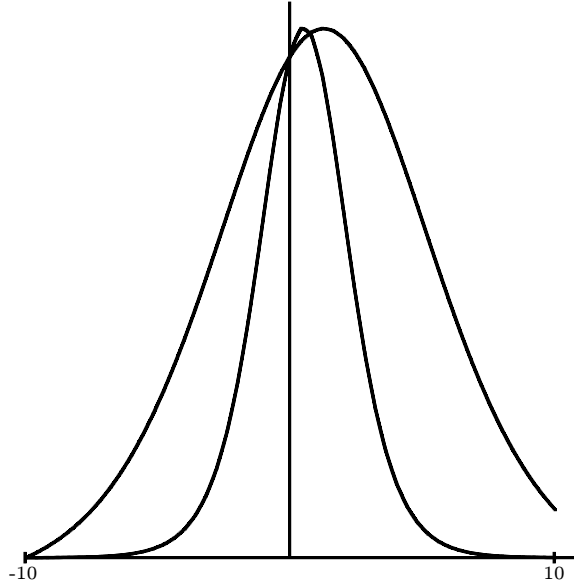


Figure 9: Two Category Response Functions (logit)

less probability to the middle categories and more probability to the outer categories than the probit model.

The logistic and the normal distribution are similar, but the variance of the logistic distribution $\Psi(u)$ is not 1 as is the case of the normal distribution $\Phi(u)$. Lord & Novick (1968, p. 299) noted that

$$|\Phi(u) - \Psi(1.7u)| < 0.01, \text{ for all } x. \quad (28)$$

Because of this closeness, results obtained under the Standard Parameterization with the logit model are likely to be close to those obtained with the probit model except for a scale factor. As will be demonstrated, the Alternative Parameterization eliminates this scale factor and makes the regression equations directly comparable and similar.

5.3 Testing the Model

The univariate probit and logit models make strong assumptions about the cumulative response functions in the form of (19) and (26), respectively. Can these assumptions be tested? The typical way of doing this is to compute a deviance, *i.e.*, the difference between $-2 \ln \hat{L}$ for the model and the same quantity for another more general model, where \hat{L} is the maximum value of (24). This kind of deviance has not been implemented in PRELIS because it is not obvious what the more general model should be. However, PRELIS prints the value of $-2 \ln \hat{L}$ so one can compare this for different models. We have also implemented a test of the hypothesis that $\gamma = \mathbf{0}$, *i.e.*, that all regression coefficients are zero. This can also be regarded as a measure of how much better the model fits than the model with no covariates. This will be illustrated in Section 5.10.

5.4 Bivariate Probit Regression

The normal distribution generalizes naturally to the bivariate and multivariate case. The logistic distribution function, however, does not have any convenient generalization to the bivariate and multivariate case. For this reason I consider only the case of underlying bivariate and multivariate normality in what follows.

Consider two ordinal variables y_g and y_h with underlying continuous variables y_g^* and y_h^* , respectively. The equations to be estimated are

$$y_g^* = \alpha_g + \boldsymbol{\gamma}'_g \mathbf{x} + z_g, \quad (29)$$

$$y_h^* = \alpha_h + \boldsymbol{\gamma}'_h \mathbf{x} + z_h, \quad (30)$$

where α_g and α_h are intercept terms, $\boldsymbol{\gamma}_g$ and $\boldsymbol{\gamma}_h$ are vectors of regression coefficients, and z_g and z_h are error terms. It is assumed that z_g and z_h have a bivariate normal distribution with means zero and covariance matrix

$$\begin{pmatrix} \psi_g^2 & \\ \psi_{gh} & \psi_h^2 \end{pmatrix}.$$

In the Standard Parameterization this is a correlation matrix with correlation ρ_{gh} . Variable y_g has thresholds

$$\boldsymbol{\tau}_g = (\tau_{g,1}, \tau_{g,2}, \dots, \tau_{g,m_g-1}),$$

and variable y_h has thresholds

$$\boldsymbol{\tau}_h = (\tau_{h,1}, \tau_{h,2}, \dots, \tau_{h,m_h-1}).$$

The probability that an individual i with covariates \mathbf{x}_i responds in category a on y_g and in category b on y_h is

$$\pi_{igh,ab} = Pr\{y_{ig} = a, y_{ih} = b \mid \mathbf{x}_i\} = \int_{\tau_{ig,a-1}^*}^{\tau_{ig,a}^*} \int_{\tau_{ih,b-1}^*}^{\tau_{ih,b}^*} \phi^{(2)}(u, v, \rho_{gh}) dudv, \quad (31)$$

where

$$\tau_{ig,a}^* = \frac{\tau_{ga} - \alpha_g - \boldsymbol{\gamma}'_g \mathbf{x}_i}{\psi_g}. \quad (32)$$

and $\phi^{(2)}(u, v, \rho)$ is the density function of the standardized bivariate normal distribution with correlation ρ . The parameter vector is

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_g, \boldsymbol{\theta}_h, \rho_{gh}),$$

where

$$\boldsymbol{\theta}_g = (\boldsymbol{\tau}_g, \alpha_g, \boldsymbol{\gamma}_g, \psi_g),$$

$$\boldsymbol{\theta}_h = (\boldsymbol{\tau}_h, \alpha_h, \boldsymbol{\gamma}_h, \psi_h).$$

The likelihood function is

$$L = \prod_{i=1}^N \left(\prod_{a=1}^{m_g} \prod_{b=1}^{m_h} \pi_{igh,ab}^{k_{igh,ab}} \right) p(\mathbf{x}_i), \quad (33)$$

where $k_{igh,ab} = 1$ if case i responds in category a on y_g and in category b on y_h , and $k_{igh,ab} = 0$, otherwise.

