November 29, 2007
SC705: Advanced Statistics
Instructor: Natasha Sarkisian
Class notes: Diagnostics and Model Building Strategies

Model Diagnostics

The model diagnostics and improvement strategies discussed here apply to both measurement and structural models. Remember that you should always examine your data and perform the necessary transformations before you start estimating SEM – explore univariate distributions, bivariate relationships, and multivariate models (using OLS techniques). Only after you fixed various potential problems should you proceed to estimating SEM. After you estimated such a model, there are some additional diagnostics to consider.

<u>I. Examining individual estimates</u>

1. Assessing the parameter estimates

One of the first steps is to determine the viability of the estimated values of parameters – they should exhibit the correct sign and size, and be consistent with the underlying theory. Any estimates falling outside the acceptable range (e.g. correlations higher than 1, negative variances – known as Heywood cases) indicate a problem with the model.

2. Assessing the standard errors

Another indicator of poor model fit is the presence of standard errors that are excessively large or small. If they approximate zero, the test statistic for the parameter cannot be defined; if they are extremely large, this means the parameter cannot be determined. There are no clear guidelines as to what's too large or too small because standard errors are influenced by the units of measurement of the respective variables. Inaccurate standard errors are especially common when analyses are based on the correlation matrix.

3. Statistical significance of parameter estimates

Nonsignificant parameters, with the exception of error variances, can be considered unimportant to the model, and, in the interest of parsimony, they should be deleted (provided there is a sufficient sample size to be able to rely on significance testing).

4. Squared multiple correlations

For the structural model, two sets of squared multiple correlations are calculated – a set calculated from the structural model and a set calculated from the reduced form model. Those from the structural model are the R^2 values indicating the % of variance in each endogenous variable explained by all the variables used in its model (both exogenous and other endogenous – i.e. it takes into account both betas and gammas). Those from the reduced form model are the R^2 values indicating the % of variance in each endogenous variable explained by the exogenous variables only (note: reduced form model recalculates the equations to express the endogenous variables solely in terms of the exogenous ones). It is more appropriate to report and interpret reduced form R^2, especially if you deal with non-recursive models or correlated disturbance terms.

For the measurement model (for Xs and for Ys), squared multiple correlations are R^2 values indicating the % of variance of each indicator explained by the latent factor. They serve as reliability indicators of the extent to which each observed variable measures its latent factor.

II. Assessing model as a whole -- goodness-of-fit statistics

A range of goodness-of-fit statistics exists for SEM (see handout, pp.240-241 from Maruyama, Geoffrey M. 1998. Basics of Structural Equation Modeling. Thousand Oaks, CA: Sage Publications). This diversity can be quite confusing. We'll discuss a few indices and the guidelines for using them.

1. Chi-square is the likelihood ratio test statistic. It tests the null hypothesis that the variance-covariance matrix estimated from our model doesn't differ from the observed one (i.e. that all residuals are zero). This test, however, is sensitive to sample size and detects even minor deviations when the sample size is large). Some also use chi-square to d.f. ratio – a smaller ratio indicates a better fit.

2. Non-centrality parameter (NCP, symbolized by λ) is a measure of discrepancy between the observed variance-covariance matrix and the estimated one. It is therefore a measure of "badness-of-fit."

3. RMSEA (Root Mean Square Error of Approximation). This goodness-of-fit index asks "How well would the model fit the population covariance matrix if it were available?" It measures that discrepancy per degree of freedom. Values less than .05 indicate good fit, and values as high as .08 represent reasonable errors of approximation in the population. .08-.10 indicates mediocre fit, and greater than .10 – poor fit. LISREL also reports the confidence interval for RMSEA that should be taken into account when making a judgment. LISREL also provides a p-value for this statistic; the suggested cutoff for p-value, however, is >.50. This index can be used to compare non-nested models (nested models are those that have similar structure, with the only difference being the number of free parameters).

4. Expected Cross-Validation Index (ECVI). This index tries to assess, in a single sample, the likelihood that the model cross-validates across similar-sized samples from the same population. It is usually used in a multiple-model setup, where the model with the smallest ECVI has the greatest potential for replication. At the very least, we can compare it with the values for a saturated model (the least restricted, just-identified model) and independence model (the most restricted model -- model assuming null correlations among all variables in the model). This index can also be used to compare non-nested models.

5. AIC and CAIC (Akaike's Information Criterion and Consistent Akaike's Information Criterion). These criteria address the issue of parsimony, combining the goodness-of-fit measure with the information on the number of estimated parameters. AIC carries a penalty related to the degrees of freedom but not the sample size, while CAIC takes the sample size into account as well. These two indices are usually used when comparing multiple models. The smaller values represent better fit. These indices can be used to compare non-nested models.

6. NFI (Normed Fit Index) has been the criterion of choice for a long time, but recent evidence showed that it has tendency to underestimate fit in small samples. This index compares fits of two different (nested) models (the default presented in LISREL is the null model). Values for NFI range from zero to 1. A values of >.90 indicates an acceptable fit. The NNFI takes the complexity of model (number of parameters) into account, but because it's not normed (can go beyond 1) it is difficult to interpret. The Parsimony Normed Fit Index (PNFI) also attempts to adjust the NFI; it is normed but typically, parsimony-based indices have substantially lower values that the threshold levels generally perceived as acceptable for other normed indices of fit. The CFI (Comparative Fit Index) takes sample size into account so it avoids the problems of NFI. The numeric value of CFI is interpreted the same way as for NFI. The IFI (Incremental Index of Fit) was developed to address both the issue of parsimony and sample size, it is also interpreted the same way. Finally, the Relative Fit Index (RFI) is algebraically equivalent to CFI.

7. Critical N (CN) focuses directly on the adequacy of sample size rather than on model fit. It tests what sample size would be sufficient to yield an adequate model fit for a chi-square test. CN value in excess of 200 is indicative of a model that adequately represents the sample data.

8. The Root Mean Square Residual (RMR) represents the average residual value. It is best interpreted in the metric of correlation matrix (i.e. it is the residual for correlations rather than covariances). In a well-fitting model, this value will be small – .05 or less.

9. GFI (Goodness of Fit Index) is an absolute measure – it measures the amount of variance and covariance explained by the model (compared with null model). AGFI is similar, but it adjusts for the number of degrees of freedom in the specified model (i.e. accounts for parsimony). Both indices range from 0 to 1, with values close to 1 being indicative of good fit (although theoretically, it is possible for them to be negative as well when the model is worse than no model at all). PGFI (Parsimony Goodness of Fit Index) takes into account the number of estimated parameters when assessing goodness-of-fit. As mentioned above, parsimony-based indices have substantially lower values that the threshold levels generally perceived as acceptable for other normed indices of fit.

III. Evaluating model misspecifications

To identify potential model misfit, one can examine residuals and modification indices.

