LAB NOTES: EXAMPLES OF PRELIS RUNS

PRELIS 2 is a data preprocessor for processing data in preparation for estimating a structural equation model in LISREL 8 or 9. For information on reading data into PRELIS, see “Lab Notes: Inputting Data into PRELIS” available on the course webpage. In this note, I will give some common examples of PRELIS runs. PRELIS provides two important tasks: (1) It provides a useful descriptive data diagnostics in preparation for an SEM model, including histograms and bivariate tables for ordinal and categorical indicators, tests of univariate and multivariate normality (including univariate tests of skewness and kurtosis), descriptives on patterns of missing values, and provides imputation methods for missing values. (2) it computes and saves moment matrices (covariance and correlation matrices) as well as asymptotic covariance or variance matrices (for ordinal and non-normal data) for input into LISREL to obtain estimates of SEM under correct scale (continuous, ordered, dichotomous) and distribution (normal vs. non-normal) assumptions. LECTURE 5 presents a description of PRELIS and LISREL commands (as well as a brief overview of SIMPLIS commands). Here are a couple of illustrations of major PRELIS runs, with a couple of examples of resulting LISREL runs. Annotations are in bold font. At the end of this memo are annotated output from the two PRELIS outputs.

Example 1: Estimating a covariance matrix under the assumption of interval scales and multivariate normality. **(PRELIS Commend file “preexch1.PR2.”)**

!PRELIS Exchange Assume Interval: exch1 **All comments are preceded by ! This is the title**
DA NI=4 **There are four input variables; I didn’t specify an N,**
LA **Labels to follow** **so PRELIS will determine that for me.**
watched borrow help asked **Here are the four labels**
! How often have you watched your neighbor's property when they were out of town? **These details**
! How often have you borrowed tools or small food items from your neighbors? **(comments)**
! How often have you helped a neighbor with a problem? **are for me (to**
! How often have you asked neighbors about personal things like child rearing or jobs? **jog my**
RA FI=h:\529data\lisrel103107-ob.dat **memory).**
CO 1-4 **The raw data are in a file called lisrel103107-ob.dat**
OU MA=CM CM=h:\529data\lisrel-ob.CM **All four variables are treated as continuous**.
I’m asking to compute a covariance matrix of the four observed variables and save them in a file called lisrel-ob.CM, which can be read into LISREL

Example 2: Treating all four indicators as ordinal variables and computing a 4 x 4 polychoric correlation matrix (with 10 elements) and a 10 x 10 asymptotic covariance matrix (with 55 elements) for those 10 polychoric correlations. **(PRELIS Commend file “preexch2.PR2.”)**

!PRELIS Exchange Assume Ordinal: exch2
DA NI=4
LA
watched borrow help asked
! How often have you watched your neighbor's property when they were out of town?
! How often have you borrowed tools or small food items from your neighbors?
! How often have you helped a neighbor with a problem?
! How often have you asked neighbors about personal things like child rearing or jobs?

RA FI=h:\529data\lisrel103107-ob.dat
OR 1-4
OU MA=PM PM=h:\529data\lisrel-ob.PM SA=h:\529data\ACOV.PM1 **All four variables are treated as ordinal. I'm asking to compute a polychoric correlation matrix of the four observed variables and save them in a file called lisrel-ob.PM, which can be read into LISREL. I'm also asking to compute the asymptotic covariance matrix of the polychoric correlations and saave it in a file called lisrel103107-ob.dat.**

Example 3: LISREL run that reads in the covariance matrix saved above (I moved it to a different directory on my c: drive) and estimates a one factor confirmatory factor model using ML. This model assumes observable variables are continuous and distributed multivariate normal. (**LISREL command file "ML_CFA.LS8.")**