PRELIS estimates $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_h$ from the univariate marginal distribution of y_g and y_h , respectively, as described in Section 5.1. Given these estimates, PRELIS estimates ρ_{gh} by maximizing the bivariate likelihood L in (33). Under the Alternative Parameterization, the conditional covariance between y_g^* and y_h^* is estimated as

$$\hat{\psi}_{gh} = \hat{\psi}_g \hat{\psi}_h \hat{\rho}_{gh}.$$

5.5 Multivariate Probit Regression

Let $\mathbf{y}(p \times 1)$ be a vector of ordinal variables with underlying variables \mathbf{y}^* . It is assumed that

$$\mathbf{y}^* | \mathbf{x} \sim N(\boldsymbol{\alpha} + \boldsymbol{\Gamma}\mathbf{x}, \boldsymbol{\Psi}) .$$

The rows of $\boldsymbol{\alpha}$ and $\boldsymbol{\Gamma}$ and the diagonal elements of $\boldsymbol{\Psi}$ are estimated from the univariate margins as described in Section 5.1, and the off-diagonal elements of $\boldsymbol{\Psi}$ are estimated from the bivariate margins as described in Section 5.4.

Denoting these estimates as $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\Gamma}}$, and $\hat{\boldsymbol{\Psi}}$, we have the following:

- The estimated conditional covariance matrix of \mathbf{y}^* for given \mathbf{x} is $\hat{\boldsymbol{\Psi}}$. In the Standard Parameterization this is a correlation matrix.
- The estimated unconditional covariance matrix of \mathbf{y}^* is

$$\hat{\boldsymbol{\Gamma}}\mathbf{S}_{xx}\hat{\boldsymbol{\Gamma}}' + \hat{\boldsymbol{\Psi}} ,$$

where \mathbf{S}_{xx} is the sample covariance matrix of \mathbf{x} .

- The estimated joint unconditional covariance matrix of \mathbf{y}^* and \mathbf{x} is

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\boldsymbol{\Gamma}}\mathbf{S}_{xx}\hat{\boldsymbol{\Gamma}}' + \hat{\boldsymbol{\Psi}} & \\ \mathbf{S}_{xx}\hat{\boldsymbol{\Gamma}}' & \mathbf{S}_{xx} \end{pmatrix} . \quad (34)$$

The relationship between the Standard and Alternative Parameterizations can be expressed in matrix form as follows. Let \mathbf{D} be the diagonal matrix of order $p \times p$

$$\mathbf{D} = \text{diag}\left(\frac{1}{\tau_{1,2} - \tau_{1,1}}, \frac{1}{\tau_{2,2} - \tau_{2,1}}, \dots, \frac{1}{\tau_{p,2} - \tau_{p,1}}\right) , \quad (35)$$

and let \mathbf{y}_S^* and \mathbf{y}_A^* denote the vector of underlying variables in the Standard and Alternative Parameterizations, respectively. Then

$$\mathbf{y}_A^* = \boldsymbol{\alpha} + \mathbf{D}\mathbf{y}_S^* , \quad (36)$$

where

$$\boldsymbol{\alpha} = \left(\frac{-\tau_{1,1}}{\tau_{1,2} - \tau_{1,1}}, \frac{-\tau_{2,1}}{\tau_{2,2} - \tau_{2,1}}, \dots, \frac{-\tau_{p,1}}{\tau_{p,2} - \tau_{p,1}}\right)' .$$

Furthermore, let $\hat{\boldsymbol{\Gamma}}_S$ and $\hat{\boldsymbol{\Psi}}_S$ be the matrices $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{\Psi}}$ estimated under the Standard Parameterization, and let $\hat{\boldsymbol{\Gamma}}_A$ and $\hat{\boldsymbol{\Psi}}_A$ be the corresponding matrices estimated under the Alternative Parameterization. Then

$$\hat{\boldsymbol{\Gamma}}_A = \mathbf{D}\hat{\boldsymbol{\Gamma}}_S , \quad (37)$$

$$\hat{\boldsymbol{\Psi}}_A = \mathbf{D}\hat{\boldsymbol{\Psi}}_S\mathbf{D} . \quad (38)$$

Using the same notation for the matrix $\hat{\boldsymbol{\Sigma}}$ in (34), we have

$$\hat{\boldsymbol{\Sigma}}_A = \mathbf{D}_1\hat{\boldsymbol{\Sigma}}_S\mathbf{D}_1 , \quad (39)$$

where \mathbf{D}_1 is the diagonal matrix of order $p + q \times p + q$

$$\mathbf{D}_1 = \text{diag}\left(\frac{1}{\tau_{1,2} - \tau_{1,1}}, \frac{1}{\tau_{2,2} - \tau_{2,1}}, \dots, \frac{1}{\tau_{p,2} - \tau_{p,1}}, 1, 1, \dots, 1\right) . \quad (40)$$

PRELIS can also estimate the asymptotic covariance matrix of $\hat{\boldsymbol{\Sigma}}$.

There is no latent variable model (LISREL model) imposed on the $\hat{\boldsymbol{\Sigma}}$ in (34). It is an unconstrained covariance matrix just as a sample covariance matrix \mathbf{S} for continuous variables. It can therefore be used for modeling in LISREL just as if \mathbf{y}^* and \mathbf{x} were directly observed. This is illustrated in Section 5.12.

5.6 PRELIS Implementation

The features described here have been implemented in PRELIS 2.52 which is available with LISREL 8.52 in June 2002. In PRELIS 2.51 or earlier versions, probit regression was available via the FI command, see Jöreskog & Sörbom (1999a, pp. 180–183) and logit regression was not available at all.

I illustrate the case of 3 ordinal variables and 4 covariates. Let Y1 Y2 Y3 be the names of the ordinal variables and let X1 X2 X3 X4 be the names of the covariates.

Probit regression of Y1 is obtained by the PRELIS command

```
PR Y1 on X1 X2 X3 X4
```

Similarly, logit regression of Y1 is obtained by the PRELIS command

```
LR Y1 on X1 X2 X3 X4
```

One can select any subset of y -variables and any subset of x -variables to be included in the equation. Thus, one can obtain the univariate probit or logit regression for all the ordinal variables simultaneously. For example,

```
PR Y1 Y2 Y3 on X1 X2 X3 X4
```

will give three univariate probit regressions. Note the word on (or ON) separating the ordinal variables from the covariates.

One can have several PR and/or LR commands in the same input file. All x -variables used as covariates must be declared continuous before the first PR or LR command, or else they must have at least 16 different values.