1. Residuals

These are obtained using RS option in the output command (OU RS). For example, for the last model, we get the following information on residuals:

```
        Fitted Covariance Matrix

              SCORE1     SCORE2       SEAT     SCHOOL       PLAY
            --------   --------   --------   --------   --------
   SCORE1       1.00
   SCORE2       0.28       1.00
     SEAT       0.16       0.10       1.00
   SCHOOL       0.10       0.06       0.52       1.00
     PLAY       0.11       0.07       0.59       0.36       1.00
```

Fitted Residuals

	SCORE1	SCORE2	SEAT	SCHOOL	PLAY
SCORE1	0.00				
SCORE2	0.00	0.00			
SEAT	0.00	0.00	0.00		
SCHOOL	-0.07	-0.02	0.00	0.00	
PLAY	0.04	-0.02	0.00	0.00	0.00

Summary Statistics for Fitted Residuals

```
Smallest Fitted Residual =   -0.07
  Median Fitted Residual =    0.00
 Largest Fitted Residual =    0.04
```

Stemleaf Plot

```
- 0|7
- 0|2200000000000
  0|4
```

Standardized Residuals

	SCORE1	SCORE2	SEAT	SCHOOL	PLAY
SCORE1	- -				
SCORE2	- -	- -			
SEAT	0.10	0.32	- -		
SCHOOL	-0.84	-0.22	0.43	- -	
PLAY	0.58	-0.21	-0.79	0.20	- -

Summary Statistics for Standardized Residuals

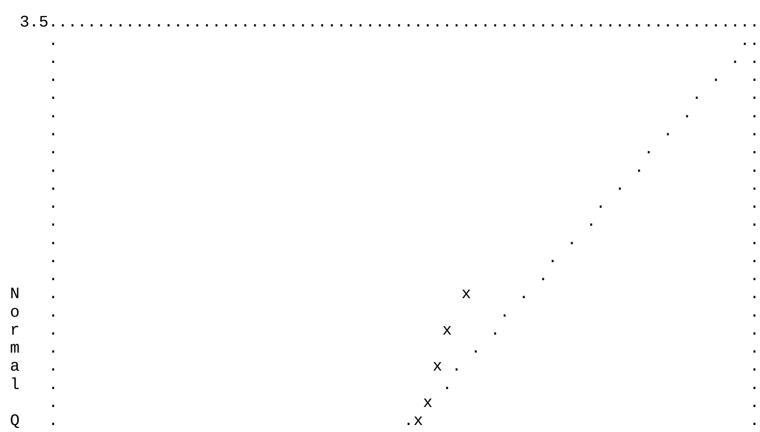
```
Smallest Standardized Residual =   -0.84
  Median Standardized Residual =    0.00
 Largest Standardized Residual =    0.58
```

Stemleaf Plot

```
- 0|88
- 0|22000000
  0|1234
  0|6
```

DA NI=5 NO=100 MA=KM

Qplot of Standardized Residuals



```
u   .                                  .                                    .
a   .                                x                                      .
n   .                               .                                       .
t   .                              . x                                      .
i   .                            .                                          .
l   .                          .x                                           .
e   .                         .                                             .
s   .                       .   x                                           .
    .                     .                                                 .
    .                    .                                                  .
    .                  .                                                    .
    .                .                                                      .
    .               .                                                       .
    .             .                                                         .
    .           .                                                           .
    .          .                                                            .
    .        .                                                              .
    .      .                                                                .
    .     .                                                                 .
    .   .                                                                   .
    . .                                                                     .
-3.5.........................................................................
  -3.5                                                                     3.5
                               Standardized Residuals
```

Using this information, we evaluate the normality of residuals distribution, and assess whether the distribution is symmetric. The Q-plot is especially useful: deviations from the straight line may indicate that (a) the model is in some ways misspecified, (b) the data are non-normally distributed, or (c) there are some nonlinear relationships. Note that along with regular residuals, we obtained standardized residuals. These are residuals divided by their standard error. These are, therefore, analogous to z-scores. Values >2.58 are considered large. These are indicative of possible fit problems in the model.

2. Modification indices. These are obtained using the output command as well (OU MI); they can be conceptualized as individual chi-square statistics for specific parameters (with 1 d.f., so the critical value is 3.841 for .05 level, 6.635 for .01 level, 10.828 for .001 level). They are estimated for all the parameters that were not freely estimated (i.e. were fixed), and represent the expected drop in the overall chi-square if the parameter were to be freely estimated. Therefore, all freely estimated parameters automatically have MI values equal to zero. Let's examine these for our last model:

```
 Modification Indices and Expected Change

         Modification Indices for LAMBDA-X
             ABILITY       PEER
            --------   --------
   SCORE1        - -        - -
   SCORE2        - -        - -
     SEAT       0.04        - -
   SCHOOL       0.63        - -
     PLAY       0.18        - -

         Expected Change for LAMBDA-X
             ABILITY       PEER
            --------   --------
   SCORE1        - -        - -
   SCORE2        - -        - -
     SEAT       0.04        - -
   SCHOOL      -0.15        - -
     PLAY       0.08        - -

 No Non-Zero Modification Indices for PHI

         Modification Indices for THETA-DELTA
```

	SCORE1	SCORE2	SEAT	SCHOOL	PLAY
SCORE1	- -				
SCORE2	- -	- -			
SEAT	0.00	0.09	- -		
SCHOOL	0.65	0.00	0.18	- -	
PLAY	0.43	0.14	0.63	0.04	- -

Expected Change for THETA-DELTA

	SCORE1	SCORE2	SEAT	SCHOOL	PLAY
SCORE1	- -				
SCORE2	- -	- -			
SEAT	0.00	0.03	- -		
SCHOOL	-0.07	0.00	0.16	- -	
PLAY	0.05	-0.03	-0.40	0.04	- -

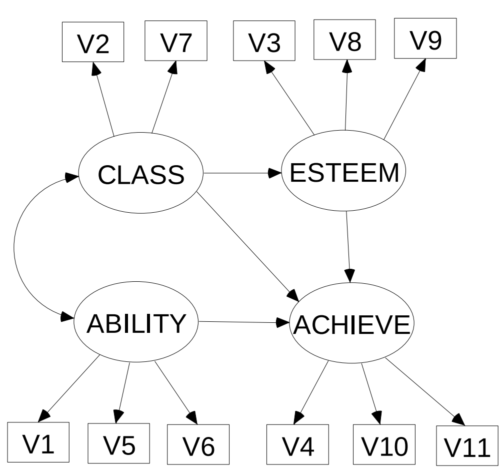
Maximum Modification Index is   0.65 for Element (4, 1) of THETA-DELTA

The modification indices reflect the predicted changes in chi-square, and the expected changes reflect the predicted changes in the coefficients. Based on these, we can respecify the model, but we should be careful because there is always a risk of overfitting the model to the data.

Estimating Combined Models

To combine the structural model and measurement model, we will typically first validate the measurement model, and then combine the two sets of equations.
Structural model:
$\eta = B\eta + \Gamma\xi + \zeta$
Measurement model:
$X = \Lambda_x \xi + \delta$
$Y = \Lambda_y \eta + \varepsilon$

Example with hypothetical data:
Social class (V2, V7), ability (V1, V5, V6), self-esteem (V3, V8, V9), school achievement (V4, V10, V11). Suppose we theorize that CLASS→ESTEEM, CLASS→ ACHIEVE, ABILITY→ACHIEVE, and ESTEEM→ACHIEVE. We also allow for a prior correlation of CLASS and ABILITY. Here's the model we envision:



Correlation Matrix (N=414):

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
V1	1.00										
V2	.345	1.00									
V3	.287	.377	1.00								
V4	.252	.579	.114	1.00							
V5	.637	.335	.253	.255	1.00						
V6	.768	.339	.302	.278	.600	1.00					
V7	.254	.703	.337	.591	.312	.313	1.00				
V8	.166	.429	.724	.143	.264	.216	.364	1.00			
V9	.104	.411	.506	.142	.199	.163	.551	.630	1.00		
V10	.247	.601	.202	.880	.288	.309	.621	.290	.199	1.00	
V11	.208	.526	.127	.827	.253	.231	.676	.155	.234	.808	1.00

Standard Deviations:

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
SD	13.9	13.4	10.0	15.9	9.4	4.9	5.9	11.1	5.9	11.8	7.9

Step 1. Assessing measurement model.

To assess the measurement model, it is best that we specify the entire model in terms of X and ξ.