!LISREL Exchange CFA 1 err-corr ML_CFA1A
DA NI=4 NO=4670 MA=CM **I am analyzing a covariance matrix**
CM FI=c:\529\lisrel-ob.CM **I am reading in the covariance matrix from above.**
LA
*
watched borrow help asked
! How often have you watched your neighbor's property when they were out of town?
! How often have you borrowed tools or small food items from your neighbors?
! How often have you helped a neighbor with a problem?
! How often have you asked neighbors about personal things like child rearing or jobs?
SE
1 2 3 4
MO NY=4 NE=1 LY=FU, FI PS=SY, FR TE=SY, FI
VA 1.0 LY 1 1
FR LY 2 1 LY 3 1 LY 4 1
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 2 4
PATH DIAGRAM
OU ME=ML SE TV SC RS MI **I'm asking for maximum likelihood estimation**

Example 4: LISREL run that reads in the polychoric correlation matrix and asymptotic covariance matrix saved above (I moved it to a different directory on my c: drive) and estimates a one factor confirmatory factor model using WLS. This model assumes observable variables are each measured on ordinal scales and uses WLS to obtain consistent and asymptotically efficient estimates without the assumption of multivariate normality of observed variables. (**LISREL command file "ML_CFA.LS8.")**

!LISREL Exchange CFA 1 err-corr ML_CFA1A
DA NI=4 NO=4670 MA=PM **I'm analyzing a polychoric correlation matrix.**
PM=c:\529\lisrel-ob.PM **I'm reading in the polychoric matrix from above.**
AC=c:\529\ACOV.PM1 **I'm reading in the asymptotic covariance matrix from above.**
LA

```
*
watched borrow help asked
! How often have you watched your neighbor's property when they were out of town?
! How often have you borrowed tools or small food items from your neighbors?
! How often have you helped a neighbor with a problem?
! How often have you asked neighbors about personal things like child rearing or jobs?
SE
1 2 3 4
MO NY=4 NE=1 LY=FU, FI PS=SY, FR TE=SY, FI
VA 1.0 LY 1 1
FR LY 2 1 LY 3 1 LY 4 1
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 2 4
PATH DIAGRAM
OU ME=WL SE TV SC RS MI
```

I'm choosing weighted least squares (WL) as an estimator, which requires that I have read the moment matrix and the asymptotic covariance matrix.

DATE: 05/09/2012
TIME: 23:23

P R E L I S 2.80

BY

Karl G. Jöreskog & Dag Sörbom

The following lines were read from file H:\529 examples\preexch1.LS8:

!PRELIS Exchange Assume Interval: exch1

DA NI=4 — **I am inputting four observed variables**

LA — **I'm giving labels to the variables in free format (the default)**

watched borrow help asked

! How often have you watched your neighbor's property when they were out of town?
! How often have you borrowed tools or small food items from your neighbors?
! How often have you helped a neighbor with a problem?
! How often have you asked neighbors about personal things like child rearing or jobs?

Here I'm writing out the entire survey question, so I have it in the output. Note the exclamation point means this is a comment line, not a LISREL command.

RA FI=h:\529data\lisrel103107-ob.dat — **Here, I'm reading the raw data file RA(w) with FI(lename) = lisrel103107-ob.dat in folder h:\529data**

CO 1-4 — **I am treating all four variables as continuous variables This will give me skewness and kurtosis measures.**

OU MA=CM CM=h:\529data\lisrel-ob.CM — **This is the OU(put) line, in which I am specifying that I want to compute a covariance matrix (MA=CM) and save it with filename lisrel-ob.cm in the folder h:\529data\.**

Total Sample Size = 4670 — **PRELIS gives me the sample size.**

Univariate Summary Statistics for Continuous Variables

Variable	Mean	St. Dev.	T-Value	Skewness	Kurtosis	Minimum	Freq.	Maximum	Freq.
watched	2.015	0.766	179.663	-0.025	-1.297	1.000	1337	3.000	1406
borrow	1.722	0.685	171.822	0.419	-0.851	1.000	1925	3.000	626
help	2.046	0.606	230.755	-0.022	-0.289	1.000	754	3.000	971
asked	1.633	0.699	159.720	0.644	-0.764	1.000	2310	3.000	598

T-value gives the t-statistic for testing whether the mean is zero. Skewness is a sample estimate of the standardized population third moment about the mean ($\mu_3 = E(X - \mu_1)^3$ For normally-distributed skewness is zero. If it is positive, the distribution is skewed to the left; if it is negative, the distribution is skewed to the right. Kurtosis is a sample estimate of the standardized population fourth moment about the mean $\mu_4 = E(X - \mu_1)^4$ minus three. For normally distributed variables, kurtosis is equal to three. If it is greater than 3, it has positive kurtosis, resulting in thinner tails than normal; if it is less than 3, it has negative kurtosis, resulting in fatter tails than normal. The measure of kurtosis here is standardized and then subtracts 3. Therefore, a normally distributed variable will have kurtosis of zero; positive values means positive kurtosis (thin tails); negative values means negative kurtosis (fat tails).