The Standard Parameterization is used by default. To obtain the Alternative Parameterization put AP on the Output line. The PR or LR command produces only univariate probit or logit regressions. Thus an MA value specified on the Output line has no meaning. To obtain the matrix $\hat{\Sigma}$ in (34), use an FI command and put MA=CM on the Output line. No other value of MA is meaningful since $\hat{\Sigma}$ is a covariance matrix even in the Standard Parameterization. There are two reasons why the covariance matrix $\hat{\Sigma}$ in (34) is not computable with PR or LR commands:

- Since one can have several PR or LR commands in the same PRELIS command file, there is no way PRELIS will know which covariance matrix to compute.
- Since the logistic distribution does not generalize to the multivariate case, the covariance matrix can only be estimated under multivariate normality. It would be odd to estimate the univariate parameters α and γ under the logistic distribution and then estimate the covariance of the error terms under multivariate normality.

The various alternatives are illustrated in the sections that follow.

5.7 A Small Example

Before proceeding to analyze the Efficacy variables, I illustrate the various alternatives by means of a small example based on generated data. File ORDATA.RAW contains data in free format on one ordinal variable y and two covariates x_1 and x_2 . To estimate the probit regression in the Standard Parameterization, use the following PRELIS command file (file ORDATA.PR2):

```
Data Ninputvars = 3
Labels
Y X1 X2
Rawdata = ORDATA.RAW
Continuous X1 X2
PR Y on X1 X2
Output
```

The probit regression is estimated as

Thresholds: -1.910 -0.962 0.662 2.050

$$Y = 1.006*X1 + 2.028*X2 + \text{Error}, R2 = 0.838$$

(0.0860)	(0.119)
11.696	16.997

To estimate the same regression in the Alternative Parameterization, just put AP on the Output line. This gives the following results:

Thresholds: 0.0 1.0 2.714 4.179

$$Y = 2.016 + 1.061*X1 + 2.141*X2 + \text{Error}, R2 = 0.838$$

(0.0907)	(0.126)
11.696	16.997

Note that

- The regression coefficients in the Standard and Alternative Parameterizations are different but the t -values are the same.
- R^2 is the same.
- Although different, the regression coefficients are rather close. However, this is just a coincidence that occurs because $\hat{\tau}_2 - \hat{\tau}_1$ is close to 1.

To use logit regression, put LR instead of PR. Logit regression gives the following results in the Standard Parameterization:

Thresholds: -3.417 -1.746 1.216 3.687

$$Y = 1.790*X1 + 3.631*X2 + \text{Error}, R2 = 0.943$$

(0.154)	(0.227)
11.599	15.977

Comparing the standard solutions for probit and logit regression, it is seen that the regression coefficients are quite different. However, a closer look shows that the regression coefficients of the logit equation are approximately 1.79 times those of the probit regression. This confirms the statement made earlier that the regression coefficients will be roughly proportional. The scale factor 1.79 may require some further explanation. The factor 1.7 in (28) should be regarded as an approximate population quantity, whereas the scale factor 1.79 is estimated from a random sample of 400 observations.

That the results of the probit and logit regressions are close can be seen much better if one uses the Alternative Parameterization. The result of logit regression in the Alternative Parameterization is:

Thresholds: 0.0 1.0 2.773 4.251

$$Y = 2.045 + 1.071*X1 + 2.173*X2 + \text{Error}, R2 = 0.943$$

(0.0924)	(0.136)
11.599	15.977

As can be seen, this is quite similar to the corresponding regression equation for the probit model.

5.8 Data Screening

I now return to the analysis of the Efficacy variables in the Political Action Survey described in Section 1. The data file for this illustration is `USA.RAW`. This contains 10 variables in free format. The first six are the six Efficacy variables; the other four variables are (the original variable names are given in parenthesis):

`YOB` Year of birth with *Don't Know* coded as 1998 and *No Answer* coded as 1999 (`V0146`). Recall that the interviews were done in 1974.

`GENDER` Gender coded as 1 for *Male*, 2 for *Female*, and 9 for *No Answer* (`V0283`).

`LEFTRIGH` A left-right scale from 1 to 10 with *Don't Know* coded as 98 and *No Answer* coded as 99 (`V0020`).

`EDUCAT` Education coded as 1 for *Compulsory level only*, 2 for *Middle level*, 3 for *Higher or Academic level*, and 9 for *No Answer* (`V0214`)

As always, it is a good idea to begin with a data screening. This can be done by running the following PRELIS command file (file `ORD51.PR2`)

```
Screening the Data in USA.RAW
Data Ninputvars = 10
Labels
NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST YOB GENDER LEFTRIGH EDUCAT
Rawdata = USA.RAW
Clabels NOSAY - INTEREST 1=AS 2=A 3=D 4=DS 8=DK 9=NA
Clabels GENDER 1=MALE 2=FEMA 9=NA
Clabels LEFTRIGH 98=DK 99=NA
Clabels EDUCAT 1=COMP 2=MIDD 3=HIGH 9=NA
Output
```

The output reveals that

- There are 1719 cases in the USA sample, 736 males and 983 females.
- The marginal distributions of the six efficacy variables are those reported in Section 1.
- There are more than 15 different birthyears in the sample. The oldest person was born in 1882. Only one person did not report his/her birthyear and nobody reported not knowing his/her year of birth.
- As many as 547 persons or 31.8% did not place themselves on the left-right scale.
- Only 8 persons did not answer the education question.

Before one can proceed one must decide how to treat the *Don't Know* and *No Answer* responses. In Section 1, I discussed various alternative ways of dealing with missing values. I do not want to repeat that discussion here. The major difficulty is to decide how to treat the 547 people who did not answer the `LEFTRIGH` variable. Does this mean that these people are in the middle of the scale, or that the concept of left-right has no meaning for them, or what? I do not know. So I will treat them as having provided no information.

File `ORD51A.PR2` eliminates all cases with *Don't Know* and *No Answer* responses (listwise deletion) and saves the data on all complete cases in a PRELIS system file called `USA.PSF`. In addition, `AGE` is computed as `1974 - YOB`. This is a proxy for age. The output file shows that the resulting listwise sample size is 1076. Thus, 643 cases were lost.