$X = \Lambda_x \xi + \delta$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 \\ 0 & \lambda_{52} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \lambda_{73} & 0 \\ 0 & 0 & \lambda_{83} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \lambda_{10\,4} \\ 0 & 0 & 0 & \lambda_{11\,4} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \\ \delta_{11} \end{bmatrix}$$

Other matrices involved:

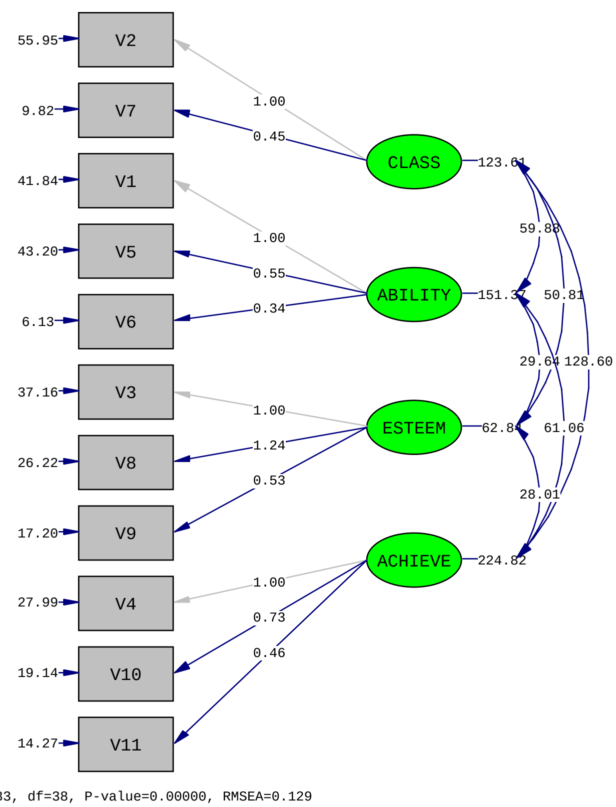
Θ_δ (11x11 matrix of variances and covariances of measurement errors δ) – measurement errors vary but they do not covary; therefore we use LISREL default (diagonal and free).

Φ (4x4 matrix of variances and covariances of latent variables ξ) – in a pure measurement model, we allow all latent variables to covary; therefore we use LISREL default (symmetric, free).

```
DA NI=11 NO=414 MA=CM
LA
V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11
KM SY
1.00
.345 1.00
.287 .377 1.00
```

```
 .252 .579 .114 1.00
 .637 .335 .253 .255 1.00
 .768 .339 .302 .278 .600 1.00
 .254 .703 .337 .591 .312 .313 1.00
 .166 .429 .724 .143 .264 .216 .364 1.00
 .104 .411 .506 .142 .199 .163 .551 .630 1.00
 .247 .601 .202 .880 .288 .309 .621 .290 .199 1.00
 .208 .526 .127 .827 .253 .231 .676 .155 .234 .808 1.00
 SD
 13.9 13.4 10.0 15.9 9.4 4.9 5.9 11.1 5.9 11.8 7.9
 SE
 2 7 1 5 6 3 8 9 4 10 11
 MO NX=11 NK=4 LX=FU,FI PH=SY,FR
 LK
 CLASS ABILITY ESTEEM ACHIEVE
 FR LX 2 1 LX 4 2 LX 5 2 LX 7 3 LX 8 3 LX 10 4 LX 11 4
 VA 1.0 LX 1 1 LX 3 2 LX 6 3 LX 9 4
PD
OU
```

Output:




```
 DA NI=11 NO=414 MA=CM
                          Number of Input Variables 11
                          Number of Y - Variables    0
                          Number of X - Variables   11
                          Number of ETA - Variables  0
                          Number of KSI - Variables  4
                          Number of Observations   414

 DA NI=11 NO=414 MA=CM

        Covariance Matrix
```

	V2	V7	V1	V5	V6	V3
V2	179.56					
V7	55.58	34.81				
V1	64.26	20.83	193.21			
V5	42.20	17.30	83.23	88.36		
V6	22.26	9.05	52.31	27.64	24.01	
V3	50.52	19.88	39.89	23.78	14.80	100.00
V8	63.81	23.84	25.61	27.55	11.75	80.36
V9	32.49	19.18	8.53	11.04	4.71	29.85
V4	123.36	55.44	55.69	38.11	21.66	18.13
V10	95.03	43.23	40.51	31.94	17.87	23.84
V11	55.68	31.51	22.84	18.79	8.94	10.03

Covariance Matrix

	V8	V9	V4	V10	V11
V8	123.21				
V9	41.26	34.81			
V4	25.24	13.32	252.81		
V10	37.98	13.85	165.11	139.24	
V11	13.59	10.91	103.88	75.32	62.41

DA NI=11 NO=414 MA=CM
Parameter Specifications

LAMBDA-X

	CLASS	ABILITY	ESTEEM	ACHIEVE
V2	0	0	0	0
V7	1	0	0	0
V1	0	0	0	0
V5	0	2	0	0
V6	0	3	0	0
V3	0	0	0	0
V8	0	0	4	0
V9	0	0	5	0
V4	0	0	0	0
V10	0	0	0	6
V11	0	0	0	7

PHI

	CLASS	ABILITY	ESTEEM	ACHIEVE
CLASS	8			
ABILITY	9	10		
ESTEEM	11	12	13	
ACHIEVE	14	15	16	17

THETA-DELTA

V2	V7	V1	V5	V6	V3
18	19	20	21	22	23

THETA-DELTA

V8	V9	V4	V10	V11
24	25	26	27	28

DA NI=11 NO=414 MA=CM
Number of Iterations = 7

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	CLASS	ABILITY	ESTEEM	ACHIEVE
V2	1.00	- -	- -	- -
V7	0.45 (0.02) 18.81	- -	- -	- -
V1	- -	1.00	- -	- -
V5	- -	0.55 (0.03) 16.01	- -	- -
V6	- -	0.34 (0.02) 19.29	- -	- -
V3	- -	- -	1.00	- -
V8	- -	- -	1.24 (0.07) 17.03	- -
V9	- -	- -	0.53 (0.04) 14.59	- -
V4	- -	- -	- -	1.00
V10	- -	- -	- -	0.73 (0.02) 34.47
V11	- -	- -	- -	0.46 (0.02) 29.45

PHI

	CLASS	ABILITY	ESTEEM	ACHIEVE
CLASS	123.61 (12.63) 9.79			
ABILITY	59.88 (8.71) 6.88	151.37 (14.30) 10.58		
ESTEEM	50.81 (6.23) 8.15	29.64 (5.85) 5.07	62.84 (6.99) 9.00	
ACHIEVE	128.60 (11.95) 10.76	61.06 (10.52) 5.81	28.01 (6.72) 4.17	224.82 (17.77) 12.65

THETA-DELTA

V2	V7	V1	V5	V6	V3
55.95	9.82	41.84	43.20	6.13	37.16
(5.80)	(1.11)	(6.39)	(3.50)	(0.80)	(3.71)
9.65	8.88	6.55	12.33	7.69	10.03

THETA-DELTA

V8	V9	V4	V10	V11
26.22	17.20	27.99	19.14	14.27
(4.37)	(1.44)	(3.72)	(2.16)	(1.23)
6.01	11.97	7.53	8.84	11.61

Squared Multiple Correlations for X - Variables

V2	V7	V1	V5	V6	V3
0.69	0.72	0.78	0.51	0.74	0.63

Squared Multiple Correlations for X - Variables

V8	V9	V4	V10	V11
0.79	0.51	0.89	0.86	0.77

Goodness of Fit Statistics

Degrees of Freedom = 38
Minimum Fit Function Chi-Square = 293.34 (P = 0.0)
Normal Theory Weighted Least Squares Chi-Square = 300.83 (P = 0.0)
Estimated Non-centrality Parameter (NCP) = 262.83
90 Percent Confidence Interval for NCP = (211.30 ; 321.85)

Minimum Fit Function Value = 0.71
Population Discrepancy Function Value (F0) = 0.64
90 Percent Confidence Interval for F0 = (0.51 ; 0.78)
Root Mean Square Error of Approximation (RMSEA) = 0.13
90 Percent Confidence Interval for RMSEA = (0.12 ; 0.14)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.00

Expected Cross-Validation Index (ECVI) = 0.86
90 Percent Confidence Interval for ECVI = (0.74 ; 1.01)
ECVI for Saturated Model = 0.32
ECVI for Independence Model = 10.38