Test of Univariate Normality for Continuous Variables

This gives significance tests for whether skewness and kurtosis depart from zero (normality) .

Variable	Skewness Z-Score	Skewness P-Value	Kurtosis Z-Score	Kurtosis P-Value	Skewness and Kurtosis Chi-Square	Skewness and Kurtosis P-Value
watched	-0.699	0.485	-18.098	0.000	328.027	0.000
borrow	11.258	0.000	-11.875	0.000	267.747	0.000
help	-0.606	0.545	-4.027	0.000	16.587	0.000
asked	16.518	0.000	-10.662	0.000	386.525	0.000

Relative Multivariate Kurtosis = 0.959

This is Mardia's measure of multivariate kurtosis (see lecture notes p. 201).

Test of Multivariate Normality for Continuous Variables

This gives significance tests for Mardia's measure of multivariate skewness and kurtosis (see lecture notes p. 201).

Skewness Value	Skewness Z-Score	Skewness P-Value	Kurtosis Value	Kurtosis Z-Score	Kurtosis P-Value	Skewness and Kurtosis Chi-Square	Skewness and Kurtosis P-Value
0.826	20.787	0.000	23.026	-5.157	0.000	458.692	0.000

Histograms for Continuous Variables

watched

Frequency	Percentage	Lower Class Limit
1337	28.6	1.000
0	0.0	1.200
0	0.0	1.400
0	0.0	1.600
1927	41.3	1.800
0	0.0	2.000
0	0.0	2.200
0	0.0	2.400
0	0.0	2.600
1406	30.1	2.800

borrow

Frequency	Percentage	Lower Class Limit
1925	41.2	1.000
0	0.0	1.200
0	0.0	1.400
0	0.0	1.600
2119	45.4	1.800
0	0.0	2.000
0	0.0	2.200
0	0.0	2.400
0	0.0	2.600
626	13.4	2.800

help

Frequency	Percentage	Lower Class Limit
754	16.1	1.000
0	0.0	1.200
0	0.0	1.400
0	0.0	1.600
0	0.0	1.800
2945	63.1	2.000
0	0.0	2.200
0	0.0	2.400
0	0.0	2.600
971	20.8	2.800

asked

Frequency	Percentage	Lower Class Limit
2310	49.5	1.000
0	0.0	1.200
0	0.0	1.400
0	0.0	1.600
0	0.0	1.800
1762	37.7	2.000
0	0.0	2.200
0	0.0	2.400
0	0.0	2.600
598	12.8	2.800

Covariance Matrix

This is the covariance matrix of observed variables, assuming interval scales.

	watched	borrow	help	asked
watched	0.587			
borrow	0.226	0.469		
help	0.226	0.173	0.367	
asked	0.172	0.213	0.159	0.488

Means

watched	borrow	help	asked
2.015	1.722	2.046	1.633

Standard Deviations

watched	borrow	help	asked
0.766	0.685	0.606	0.699

```
DATE: 05/17/2012
                                    TIME: 16:47

                                 P R E L I S  2.80

                                         BY

                          Karl G. Jöreskog & Dag Sörbom



                     This program is published exclusively by
                     Scientific Software International, Inc.
                        7383 N. Lincoln Avenue, Suite 100
                         Lincolnwood, IL 60712, U.S.A.
             Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140
         Copyright by Scientific Software International, Inc., 1981-2006
           Use of this program is subject to the terms specified in the
                          Universal Copyright Convention.
                            Website: www.ssicentral.com

 The following lines were read from file H:\529 examples\preexch2.LS8:

 !PRELIS Exchange Assume Ordinal: exch2
 DA NI=4
 LA
 watched borrow help asked
 ! How often have you watched your neighbor's property when they were out of town?
 ! How often have you borrowed tools or small food items from your neighbors?
 ! How often have you helped a neighbor with a problem?
 ! How often have you asked neighbors about personal things like child rearing or jobs?
 RA FI=h:\529data\lisrel103107-ob.dat
```