5.9 Probit Regression of NOSAY

The following PRELIS command file (file ORD52.PR2) estimates the probit regression of NOSAY (as a y -variable) on GENDER, LEFTRIGH, EDUCAT, and AGE (as x -variables – covariates) using the Alternative Parameterization.

```
Probit Regression of NOSAY
SY=USA.PSF
Continuous GENDER - AGE
PR NOSAY on GENDER - AGE
Output AP
```

The output gives the following information about the probit regression:

```
Univariate Probit Regression for    NOSAY
Alternative Parameterization
```

```
Thresholds: 0.0 1.0 2.825
```

```

NOSAY = 1.403 + 0.00927*GENDER + 0.0437*LEFTRIGH + 0.371*EDUCAT
          (0.0696)          (0.0189)          (0.0534)
          0.133           2.308           6.946

+ 0.00467*AGE + Error, R2 = 0.0586
          (0.00213)
          2.187
```

Because the t -values for LEFTRIGH, EDUCAT, and AGE are all positive and larger than 2, this means that people on the right of the left-right scale, people with higher education and older people have a tendency to respond higher on the ordinal scale for NOSAY, that is, they are likely to disagree or disagree strongly to NOSAY. This seems quite plausible.

The corresponding logit regression is obtained by replacing PR with LR. The resulting regression equation is

```
Univariate Logit Regression for    NOSAY
Alternative Parameterization
```

```
Thresholds: 0.0 1.0 2.771
```

```

NOSAY = 1.369 - 0.00285*GENDER + 0.0492*LEFTRIGH + 0.379*EDUCAT
          (0.0693)          (0.0188)          (0.0538)
          -0.0411          2.616           7.043

+ 0.00456*AGE + Error, R2 = 0.171
          (0.00212)
          2.147
```

The logit regression is very similar to the probit regression, but note that R^2 is larger for the logit model than for the probit model. I will discuss the issue of the fit of the probit vs logit model in the next Section.

In addition to the ordinary output file ORD52.OUT, the run of ORD52.PR2 gives another output file ORD52.FIT giving information about the fit of the probit and logit regressions. For the initial probit regression of NOSAY, this file looks like this:

Variable	-2lnL	Chi-square	df
NOSAY	2344.591	53.512	4

This does not give much information; only that the four covariates fit much better than no covariate at all. However, consider entering the covariates stepwise one at a time using the following input file (ORD52A.PR2:

```

Probit Regression of NOSAY
SY=USA.PSF
Continuous GENDER - AGE
PR NOSAY on GENDER
PR NOSAY on GENDER AGE
PR NOSAY on GENDER AGE LEFTRIGH
PR NOSAY on GENDER AGE LEFTRIGH EDUCAT
Output AP

```

The file ORD52A.FIT looks like this

Variable	-2lnL	Chi-square	df	Covariates
NOSAY	2398.100	0.004	1	GENDER
NOSAY	2395.940	2.163	2	GENDER AGE
NOSAY	2393.094	5.009	3	GENDER AGE LEFTRIGH
NOSAY	2344.591	53.512	4	GENDER AGE LEFTRIGH EDUCAT

This can be interpreted as follows. GENDER is no better than no covariate at all, *i.e.*, GENDER alone cannot be used to predict NOSAY. If AGE is used together with GENDER there is no significant improvement in fit. GENDER and AGE alone does not predict NOSAY. If LEFTRIGH is added to the equation, there is still no significant improvement in fit because $5.009 - 2.163 = 2.846$ is not significant as a chi-square with one degree of freedom. If EDUCAT is added to the equation, there is a highly significant improvement in fit. This suggest that EDUCAT is the best predictor of NOSAY. These findings are confirmed in the output file ORD52A.OUT.

How come that LEFTRIGH and AGE are significant in the last equation whereas they are not significant in any equation that does not include EDUCAT? The reason is that EDUCAT is correlated with LEFTRIGH and AGE thereby generating interactive effects of LEFTRIGH and AGE. This can be seen by entering the covariates one at a time in the opposite order, see file ORD52B.PR2. The fit file from this run is

Variable	-2lnL	Chi-square	df	Covariates
NOSAY	2357.968	40.136	1	EDUCAT
NOSAY	2349.465	48.638	2	EDUCAT LEFTRIGH
NOSAY	2344.609	53.494	3	EDUCAT LEFTRIGH AGE
NOSAY	2344.591	53.512	4	EDUCAT LEFTRIGH AGE GENDER

Recall that chi-square is a test of the hypothesis that none of the covariates has any effect. This hypothesis is rejected for any equation with EDUCAT included. Note that chi-square increases considerably when LEFTRIGH is added to EDUCAT and when AGE is added to EDUCAT and LEFTRIGH but not when GENDER is added. The chi-square difference $48.638 - 40.136 = 8.502$ with one degree of freedom is a test of the hypothesis that LEFTRIGH has no effect, given that EDUCAT is included. This hypothesis is rejected. Thus, LEFTRIGH should be included with EDUCAT. Similarly, the chi-square difference $53.494 - 48.638 = 4.856$ with one degree of freedom is a test of the hypothesis that AGE has no effect, given that EDUCAT and LEFTRIGH are included. This hypothesis is also rejected (at the 5% level). Thus, AGE should be included with EDUCAT and LEFTRIGH. But one cannot reject the

hypothesis that GENDER has no effect, given that EDUCAT, LEFTRIGH, and AGE are included in the equation because $53.512 - 53.494 = 0.018$ is not significant.

Using (22) and the estimated parameter values, one can compute estimated category probabilities for any specified set of covariate values. I illustrate this for the probit and logit regressions of NOSAY. Table 13 gives estimated probabilities for 16 different combinations of the four covariates for the probit model and Table 14 gives the same probabilities estimated under the logit model.