Chi-Square for Independence Model with 55 Degrees of Freedom = 4266.81
Independence AIC = 4288.81
Model AIC = 356.83
Saturated AIC = 132.00
Independence CAIC = 4344.10
Model CAIC = 497.56
Saturated CAIC = 463.71

Normed Fit Index (NFI) = 0.93
Non-Normed Fit Index (NNFI) = 0.91
Parsimony Normed Fit Index (PNFI) = 0.64
Comparative Fit Index (CFI) = 0.94
Incremental Fit Index (IFI) = 0.94
Relative Fit Index (RFI) = 0.90

Critical N (CN) = 87.11

Root Mean Square Residual (RMR) = 4.53
Standardized RMR = 0.051
Goodness of Fit Index (GFI) = 0.88
Adjusted Goodness of Fit Index (AGFI) = 0.80
Parsimony Goodness of Fit Index (PGFI) = 0.51

Time used: 0.047 Seconds

DA NI=11 NO=414 MA=CM

Fitted Covariance Matrix

	V2	V7	V1	V5	V6	V3
V2	179.56					
V7	55.58	34.81				
V1	59.88	26.93	193.21			
V5	32.71	14.71	82.68	88.36		
V6	20.58	9.25	52.02	28.41	24.01	
V3	50.81	22.85	29.64	16.19	10.19	100.00
V8	63.12	28.38	36.82	20.11	12.65	78.07
V9	26.90	12.09	15.69	8.57	5.39	33.27
V4	128.60	57.82	61.06	33.35	20.99	28.01
V10	93.99	42.26	44.63	24.38	15.34	20.47
V11	59.51	26.76	28.26	15.43	9.71	12.96

Fitted Covariance Matrix

	V8	V9	V4	V10	V11
V8	123.21				
V9	41.33	34.81			
V4	34.80	14.83	252.81		
V10	25.43	10.84	164.32	139.24	
V11	16.10	6.86	104.04	76.04	62.41

Fitted Residuals

	V2	V7	V1	V5	V6	V3
V2	0.00					
V7	0.00	0.00				
V1	4.37	-6.10	0.00			
V5	9.49	2.60	0.55	0.00		
V6	1.68	-0.20	0.29	-0.78	0.00	
V3	-0.29	-2.96	10.25	7.59	4.61	0.00
V8	0.69	-4.54	-11.21	7.43	-0.91	2.30
V9	5.59	7.09	-7.16	2.47	-0.68	-3.41
V4	-5.24	-2.38	-5.37	4.76	0.67	-9.88
V10	1.04	0.97	-4.12	7.57	2.53	3.36
V11	-3.83	4.75	-5.42	3.35	-0.77	-2.93

Fitted Residuals

	V8	V9	V4	V10	V11
V8	0.00				
V9	-0.07	0.00			
V4	-9.56	-1.51	0.00		
V10	12.55	3.02	0.78	0.00	
V11	-2.51	4.04	-0.16	-0.72	0.00

Summary Statistics for Fitted Residuals

Smallest Fitted Residual = -11.21

```
  Median Fitted Residual =    0.00
 Largest Fitted Residual =   12.55
```

Stemleaf Plot

```
-10|2
- 8|96
- 6|21
- 4|44251
- 2|840954
- 0|598877322100000000000
  0|36778007
  2|3556044
  4|046886
  6|1466
  8|5
 10|3
 12|6
```

```
        Standardized Residuals
                 V2         V7         V1         V5         V6         V3
              --------   --------   --------   --------   --------   --------
      V2        - -
      V7        - -        - -
      V1       1.07      -3.61        - -
      V5       2.34       1.49       0.89        - -
      V6       1.08      -0.32       2.76      -2.91        - -
      V3      -0.09      -2.12       2.68       2.25       3.26        - -
      V8       0.25      -3.99      -3.61       2.21      -0.75       6.40
      V9       2.43       7.18      -2.70       1.15      -0.71      -5.20
      V4      -1.96      -2.17      -1.31       1.02       0.42      -2.33
     V10       0.47       1.07      -1.26       2.14       2.00       1.03
     V11      -2.02       5.94      -2.03       1.33      -0.77      -1.22
```

```
        Standardized Residuals
                 V8         V9         V4        V10        V11
              --------   --------   --------   --------   --------
      V8        - -
      V9      -0.20        - -
      V4      -3.00      -0.50        - -
     V10       4.89       1.33       3.03        - -
     V11      -1.18       2.51      -0.45      -2.14        - -
```

Summary Statistics for Standardized Residuals

```
Smallest Standardized Residual =   -5.20
  Median Standardized Residual =    0.00
 Largest Standardized Residual =    7.18
```

Stemleaf Plot

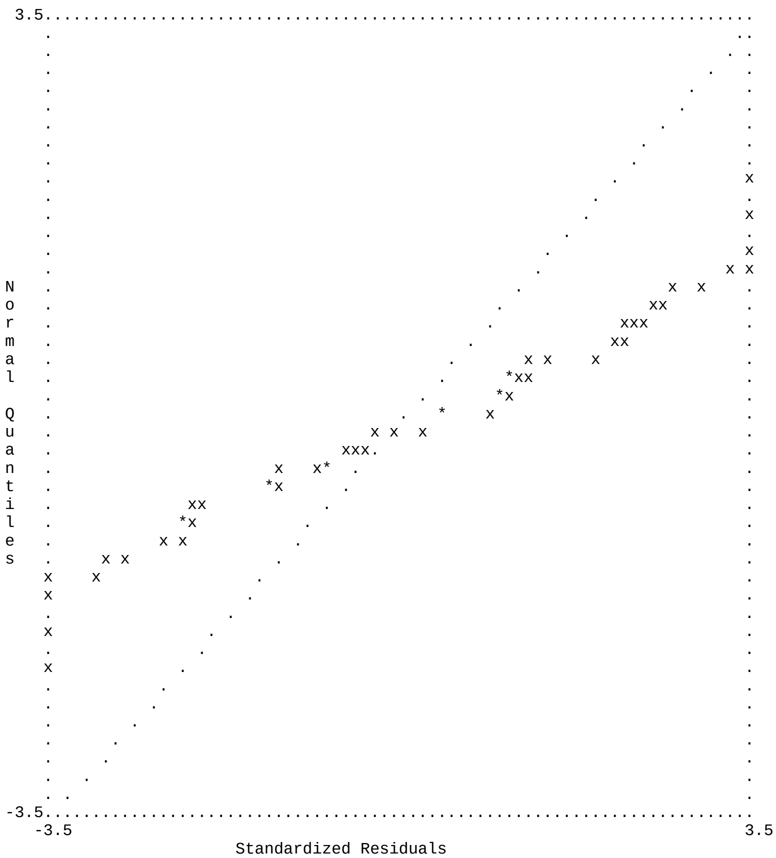
```
- 4|20
- 2|660973211000
- 0|3322887543210000000000
  0|2459001112335
  2|01233457803
  4|99
  6|42
```

Largest Negative Standardized Residuals

```
Residual for       V1 and       V7  -3.61
Residual for       V6 and       V5  -2.91
```

```
Residual for       V8 and       V7  -3.99
Residual for       V8 and       V1  -3.61
Residual for       V9 and       V1  -2.70
Residual for       V9 and       V3  -5.20
Residual for       V4 and       V8  -3.00
Largest Positive Standardized Residuals
Residual for       V6 and       V1   2.76
Residual for       V3 and       V1   2.68
Residual for       V3 and       V6   3.26
Residual for       V8 and       V3   6.40
Residual for       V9 and       V7   7.18
Residual for      V10 and       V8   4.89
Residual for      V10 and       V4   3.03
Residual for      V11 and       V7   5.94
```