Reading in Raw data

```
 OR 1-4
 OU MA=PM PM=h:\529data\lisrel-ob.PM SA=h:\529data\ACOV.PM1
```

PM Saves the matrix of polychoric correlations in file lisrel-ob.PM in folder h:\529data\ and the asymptotic covariance matrix of the polychoric correlations in file ACOV.PM in the same file.

```
 Total Sample Size =   4670

 Univariate Marginal Parameters

 Variable     Mean St. Dev.   Thresholds
 --------     ---- --------   ----------
  watched    0.000    1.000  -0.564   0.521
   borrow    0.000    1.000  -0.222   1.107
     help    0.000    1.000  -0.988   0.814
    asked    0.000    1.000  -0.013   1.136
```

These are thresholds for the three-category ordinal variables

```
 Univariate Distributions for Ordinal Variables
```

Histograms for the ordinal variables

```
  watched Frequency Percentage Bar Chart
      1    1337        28.6
      2    1927        41.3
      3    1406        30.1

   borrow Frequency Percentage Bar Chart
      1    1925        41.2
      2    2119        45.4
      3     626        13.4

     help Frequency Percentage Bar Chart
      1     754        16.1
      2    2945        63.1
      3     971        20.8

    asked Frequency Percentage Bar Chart
      1    2310        49.5
      2    1762        37.7
      3     598        12.8

 There are 75 distinct response patterns, see FREQ-file.
```

Patterns in the data

```
 The 20 most common patterns are :
```

502	2	2	2	2
372	1	1	1	1
336	2	2	2	1
331	1	1	2	1
330	2	1	2	1
216	3	2	2	2
182	3	3	3	3
172	2	1	2	2
148	1	2	2	1
132	3	2	2	1
118	1	2	2	2
118	1	1	2	2
105	3	1	2	1
105	3	2	3	2
102	2	1	1	1
82	3	3	3	2
73	2	2	2	3
72	3	2	3	1
69	3	1	3	1
67	2	2	3	2

Correlations and Test Statistics

(PE=Pearson Product Moment, PC=Polychoric, PS=Polyserial)

Variable vs. Variable	Correlation	Test of Model Chi-Squ.	D.F.	P-Value	Test of Close Fit RMSEA	P-Value
borrow vs. watched	0.539 (PC)	74.264	3	0.000	0.071	1.000
help vs. watched	0.611 (PC)	64.231	3	0.000	0.066	1.000
help vs. borrow	0.536 (PC)	138.475	3	0.000	0.098	0.561
asked vs. watched	0.410 (PC)	59.425	3	0.000	0.063	1.000
asked vs. borrow	0.557 (PC)	62.373	3	0.000	0.065	1.000
asked vs. help	0.491 (PC)	71.643	3	0.000	0.070	1.000

This is a test of bivariate normality between the two latent variables; in each case we reject the null of bivariate normality; in each case we reject he hypothesis of normality; RMSEA < .05 suggests a reasonable fit in a large sample

Percentage of Tests Exceeding 0.5% Significance Level: 0.0%

Percentage of Tests Exceeding 1.0% Significance Level: 0.0%

Percentage of Tests Exceeding 5.0% Significance Level: 0.0%

Correlation Matrix

	watched	borrow	help	asked
watched	1.000			
borrow	0.539	1.000		
help	0.611	0.536	1.000	
asked	0.410	0.557	0.491	1.000

Means

watched	borrow	help	asked
0.000	0.000	0.000	0.000

Standard Deviations

watched	borrow	help	asked
1.000	1.000	1.000	1.000

The Problem used 5184 Bytes (= 0.0% of available workspace)