Table 13: Estimated Category Probabilities (probit)

Covariates				Probabilities			
Gender	LeftRight	Education	Age	AS	A	D	DS
1	2	1	30	0.026	0.138	0.622	0.215
1	2	1	60	0.018	0.114	0.611	0.257
1	2	3	30	0.004	0.041	0.483	0.472
1	2	3	60	0.003	0.031	0.440	0.527
1	8	1	30	0.014	0.095	0.595	0.296
1	8	1	60	0.010	0.076	0.570	0.345
1	8	3	30	0.002	0.024	0.401	0.573
1	8	3	60	0.001	0.017	0.356	0.626
2	2	1	30	0.025	0.136	0.621	0.218
2	2	1	60	0.018	0.112	0.610	0.260
2	2	3	30	0.004	0.040	0.480	0.476
2	2	3	60	0.002	0.030	0.437	0.530
2	8	1	30	0.013	0.093	0.594	0.299
2	8	1	60	0.009	0.074	0.568	0.348
2	8	3	30	0.002	0.023	0.398	0.577
2	8	3	60	0.001	0.017	0.353	0.629

Table 13 shows that a young male with low education and “leftist” opinion is most likely to respond *Disagree* ($P = 0.622$) to the NOSAY statement. This may be contrasted with an old male with high education and “rightist” opinion whose most likely response is *Disagree Strongly* ($P = 0.626$). It is also seen that any person is more likely to respond *Disagree* or *Disagree Strongly* than *Agree* or *Agree Strongly* no matter what his characteristics are. The probability of an *Agree Strongly* response is very small for all types of persons. Table 14 shows very similar probabilities, but note that all probabilities for *Agree Strongly* are larger than the corresponding probabilities in Table 13 and most of the probabilities for *Disagree Strongly* are larger in Table 14 than in Table 13. This is in line with the remark made earlier that the logit model gives more probability to the outer categories than the probit model.

5.10 Probit and Logit Regression of All Efficacy Variables

To analyze all the six efficacy variables jointly with the four covariates, just replace the PR line in ORD52.PR2 with (see ORD53.PR2 where the Standard Parameterization is used):

PR NOSAY - INTEREST on GENDER - AGE

A slight editing of the output file ORD53.OUT gives the following estimated probit regressions.

$$\begin{array}{l}
 \text{NOSAY} = 0.00899 \cdot \text{GENDER} + 0.0424 \cdot \text{LEFTRIGH} + 0.360 \cdot \text{EDUCAT} + 0.00453 \cdot \text{AGE} + \text{Error} \\
 \quad (0.0675) \qquad \quad (0.0184) \qquad \quad (0.0519) \qquad \quad (0.00207) \\
 \quad 0.133 \qquad \quad 2.308 \qquad \quad 6.946 \qquad \quad 2.187
 \end{array}$$

Table 14: Estimated Category Probabilities (logit)

Covariates				Probabilities			
Gender	LeftRight	Education	Age	AS	A	D	DS
1	2	1	30	0.032	0.124	0.638	0.207
1	2	1	60	0.026	0.102	0.625	0.248
1	2	3	30	0.009	0.039	0.463	0.489
1	2	3	60	0.007	0.031	0.414	0.547
1	8	1	30	0.020	0.080	0.598	0.302
1	8	1	60	0.016	0.065	0.566	0.353
1	8	3	30	0.005	0.024	0.357	0.613
1	8	3	60	0.004	0.019	0.309	0.667
2	2	1	30	0.032	0.123	0.637	0.208
2	2	1	60	0.025	0.101	0.625	0.249
2	2	3	30	0.009	0.039	0.462	0.490
2	2	3	60	0.007	0.031	0.413	0.549
2	8	1	30	0.020	0.080	0.598	0.303
2	8	1	60	0.015	0.065	0.565	0.355
2	8	3	30	0.005	0.024	0.356	0.615
2	8	3	60	0.004	0.019	0.308	0.668

$$\begin{aligned}
 \text{VOTING} &= -0.0344*\text{GENDER} - 0.0217*\text{LEFTRIGH} + 0.447*\text{EDUCAT} - 0.00634*\text{AGE} + \text{Error} \\
 &\quad (0.0667) \quad (0.0181) \quad (0.0515) \quad (0.00205) \\
 &\quad -0.516 \quad -1.195 \quad 8.673 \quad -3.096 \\
 \\
 \text{COMPLEX} &= -0.212*\text{GENDER} - 0.0233*\text{LEFTRIGH} + 0.494*\text{EDUCAT} + 0.000881*\text{AGE} + \text{Error} \\
 &\quad (0.0678) \quad (0.0184) \quad (0.0525) \quad (0.00207) \\
 &\quad -3.135 \quad -1.267 \quad 9.402 \quad 0.425 \\
 \\
 \text{NOCARE} &= -0.0402*\text{GENDER} + 0.0240*\text{LEFTRIGH} + 0.371*\text{EDUCAT} + 0.00288*\text{AGE} + \text{Error} \\
 &\quad (0.0669) \quad (0.0182) \quad (0.0514) \quad (0.00205) \\
 &\quad -0.600 \quad 1.319 \quad 7.219 \quad 1.407 \\
 \\
 \text{TOUCH} &= 0.0382*\text{GENDER} + 0.0118*\text{LEFTRIGH} + 0.290*\text{EDUCAT} + 0.00540*\text{AGE} + \text{Error} \\
 &\quad (0.0676) \quad (0.0184) \quad (0.0516) \quad (0.00207) \\
 &\quad 0.565 \quad 0.643 \quad 5.632 \quad 2.604 \\
 \\
 \text{INTEREST} &= 0.0316*\text{GENDER} - 0.00604*\text{LEFTRIGH} + 0.249*\text{EDUCAT} + 0.00467*\text{AGE} + \text{Error} \\
 &\quad (0.0672) \quad (0.0183) \quad (0.0511) \quad (0.00206) \\
 &\quad 0.470 \quad -0.331 \quad 4.864 \quad 2.266
 \end{aligned}$$

Thus,

- GENDER has a significant effect only for COMPLEX.
- LEFTRIGH is significant only for NOSAY.
- EDUCAT is significant for all the ordinal variables.
- AGE is significant for NOSAY, VOTING, TOUCH, and INTEREST. Note that the effect of AGE on VOTING is negative.

The fit file ORD53.FIT gives the following information about the fit of the probit regressions.

Variable	-2lnL	Chi-square	df	Covariates
NOSAY	2344.591	53.512	4	GENDER LEFTRIGH EDUCAT AGE
VOTING	2470.648	106.818	4	GENDER LEFTRIGH EDUCAT AGE
COMPLEX	2284.676	108.725	4	GENDER LEFTRIGH EDUCAT AGE
NOCARE	2401.907	53.767	4	GENDER LEFTRIGH EDUCAT AGE
TOUCH	2249.115	35.305	4	GENDER LEFTRIGH EDUCAT AGE
INTEREST	2338.651	26.163	4	GENDER LEFTRIGH EDUCAT AGE

The second column gives a deviance but since we have no base model to compare it with this does not provide any information about whether the probit model fits the data or not. The third and fourth column give a chi-square test of the hypothesis that all regression coefficients are zero. It is seen that this hypothesis is rejected for all ordinal variables. This is as it should be.