DA NI=11 NO=414 MA=CM
Qplot of Standardized Residuals



Modification Indices and Expected Change

Modification Indices for LAMBDA-X

	CLASS	ABILITY	ESTEEM	ACHIEVE
V2	- -	4.29	1.35	2.97
V7	- -	4.29	1.35	2.97
V1	7.21	- -	6.79	4.42
V5	6.58	- -	6.86	2.95
V6	0.67	- -	0.58	0.80
V3	1.43	13.65	- -	1.18
V8	5.46	5.39	- -	0.00

V9	18.74	1.76	- -	1.55
V4	18.46	0.19	22.95	- -
V10	7.70	2.06	26.94	- -
V11	3.71	1.40	0.07	- -

Expected Change for LAMBDA-X

	CLASS	ABILITY	ESTEEM	ACHIEVE
V2	- -	0.11	0.11	-0.13
V7	- -	-0.05	-0.05	0.06
V1	-0.14	- -	-0.17	-0.07
V5	0.10	- -	0.13	0.04
V6	0.01	- -	0.02	0.01
V3	-0.05	0.12	- -	-0.03
V8	-0.12	-0.09	- -	0.00
V9	0.12	-0.03	- -	0.02
V4	-0.28	-0.01	-0.24	- -
V10	0.13	0.04	0.20	- -
V11	0.07	-0.02	-0.01	- -

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	V2	V7	V1	V5	V6	V3
V2	- -					
V7	- -	- -				
V1	7.20	6.58	- -			
V5	0.06	0.66	0.80	- -		
V6	1.92	0.82	7.60	8.45	- -	
V3	0.04	3.42	9.99	4.30	1.40	- -
V8	1.55	30.95	5.86	6.39	0.14	40.91
V9	2.05	87.88	6.19	0.79	0.00	27.03
V4	3.79	10.01	3.55	1.99	0.12	0.25
V10	0.35	6.19	5.95	0.84	4.91	0.03
V11	15.78	59.72	0.18	1.51	2.63	0.81

Modification Indices for THETA-DELTA

	V8	V9	V4	V10	V11
V8	- -				
V9	0.04	- -			
V4	4.30	2.86	- -		
V10	48.26	15.90	9.20	- -	
V11	11.94	17.00	0.20	4.60	- -

Expected Change for THETA-DELTA

	V2	V7	V1	V5	V6	V3
V2	- -					
V7	- -	- -				
V1	9.90	-4.10	- -			
V5	-0.71	1.06	6.36	- -		
V6	-1.83	0.52	16.42	-6.97	- -	
V3	0.58	-2.45	9.06	-4.92	1.22	- -
V8	4.08	-7.94	-7.16	6.14	-0.39	42.00
V9	-2.78	7.86	-4.60	1.36	-0.03	-12.26
V4	6.06	-4.27	5.40	-3.34	-0.35	1.17
V10	1.40	-2.55	-5.38	1.67	1.75	-0.30

```
V11      -7.18       6.03      -0.73       1.76      -1.00      -1.27

   Expected Change for THETA-DELTA

               V8         V9         V4        V10        V11
         --------   --------   --------   --------   --------
    V8        - -
    V9      -0.62        - -
    V4      -4.98      -2.55        - -
   V10      12.83      -4.63      15.22        - -
   V11      -4.98       3.74      -1.23      -4.24        - -

Maximum Modification Index is   87.88 for Element ( 8, 2) of THETA-DELTA
```

Step 2. SEM with Latent Variables.

Once we are satisfied with our measurement model, the next step is to combine the measurement model and the structural model. Note that in terms of model identification, we can either count total number of parameters estimated in all matrices, or we can separately assess the measurement model and the structural model – if each individually is identified, then the entire model is identified. I would advise to use this second method because typically you cannot obtain degrees of freedom for structural model from the measurement model and vice versa.

Let's estimate a combined SEM model for our example.

Measurement model:

$X = \Lambda_x \xi + \delta$

$Y = \Lambda_y \eta + \varepsilon$

Structural model:

$\eta = B\eta + \Gamma\xi + \zeta$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & 1 \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \beta_{21} & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \gamma_{12} & 0 \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

Other matrices involved:

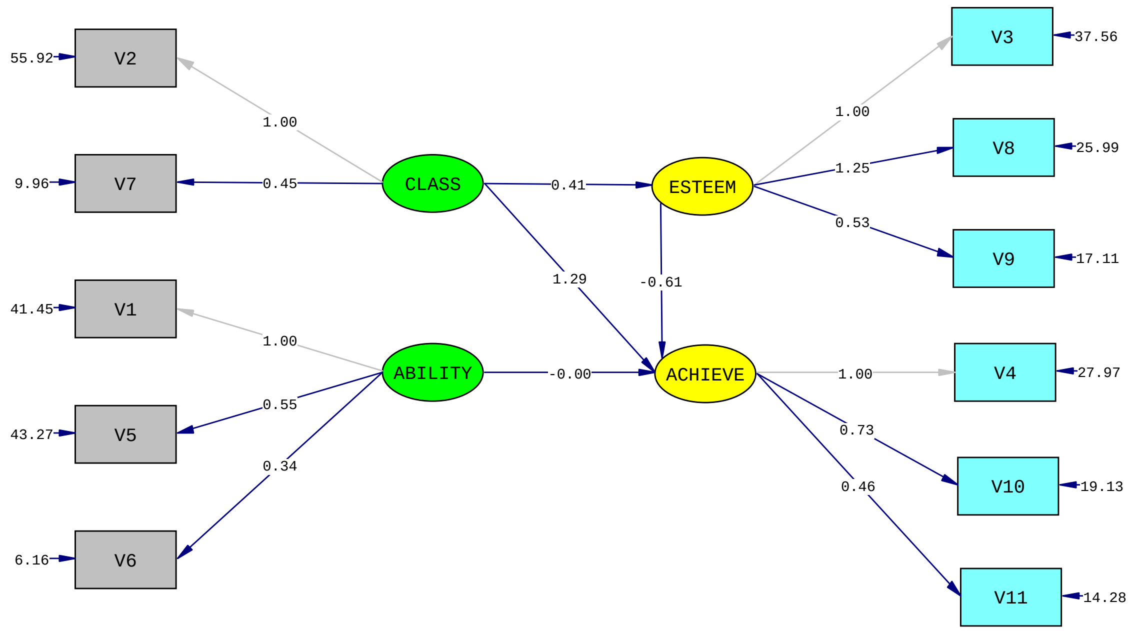
Θ_δ (5x5 matrix of variances and covariances of measurement errors δ) – measurement errors vary but they do not covary; we use LISREL default (diagonal and free).

Θ_ε (6x6 matrix of variances and covariances of measurement errors ε) – measurement errors vary but they do not covary; we use LISREL default (diagonal and free).

Φ (2x2 matrix of variances and covariances of exogenous variables ξ) – the two exogenous variables are correlated, so all three elements of this matrix are estimated – use LISREL default (symmetric, free).

Ψ (2x2 matrix of variances and covariances of disturbance terms ζ) – since we don't allow disturbance terms to covary, we use LISREL default (diagonal, free).

```
 DA NI=11 NO=414 MA=CM
 LA
 V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11
 KM SY
 1.00
 .345 1.00
 .287 .377 1.00
 .252 .579 .114 1.00
 .637 .335 .253 .255 1.00
 .768 .339 .302 .278 .600 1.00
 .254 .703 .337 .591 .312 .313 1.00
 .166 .429 .724 .143 .264 .216 .364 1.00
 .104 .411 .506 .142 .199 .163 .551 .630 1.00
 .247 .601 .202 .880 .288 .309 .621 .290 .199 1.00
 .208 .526 .127 .827 .253 .231 .676 .155 .234 .808 1.00
 SD
 13.9 13.4 10.0 15.9 9.4 4.9 5.9 11.1 5.9 11.8 7.9
 SE
 3 8 9 4 10 11  2 7 1 5 6
 MO NX=5 NK=2 NY=6 NE=2 LX=FU, FI LY=FU, FI PH=SY,FR PS=DI,FR TD=DI,FR TE=DI,FR
BE=FU,FI GA=FU,FR
 LK
 CLASS ABILITY
 LE
 ESTEEM ACHIEVE
 FR LX 2 1 LX 4 2 LX 5 2 LY 2 1 LY 3 1 LY 5 2 LY 6 2 BE 2 1
 FI GA 1 2
 VA 1.0 LX 1 1 LX 3 2 LY 1 1 LY 4 2
 PD
 OU RS MI
```



DA NI=11 NO=414 MA=CM

```
Number of Input Variables 11
Number of Y - Variables    6
Number of X - Variables    5
Number of ETA - Variables  2
Number of KSI - Variables  2
Number of Observations   414
```

DA NI=11 NO=414 MA=CM

Covariance Matrix

	V3	V8	V9	V4	V10	V11
V3	100.00					
V8	80.36	123.21				
V9	29.85	41.26	34.81			
V4	18.13	25.24	13.32	252.81		
V10	23.84	37.98	13.85	165.11	139.24	
V11	10.03	13.59	10.91	103.88	75.32	62.41
V2	50.52	63.81	32.49	123.36	95.03	55.68
V7	19.88	23.84	19.18	55.44	43.23	31.51
V1	39.89	25.61	8.53	55.69	40.51	22.84
V5	23.78	27.55	11.04	38.11	31.94	18.79
V6	14.80	11.75	4.71	21.66	17.87	8.94

Covariance Matrix

	V2	V7	V1	V5	V6
V2	179.56				
V7	55.58	34.81			
V1	64.26	20.83	193.21		
V5	42.20	17.30	83.23	88.36	
V6	22.26	9.05	52.31	27.64	24.01

DA NI=11 NO=414 MA=CM

Parameter Specifications

LAMBDA-Y

	ESTEEM	ACHIEVE
V3	0	0
V8	1	0
V9	2	0
V4	0	0
V10	0	3
V11	0	4

LAMBDA-X

	CLASS	ABILITY
V2	0	0
V7	5	0
V1	0	0
V5	0	6
V6	0	7

BETA

	ESTEEM	ACHIEVE
ESTEEM	0	0
ACHIEVE	8	0

GAMMA

	CLASS	ABILITY
ESTEEM	9	0
ACHIEVE	10	11

PHI

	CLASS	ABILITY
CLASS	12	
ABILITY	13	14

PSI

ESTEEM	ACHIEVE
15	16

THETA-EPS

V3	V8	V9	V4	V10	V11
17	18	19	20	21	22

THETA-DELTA

V2	V7	V1	V5	V6
23	24	25	26	27

```
DA NI=11 NO=414 MA=CM

Number of Iterations =  7

LISREL Estimates (Maximum Likelihood)

        LAMBDA-Y

             ESTEEM    ACHIEVE
           --------   --------
      V3       1.00        - -

      V8       1.25        - -
             (0.07)
              16.97

      V9       0.53        - -
             (0.04)
              14.60

      V4        - -       1.00

     V10        - -       0.73
                        (0.02)
                         34.47

     V11        - -       0.46
                        (0.02)
                         29.44


        LAMBDA-X

              CLASS    ABILITY
           --------   --------
      V2       1.00        - -

      V7       0.45        - -
             (0.02)
              18.82

      V1        - -       1.00

      V5        - -       0.55
                        (0.03)
                         16.00

      V6        - -       0.34
                        (0.02)
                         19.27


        BETA

             ESTEEM    ACHIEVE
           --------   --------
  ESTEEM        - -        - -

 ACHIEVE      -0.61        - -
             (0.11)
              -5.52
```

GAMMA

	CLASS	ABILITY
ESTEEM	0.41	- -
	(0.04)	
	9.83	
ACHIEVE	1.29	0.00
	(0.10)	(0.06)
	12.94	-0.03

Covariance Matrix of ETA and KSI

	ESTEEM	ACHIEVE	CLASS	ABILITY
ESTEEM	62.44			
ACHIEVE	27.94	224.84		
CLASS	51.13	128.58	123.64	
ABILITY	25.21	63.19	60.97	151.76

PHI

	CLASS	ABILITY
CLASS	123.64	
	(12.62)	
	9.80	
ABILITY	60.97	151.76
	(8.71)	(14.32)
	7.00	10.60

PSI

Note: This matrix is diagonal.

ESTEEM	ACHIEVE
41.29	75.70
(4.98)	(10.22)
8.29	7.41

Squared Multiple Correlations for Structural Equations

ESTEEM	ACHIEVE
0.34	0.66

Squared Multiple Correlations for Reduced Form

ESTEEM	ACHIEVE
0.34	0.59

Reduced Form

CLASS	ABILITY

ESTEEM	0.41 (0.04) 9.83	- -
ACHIEVE	1.04 (0.08) 13.72	0.00 (0.06) -0.03

THETA-EPS

V3	V8	V9	V4	V10	V11
37.56 (3.72) 10.10	25.99 (4.37) 5.95	17.11 (1.43) 11.94	27.97 (3.71) 7.53	19.13 (2.16) 8.84	14.28 (1.23) 11.61

Squared Multiple Correlations for Y - Variables

V3	V8	V9	V4	V10	V11
0.62	0.79	0.51	0.89	0.86	0.77

THETA-DELTA

V2	V7	V1	V5	V6
55.92 (5.78) 9.68	9.96 (1.10) 9.04	41.45 (6.39) 6.48	43.27 (3.51) 12.33	6.16 (0.80) 7.72

Squared Multiple Correlations for X - Variables

V2	V7	V1	V5	V6
0.69	0.71	0.79	0.51	0.74

Goodness of Fit Statistics

Degrees of Freedom = 39
Minimum Fit Function Chi-Square = 294.56 (P = 0.0)
Normal Theory Weighted Least Squares Chi-Square = 303.30 (P = 0.0)
Estimated Non-centrality Parameter (NCP) = 264.30
90 Percent Confidence Interval for NCP = (212.57 ; 323.51)

Minimum Fit Function Value = 0.71
Population Discrepancy Function Value (F0) = 0.64
90 Percent Confidence Interval for F0 = (0.51 ; 0.78)
Root Mean Square Error of Approximation (RMSEA) = 0.13
90 Percent Confidence Interval for RMSEA = (0.11 ; 0.14)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.00

Expected Cross-Validation Index (ECVI) = 0.87
90 Percent Confidence Interval for ECVI = (0.74 ; 1.01)
ECVI for Saturated Model = 0.32
ECVI for Independence Model = 10.38

Chi-Square for Independence Model with 55 Degrees of Freedom = 4266.81
Independence AIC = 4288.81
Model AIC = 357.30

```
                    Saturated AIC = 132.00
                 Independence CAIC = 4344.10
                     Model CAIC = 493.00
                   Saturated CAIC = 463.71

                Normed Fit Index (NFI) = 0.93
              Non-Normed Fit Index (NNFI) = 0.91
          Parsimony Normed Fit Index (PNFI) = 0.66
              Comparative Fit Index (CFI) = 0.94
              Incremental Fit Index (IFI) = 0.94
               Relative Fit Index (RFI) = 0.90

                   Critical N (CN) = 88.53


         Root Mean Square Residual (RMR) = 4.73
                  Standardized RMR = 0.053
             Goodness of Fit Index (GFI) = 0.88
        Adjusted Goodness of Fit Index (AGFI) = 0.80
       Parsimony Goodness of Fit Index (PGFI) = 0.52

DA NI=11 NO=414 MA=CM
```

Fitted Covariance Matrix

	V3	V8	V9	V4	V10	V11
V3	100.00					
V8	77.91	123.21				
V9	33.24	41.48	34.81			
V4	27.94	34.87	14.88	252.80		
V10	20.42	25.48	10.87	164.33	139.24	
V11	12.93	16.13	6.88	104.03	76.03	62.41
V2	51.13	63.80	27.22	128.58	93.98	59.49
V7	22.92	28.60	12.20	57.65	42.13	26.67
V1	25.21	31.46	13.42	63.19	46.19	29.24
V5	13.74	17.15	7.32	34.44	25.17	15.94
V6	8.65	10.79	4.60	21.67	15.84	10.03