For comparison, I give the fit statistics for the logit regressions obtained by putting LR instead of PR in ORD53.PR2, see file ORD53A.PR2.

Variable	-2lnL	Chi-square	df	Covariates
NOSAY	2340.790	57.313	4	GENDER LEFTRIGH EDUCAT AGE
VOTING	2467.314	110.152	4	GENDER LEFTRIGH EDUCAT AGE
COMPLEX	2279.764	113.637	4	GENDER LEFTRIGH EDUCAT AGE
NOCARE	2392.594	63.081	4	GENDER LEFTRIGH EDUCAT AGE
TOUCH	2244.766	39.653	4	GENDER LEFTRIGH EDUCAT AGE
INTEREST	2334.276	30.538	4	GENDER LEFTRIGH EDUCAT AGE

Does the logit model fit better than the probit model? The answer is Yes, Yes, Yes, Yes, Yes, and Yes. The two models have the same number of parameters but the deviance is smaller for the logit model than for the probit model for all variables. Take NOSAY, for example. The difference in deviance is $2344.591 - 2340.790 = 3.801$. Note that one can obtain the same number as the difference between the two chi-squares in the reverse order: $57.313 - 53.512 = 3.801$.

5.11 Estimating the Joint Covariance Matrix

To estimate the joint covariance matrix of the continuous variables underlying the ordinal variables and the covariates as defined in Section 5.5, one must use a `Fixedvariables` command (or `FI` command for short), see Jöreskog & Sörbom (1999a, pp. 180–183). Instead of `Fixedvariables` one can write `Covariates`. In addition to all probit regressions, these commands give estimates of the conditional covariance matrix and the joint unconditional covariance matrix as defined in Section 5.5. File ORD54.PR2 illustrates this using the Standard Parameterization. It also shows how one can obtain the asymptotic covariance matrix of the joint unconditional covariance matrix. File ORD54.PR2 is

```
Computing Covariance Matrix
SY=USA.PSF
Covariates: GENDER - AGE
Output MA=CM CM=USA.CM AC=USA.ACC WP
```

All variables specified on the `Covariates:` line are automatically treated as continuous variables. All other variables are assumed to be ordinal.

The output file ORD54.OUT gives the conditional covariance matrix as

	NOSAY	VOTING	COMPLEX	NOCARE	TOUCH	INTEREST
	-----	-----	-----	-----	-----	-----
NOSAY	1.000					
VOTING	0.284 (0.034) 8.312	1.000				
COMPLEX	0.270 (0.035) 7.746	0.204 (0.035) 5.766	1.000			
NOCARE	0.567 (0.027) 21.089	0.223 (0.035) 6.412	0.379 (0.032) 11.720	1.000		
TOUCH	0.367 (0.033) 11.106	0.206 (0.035) 5.824	0.274 (0.035) 7.874	0.637 (0.024) 26.169	1.000	
INTEREST	0.460 (0.030) 15.117	0.200 (0.035) 5.677	0.305 (0.034) 8.963	0.657 (0.023) 28.183	0.674 (0.023) 29.221	1.000

In this case, when the Standard Parameterization is used, this is the correlation matrix of the error terms. All correlations are highly significant. This means that the covariates alone do not account for the correlations of the ordinal variables (or more correctly the variables underlying the ordinal variables). This is not surprising since we know from Section 2 that we need the latent variables *Efficacy* and *Respons* to account for these correlations. In Section 5.12 I will use these latent variables as well.

The output `ORD54.OUT` also gives the joint covariance matrix of the variables underlying the ordinal variables and the covariates. This is too large to list here. It is saved in the file `USA.CM` and its asymptotic covariance matrix is saved in the file `USA.ACC`. The covariance matrix `USA.CM` is an unconstrained covariance matrix just as a sample covariance matrix for continuous variables. It can therefore be used for modeling in LISREL just as if all variables were continuous. The only restriction is that the covariates must not be treated as indicators of latent variables. In LISREL, one can estimate the model either by WLS using the inverse of `USA.ACC` as a weight matrix or by ML using `USA.ACC` to correct standard errors and chi-square for non-normality.

In Section 2, I used the ordinal *Efficacy* variables to establish a measurement model for the two latent variables *Efficacy* and *Respons*. Now I will investigate to what extent the covariates affect these two latent variables. To investigate this, one can use a MIMIC model described in Section 5.12.

5.12 A MIMIC Model for *Efficacy* and *Respons*

The idea of a MIMIC model is that a set of possibly explanatory variables (covariates) affects latent variables which are indicated by other observed variables, in this case ordinal variables. Thus there are multiple indicators and multiple causes of latent variables, see Jöreskog & Goldberger (1975). For examples of MIMIC models with continuous indicators see Jöreskog & Sörbom (1999b). The MIMIC model considered here is shown in Fig. 10.

A SIMPLIS command file for estimating the model in Fig. 10 is `ORD55.SPL`:

```
MIMIC Model
Observed Variables: NOSAY VOTING COMPLEX NOCARE TOUCH INTEREST
```