Fitted Covariance Matrix

	V2	V7	V1	V5	V6
V2	179.56				
V7	55.43	34.81			
V1	60.97	27.33	193.21		
V5	33.23	14.90	82.72	88.36	
V6	20.91	9.37	52.04	28.37	24.01

Fitted Residuals

	V3	V8	V9	V4	V10	V11
V3	0.00					
V8	2.45	0.00				
V9	-3.39	-0.22	0.00			
V4	-9.82	-9.63	-1.55	0.01		
V10	3.41	12.50	2.98	0.78	0.00	
V11	-2.90	-2.54	4.02	-0.15	-0.71	0.00
V2	-0.61	0.01	5.27	-5.22	1.05	-3.81
V7	-3.04	-4.77	6.98	-2.20	1.10	4.84
V1	14.68	-5.85	-4.89	-7.50	-5.67	-6.40
V5	10.04	10.40	3.72	3.67	6.77	2.85

V6	6.15	0.96	0.11	-0.01	2.03	-1.08

Fitted Residuals

	V2	V7	V1	V5	V6
V2	0.00				
V7	0.15	0.00			
V1	3.29	-6.50	0.00		
V5	8.96	2.40	0.51	0.00	
V6	1.35	-0.32	0.27	-0.73	0.00

Summary Statistics for Fitted Residuals

```
Smallest Fitted Residual =   -9.82
  Median Fitted Residual =    0.00
 Largest Fitted Residual =   14.68
```

Stemleaf Plot

```
- 8|86
- 6|554
- 4|87298
- 2|840952
- 0|6177632100000000000000
  0|113580114
  2|045903477
  4|083
  6|280
  8|0
 10|04
 12|5
 14|7
```

Standardized Residuals

	V3	V8	V9	V4	V10	V11
V3	- -					
V8	6.81	- -				
V9	-5.15	-0.64	- -			
V4	-2.30	-3.04	-0.52	1.10		
V10	1.04	4.89	1.32	3.02	1.10	
V11	-1.20	-1.20	2.50	-0.43	-2.12	1.10
V2	-0.18	0.00	2.30	-1.93	0.48	-2.01
V7	-2.15	-4.12	7.08	-1.97	1.20	5.99
V1	2.65	-1.01	-1.44	-1.68	-1.61	-2.29
V5	2.49	2.41	1.53	0.77	1.88	1.12
V6	3.12	0.46	0.09	-0.01	1.50	-1.04

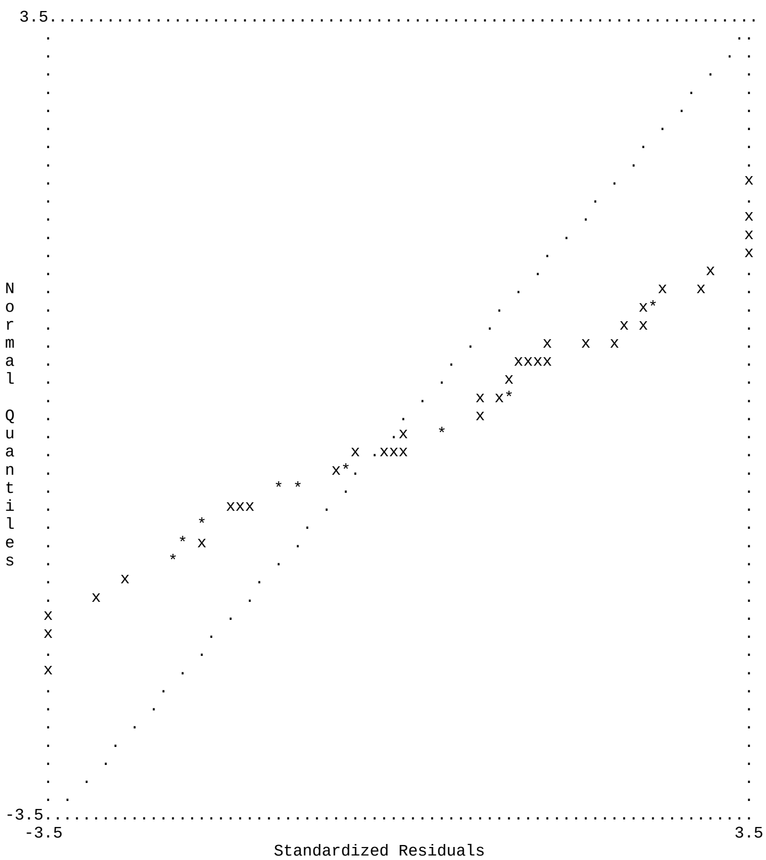
Standardized Residuals

	V2	V7	V1	V5	V6
V2	- -				
V7	1.09	- -			
V1	0.79	-3.77	- -		
V5	2.20	1.37	0.85	- -	
V6	0.86	-0.49	2.59	-2.73	- -

Summary Statistics for Standardized Residuals

Smallest Standardized Residual =   -5.15

```
  Median Standardized Residual =    0.00
 Largest Standardized Residual =    7.08

Stemleaf Plot

- 4|11
- 2|807332100
- 0|976422006554200000000000
  0|1558889011111234559
  2|234556701
  4|9
  6|081
Largest Negative Standardized Residuals
Residual for       V9 and       V3  -5.15
Residual for       V4 and       V8  -3.04
Residual for       V7 and       V8  -4.12
Residual for       V1 and       V7  -3.77
Residual for       V6 and       V5  -2.73
Largest Positive Standardized Residuals
Residual for       V8 and       V3   6.81
Residual for      V10 and       V8   4.89
Residual for      V10 and       V4   3.02
Residual for       V7 and       V9   7.08
Residual for       V7 and      V11   5.99
Residual for       V1 and       V3   2.65
Residual for       V6 and       V3   3.12
Residual for       V6 and       V1   2.59

DA NI=11 NO=414 MA=CM

Qplot of Standardized Residuals
```



```
DA NI=11 NO=414 MA=CM

Modification Indices and Expected Change

        Modification Indices for LAMBDA-Y

             ESTEEM    ACHIEVE
           --------   --------
      V3       - -       1.11
      V8       - -       0.00
      V9       - -       1.51
      V4     23.13        - -
     V10     26.98        - -
     V11      0.06        - -

        Expected Change for LAMBDA-Y

             ESTEEM    ACHIEVE
           --------   --------
      V3       - -      -0.03
      V8       - -       0.00
      V9       - -       0.02
      V4     -0.24        - -
     V10      0.20        - -
     V11     -0.01        - -

        Modification Indices for LAMBDA-X

              CLASS    ABILITY
           --------   --------
      V2       - -       2.83
      V7       - -       5.16
      V1      7.47        - -
      V5      6.71        - -
      V6      0.72        - -

        Expected Change for LAMBDA-X

              CLASS    ABILITY
           --------   --------
      V2       - -       0.08
      V7       - -      -0.05
      V1     -0.14        - -
      V5      0.10        - -
      V6      0.02        - -

        Modification Indices for BETA

             ESTEEM    ACHIEVE
           --------   --------
  ESTEEM       - -       1.19
 ACHIEVE       - -        - -

        Expected Change for BETA

             ESTEEM    ACHIEVE
           --------   --------
  ESTEEM       - -     -23.02
 ACHIEVE       - -        - -

        Modification Indices for GAMMA

              CLASS    ABILITY
```

ESTEEM	- -	1.19
ACHIEVE	- -	- -

Expected Change for GAMMA

	CLASS	ABILITY
ESTEEM	- -	0.04
ACHIEVE	- -	- -

No Non-Zero Modification Indices for PHI

No Non-Zero Modification Indices for PSI

Modification Indices for THETA-EPS

	V3	V8	V9	V4	V10	V11
V3	- -					
V8	46.35	- -				
V9	26.49	0.40	- -			
V4	0.24	4.20	2.76	- -		
V10	0.02	48.72	16.05	8.99	- -	
V11	0.85	12.42	17.01	0.19	4.52	- -

Expected Change for THETA-EPS

	V3	V8	V9	V4	V10	V11
V3	- -					
V8	44.70	- -				
V9	-12.17	-1.98	- -			
V4	1.15	-4.92	-2.50	- -		
V10	-0.28	12.88	-4.64	15.06	- -	
V11	-1.30	-5.07	3.74	-1.20	-4.20	- -

Modification Indices for THETA-DELTA-EPS

	V3	V8	V9	V4	V10	V11
V2	0.01	1.30	1.81	3.76	0.25	15.18
V7	3.74	32.76	88.24	9.62	6.31	59.78
V1	10.56	4.66	6.10	3.32	5.55	0.21
V5	4.12	6.61	0.75	2.07	0.94	1.47
V6	1.57	0.04	0.00	0.16	5.30	2.69

Expected Change for THETA-DELTA-EPS

	V3	V8	V9	V4	V10	V11
V2	0.33	3.74	-2.61	6.03	1.18	-7.04
V7	-2.58	-8.18	7.88	-4.19	-2.58	6.04
V1	9.29	-6.30	-4.54	5.21	-5.19	-0.78
V5	-4.82	6.23	1.33	-3.41	1.77	1.74
V6	1.29	-0.21	-0.03	-0.41	1.82	-1.02

Modification Indices for THETA-DELTA

	V2	V7	V1	V5	V6
V2	- -				
V7	1.19	- -			
V1	6.72	7.05	- -		
V5	0.08	0.60	0.72	- -	
V6	2.09	0.68	6.71	7.45	- -

Expected Change for THETA-DELTA

	V2	V7	V1	V5	V6

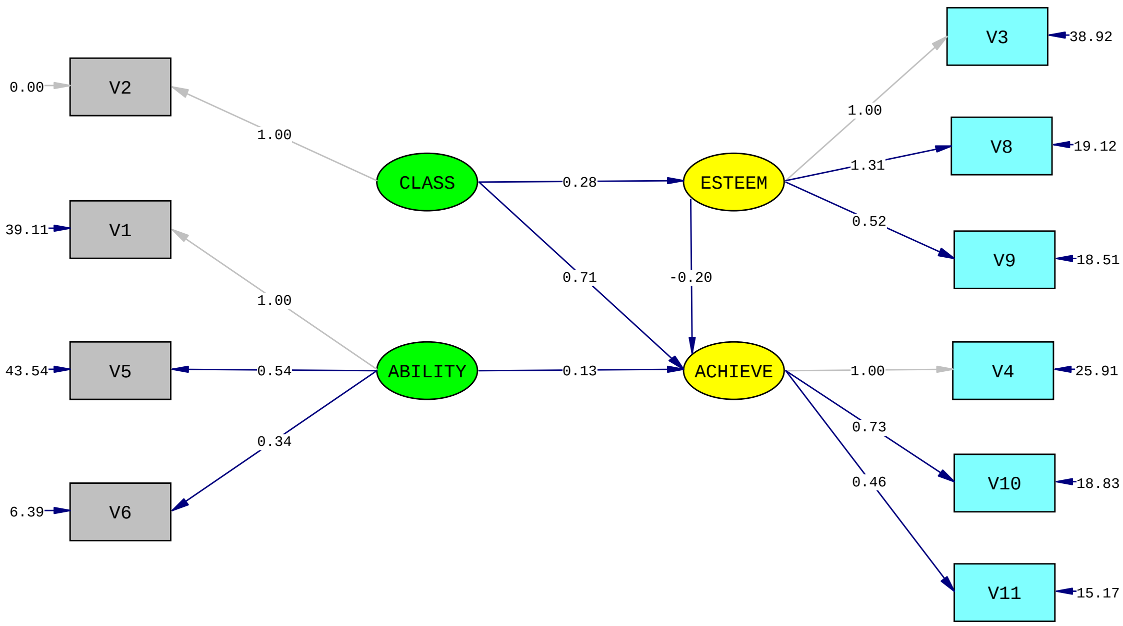
```
        --------   --------   --------   --------   --------
      V2        - -
      V7      10.75        - -
      V1       9.54      -4.23        - -
      V5      -0.83       1.01       6.08        - -
      V6      -1.91       0.47      15.61      -6.58        - -

 Maximum Modification Index is   88.24 for Element ( 2, 3) of THETA DELTA-EPSILON

                          Time used:    0.047 Seconds
```

Variables with single indicators

So far in our model all variables were latent variables with multiple indicators. But sometimes we have to use single-indicator variables in our analysis – e.g. if we have to rely on income only to measure class (or to include gender, or age into the model). An easy way to do that is to specify a single-indicator latent variable with the corresponding λ fixed to 1, and the corresponding measurement error δ fixed to 0 (i.e., we assume no measurement error). E.g. if only V2 is available as a measure of class, we specify:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$$

```
 DA NI=11 NO=414 MA=CM
 LA
 V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11
 KM SY
 1.00
 .345 1.00
 .287 .377 1.00
 .252 .579 .114 1.00
 .637 .335 .253 .255 1.00
 .768 .339 .302 .278 .600 1.00
 .254 .703 .337 .591 .312 .313 1.00
 .166 .429 .724 .143 .264 .216 .364 1.00
 .104 .411 .506 .142 .199 .163 .551 .630 1.00
 .247 .601 .202 .880 .288 .309 .621 .290 .199 1.00
 .208 .526 .127 .827 .253 .231 .676 .155 .234 .808 1.00
 SD
 13.9 13.4 10.0 15.9 9.4 4.9 5.9 11.1 5.9 11.8 7.9
 SE
 3 8 9 4 10 11  2 1 5 6 /
 MO NX=4 NK=2 NY=6 NE=2 LX=FU,FI LY=FU,FI PH=SY,FR PS=DI,FR TD=DI,FR TE=DI,FR
BE=FU,FI GA=FU,FR
 LK
 CLASS ABILITY
 LE
 ESTEEM ACHIEVE
 FR LX 3 2 LX 4 2 LY 2 1 LY 3 1 LY 5 2 LY 6 2 BE 2 1
 FI GA 1 2 TD 1 1
 VA 1.0 LX 1 1 LX 2 2 LY 1 1 LY 4 2
 PD
 OU
```



Obtaining Covariance Matrices and Using Raw Data

Let's use nys2.sav data available in HLM folder to obtain the covariance matrix. First, we import nys2.sav into LISREL, creating nys2.PSF. Upon importing the data, we can obtain the covariance matrix – it will be in the file named nys.cov.

```
!Prelis syntax
SY='C:\nys2.PSF'
OU MA=CM SM=nys.cov
```

Note that here we obtained a covariance matrix using CM option; that's the standard when all variables are continuous. But if we are dealing with ordinal, categorical, or mixed data, we need two matrices, polychoric correlation matrix (PM) and asymptotic covariance matrix (AC), e.g.:

```
!Prelis syntax
SY='C:\nys2.PSF'
OU MA=PM SM=nys.pcm AC=asymptnys.acm
```

In the analysis syntax, in the DA statement, make sure to specify MA=PM and specify the following input matrices:

```
PM=nys.pcm
AC=asymptnys.acm
```

Further, our dataset contains missing values – we have to decide how to deal with those. Listwise deletion is used as default. We could, however, impute missing values either using single or multiple imputation, or we could opt for using FIML method – Full Information Maximum Likelihood – which proceeds from raw data rather than from the covariance matrix. In that case, we make sure that missing values are identified in the .PSF file, and don't obtain any matrices in PRELIS. Instead, we input the raw data into LISREL by including the following into the DA statement:

```
RA=C:\nys2.PSF
```