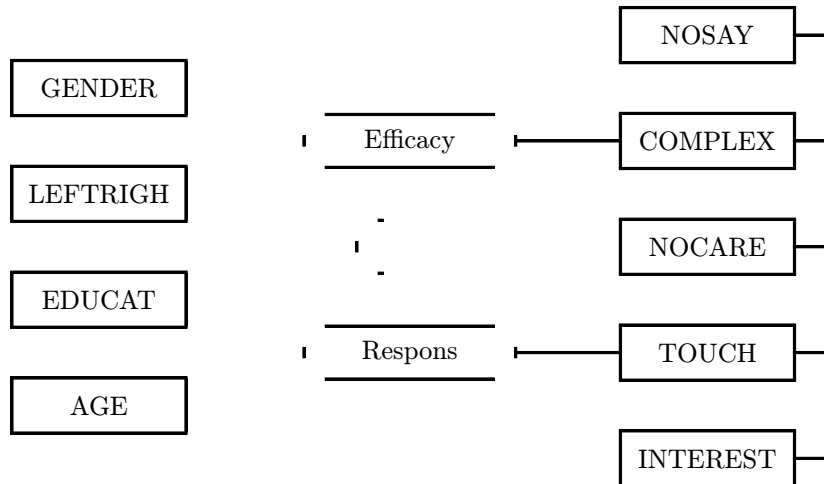


Figure 10: MIMIC Model for Efficacy and Respons

```

GENDER LEFTRIGH EDUCAT AGE
Covariance Matrix from File USA.CM
Asymptotic Covariance Matrix from File USA.ACC
Sample Size: 1076
Latent Variables: Efficacy Respons
Relationships:
  NOSAY COMPLEX NOCARE = Efficacy
  NOCARE TOUCH INTEREST = Respons
  NOSAY = 1*Efficacy
  INTEREST = 1*Respons
  Efficacy Respons = GENDER LEFTRIGH EDUCAT AGE
Let the errors of Efficacy and Respons correlate
Path Diagram

```

The output gives the structural equations as

$$\begin{aligned}
 \text{Efficacy} = & -0.19 \cdot \text{GENDER} - 0.025 \cdot \text{LEFTRIGH} + 0.43 \cdot \text{EDUCAT} - 0.00052 \cdot \text{AGE} \\
 & (0.065) \quad (0.019) \quad (0.066) \quad (0.0021) \\
 & -2.94 \quad -1.28 \quad 6.48 \quad -0.25
 \end{aligned}$$

$$\begin{aligned}
 \text{Respons} = & -0.075 \cdot \text{GENDER} - 0.031 \cdot \text{LEFTRIGH} + 0.30 \cdot \text{EDUCAT} + 0.0021 \cdot \text{AGE} \\
 & (0.068) \quad (0.021) \quad (0.057) \quad (0.0023) \\
 & -1.11 \quad -1.45 \quad 5.32 \quad 0.90
 \end{aligned}$$

which shows that GENDER has a significant effect on Efficacy and EDUCAT has significant effects on both Efficacy and Respons. LEFTRIGH and AGE have no significant effects on either of the latent variables. The fact that they are non-significant does not mean they do not exist, only that the sample size is not sufficiently large to make them significant.

The model fits the data reasonably well as judged by the following fit statistics. For this conclusion I use the information about RMSEA and the guidelines of Browne & Cudeck (1993).

```

Degrees of Freedom = 15
Minimum Fit Function Chi-Square = 40.95 (P = 0.00032)
Root Mean Square Error of Approximation (RMSEA) = 0.040
90 Percent Confidence Interval for RMSEA = (0.026 ; 0.055)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.85

```

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.
- Barnes, S.H., & Kaase, M. (Eds.) (1979) *Political action: Mass participation in five western democracies*. Beverly Hills and London: Sage Publications.
- Brown, R.L. (1994) Efficacy of the indirect approach for estimating structural equation models with missing data. *Structural Equation Modeling*, 1, 287–316.
- Browne, M. W., & Cudeck, R. (1993) Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.): *Testing Structural Equation Models*. Sage Publications.
- Campbell, A., Gurin, G., & Miller, W.A. (1954). *The voter decides*. New York: Wiley.
- Costner, H.L. (1969) Theory, deduction, and rules of correspondence. *American Journal of Sociology*, 75, 245–263.
- Craig, S.C., & Maggiotto, M.A. (1981) Measuring political efficacy. *Political Methodology*, 8, 85–109.
- Du Toit, M., & Du Toit, S. (2001) *Interactive LISREL: User's Guide*. Chicago: Scientific Software International.
- Jöreskog, K.G., & Goldberger, A.S. (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 10, 631–639.
- Jöreskog, K.G., & Sörbom, D. (1999a) PRELIS 2 *User's Reference Guide*. Chicago: Scientific Software International.
- Jöreskog, K.G., & Sörbom, D. (1999b) LISREL 8: *Structural Equation Modeling with the SIMPLIS Command Language*. Chicago: Scientific Software International.
- Jöreskog, K.G., & Sörbom, D. (1999c) LISREL 8 *User's Reference Guide*. Chicago: Scientific Software International.
- Jöreskog, K.G., Sörbom, D., Du Toit, S., & Du Toit, M. (2001) LISREL 8: *New Statistical Features*. Third printing with revisions. Chicago: Scientific Software International.
- Lord, F.M., & Novick, M.E. (1968) *Statistical Theories of Mental Test Scores*. Reading: Addison-Wesley Publishing Co.
- McCullagh, P., & Nelder, J.A. (1983) *Generalized Linear Models*. London: Chapman & Hall.
- Miller, W.E., Miller A.H., & Schneider, E.J. (1980) *American National Election Studies Data Source Book: 1952–1978*. Cambridge, Massachusetts: Harvard University Press.
- Olsson, U. (1979) Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44, 443–460.
- Quiroga, A. (1992) *Studies of the polychoric correlation and other correlation measures for ordinal variables*. Doctoral dissertation at Uppsala University.
- Rao, C.R. (1965) *Linear statistical inference and its applications* New York: Wiley.
- Steiger, J.H. (1990) Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180.

Appendix 1: Derivation of the RMSEA Measure

In this Appendix, I sketch the derivation of the RMSEA measure and the P -value for test of close fit for underlying bivariate normality.

Let a run over all cells in the bivariate contingency and let $\pi_a = \pi_a(\boldsymbol{\theta})$ be the model and p_a be the sample proportion. Suppose the sample proportions \mathbf{p} converge in probability to $\boldsymbol{\pi}_0$ as the sample size increases. Fitting the model $\boldsymbol{\pi}(\boldsymbol{\theta})$ to $\boldsymbol{\pi}_0$ gives $\hat{\boldsymbol{\theta}}_0$ and $\hat{\boldsymbol{\pi}}_0 = \boldsymbol{\pi}(\hat{\boldsymbol{\theta}}_0)$. The model holds if $\hat{\boldsymbol{\pi}}_0 = \boldsymbol{\pi}_0$ and the model does not hold if $\hat{\boldsymbol{\pi}}_0 \neq \boldsymbol{\pi}_0$. Fitting the model $\boldsymbol{\pi}(\boldsymbol{\theta})$ to the sample data \mathbf{p} gives $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\pi}} = \boldsymbol{\pi}(\hat{\boldsymbol{\theta}})$. We can now distinguish between three kinds of errors:

- Approximation Error: $\hat{\boldsymbol{\pi}}_0 - \boldsymbol{\pi}_0$
- Estimation Error: $\hat{\boldsymbol{\pi}} - \hat{\boldsymbol{\pi}}_0$
- Total Error: $\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}_0$

To define measures of these errors, I write the fit function (6) as a general fit function $F(\mathbf{p}, \mathbf{q})$ of two vectors of probabilities \mathbf{p} and \mathbf{q} and assume that F is non-negative and zero if and only if $\mathbf{p} = \mathbf{q}$. We can then define measures of the three kinds of errors as

$$\mathbf{A} = F(\boldsymbol{\pi}_0, \hat{\boldsymbol{\pi}}_0)$$

$$\mathbf{E} = F(\hat{\boldsymbol{\pi}}, \hat{\boldsymbol{\pi}}_0)$$

$$\mathbf{T} = F(\hat{\boldsymbol{\pi}}, \boldsymbol{\pi}_0)$$

Here \mathbf{A} is a constant independent of sample data, whereas \mathbf{E} and \mathbf{T} are random variables depending of sample data. \mathbf{A} is a measure of the degree of approximation in the population. I will estimate a transformation of \mathbf{A} .

Let $c = 2NF[\mathbf{p}, \boldsymbol{\pi}(\hat{\boldsymbol{\theta}})] = 2N\hat{F}$. Then, as stated previously, if the model holds exactly, c has approximately a chi-square distribution with d degrees of freedom. However, if the model does not hold exactly but holds approximately, then c has a non-central chi-square distribution with d degrees of freedom and non-centrality parameter λ , where ¹⁶

$$\lambda = 2NF[\boldsymbol{\pi}_0, \boldsymbol{\pi}(\boldsymbol{\theta}_0)] = 2NF_0 . \quad (41)$$

The non-centrality parameter λ can be estimated as

$$\hat{\lambda} = \max(c - d, 0) . \quad (42)$$

F_0 is the required measure of approximation in the population. Since it is approximately the sum of d squares, it is more convenient to define the *Root Mean Squared Error of Approximation (RMSEA)* (Steiger, 1990)

$$\epsilon_a = \sqrt{F_0/d} = \sqrt{\lambda/2Nd} , \quad (43)$$

which can be estimated as

$$\mathbf{RMSEA} = \hat{\epsilon}_a = \sqrt{\hat{\lambda}/2Nd} = \sqrt{\max[(c - d)/2Nd, 0]} , \quad (44)$$

Practical experience based on a large number of contingency tables suggests that a reasonably acceptable degree of approximation in the population is $\epsilon_a \leq 0.1$. This represents a close fit in contrast to $\epsilon_a = 0$ which represents exact fit. The P -value for a test of close fit is obtained as

$$P = 1.0 - G(c \mid 0.1^2Nd, d) , \quad (45)$$

where $G(x \mid \lambda, d)$ is the non-central chi-square distribution function with non-centrality parameter λ and degrees of freedom d .

¹⁶Sik-Yum Lee (personal communication) gave me a handwritten proof of this statement.

Appendix 2: Questions and Answers

Here are some questions often asked:

Q: Why cannot I just add the items up and use this as a measure of *Efficacy*?

A: You can. But you should first investigate whether the items measure a single unidimensional latent variable. If they do not you will have a validity problem, *i.e.*, the composite variable does not measure what you think it does.

Q: Why cannot I just dichotomize all variables and use tetrachoric correlations instead of polychoric correlations?

A: You can, but it is questionable whether you gain anything by doing so. If respondents can distinguish between *Agree* and *Agree Strongly* and between *Disagree* and *Disagree Strongly* you will throw away information. Collapsing categories is a good idea if there are many zero cells in the bivariate contingency tables.

Q: Why cannot I just use the numbers 1, 2, 3, and 4 as they are and compute an ordinary covariance matrix or correlation matrix?

A: That assumes that the numbers 1, 2, 3, and 4 are on an interval scale, *i.e.*, that 2 is twice as much as 1, 3 is three times 1, etc. But here, 1, 2, 3, and 4 are just labels for a set of ordered categories. Even if one were to assume interval scale properties, the distribution is discrete and therefore not continuous (only four values are possible; there are no values between the numbers).

Q: Why cannot I just compute the polychoric correlation matrix and use this with the ML method in LISREL? Why do I have to compute the asymptotic covariance matrix?

A: The ML method in LISREL assumes that we have a sample covariance matrix from a multivariate normal distribution. This is not the case here. The asymptotic covariance matrix is the price you pay for non-normality. The WLS method based on the polychoric correlations and its asymptotic covariance matrix gives correct standard errors and chi-squares in large samples.

Q: But the asymptotic covariance matrix requires a large sample to estimate. How large a sample do I need?

A: For a small data set like this, probably 400 is OK. With more variables the sample must be larger. It is difficult to give a general guideline because it depends on both number of categories and number of variables. If there are many zero cells in the bivariate contingency tables, the sample size should be considered too small.

If the asymptotic covariance matrix is poorly estimated due to a small sample, the inverse of it is even more severely affected thereby making the WLS method unstable. In this situation, a reasonable compromise is to fit the model by ML and correct the standard error and chi-squares by using the asymptotic covariance matrix as described in Section 2.6. Then the asymptotic covariance matrix need not be inverted.

Q: What should I do if I have many ordinal variables and a small sample?

A: What do you want to know? You may not be able to use LISREL to estimate an elaborate model using *all* your ordinal variables, but PRELIS data screening provides a lot of information about the data which could be useful.

If you have a strong measurement theory so that you know which latent variables you want to measure and which observed variables might be used as indicators of each latent variable, I suggest you analyze the indicators of one latent variable at a time to investigate if they are unidimensional. The result of this investigation will provide information about which variables to eliminate if necessary. You can then form two or three subscales by adding indicators. These subscales can then be used as indicators in a LISREL model with all the latent variables.

If you do not have a strong measurement theory so that you don't know what the ordinal variables are supposed to measure, I suggest that you do an exploratory factor analysis of the polychoric correlation matrix and then proceed as in the previous paragraph for each factor clearly identified and interpreted.