

Examples of Yuan-Bentler Corrections for Nonnormal and Missing Data

The Mplus and lavaan outputs below illustrate estimation when data are nonnormal and there are missing values present. Both packages use “MLR” to request the scaled chi-square and robust standard errors described by Yuan and Bentler (2000).¹

Mplus

Mplus VERSION 8.9
MUTHEN & MUTHEN
05/07/2025 1:09 PM

INPUT INSTRUCTIONS

```
title: CFA of three negative exchanges factors;

data: file=missing.dat; format=free;

variable: names = neg6 neg26 neg30 neg35
           neg11 neg12 neg13 neg14
           neg16 neg17 neg19 neg20;

           missing=neg6-neg20 (-99);

usevariables=neg6 neg26 neg30 neg35
           neg11 neg12 neg13 neg14
           neg16 neg17 neg19 neg20;

!Note: with missing data estimator=mlr is used to obtain robust estimates (Yuan &
!Bentler, 2000), if non-robust estimates are desired use estimator=ml;
! Missing data estimation is now the default in Version 5 and higher;

analysis: estimator=mlr;

model: hostile by neg6-neg35;
       badadv by neg11-neg14;
       demands by neg16-neg20;

output: stdyx;
```

INPUT READING TERMINATED NORMALLY

CFA of three negative exchanges factors;

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	275
Number of dependent variables	12
Number of independent variables	0
Number of continuous latent variables	3

Observed dependent variables

Continuous					
NEG6	NEG26	NEG30	NEG35	NEG11	NEG12
NEG13	NEG14	NEG16	NEG17	NEG19	NEG20

Continuous latent variables

HOSTILE	BADADV	DEMANDS
---------	--------	---------

Estimator	MLR
Information matrix	OBSERVED
Maximum number of iterations	1000

¹ Yuan K. H. and P.M. Bentler. 2000. “Three Likelihood-Based Methods for Mean and Covariance Structure Analysis with Non-Normal Missing Data.” *Sociological Methodology* 2000:165-200. Washington, D.C. American Sociological.

Convergence criterion 0.500D-04
 Maximum number of steepest descent iterations 20
 Maximum number of iterations for H1 2000
 Convergence criterion for H1 0.100D-03

Input data file(s)
 missing.dat

Input data format FREE

SUMMARY OF DATA

Number of missing data patterns 7

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

	Covariance Coverage				
	NEG6	NEG26	NEG30	NEG35	NEG11
NEG6	0.971				
NEG26	0.967	0.996			
NEG30	0.967	0.993	0.996		
NEG35	0.971	0.996	0.996	1.000	
NEG11	0.967	0.964	0.964	0.967	0.967
NEG12	0.971	0.967	0.967	0.971	0.967
NEG13	0.967	0.964	0.964	0.967	0.964
NEG14	0.971	0.967	0.967	0.971	0.967
NEG16	0.971	0.967	0.967	0.971	0.967
NEG17	0.716	0.716	0.713	0.716	0.713
NEG19	0.971	0.996	0.996	1.000	0.967
NEG20	0.971	0.996	0.996	1.000	0.967

	Covariance Coverage				
	NEG12	NEG13	NEG14	NEG16	NEG17
NEG12	0.971				
NEG13	0.967	0.967			
NEG14	0.971	0.967	0.971		
NEG16	0.971	0.967	0.971	0.971	
NEG17	0.716	0.713	0.716	0.716	0.716
NEG19	0.971	0.967	0.971	0.971	0.716
NEG20	0.971	0.967	0.971	0.971	0.716

	Covariance Coverage	
	NEG19	NEG20
NEG19	1.000	
NEG20	1.000	1.000

UNIVARIATE SAMPLE STATISTICS

UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS

Variable/ Sample Size	Mean/ Variance	Skewness/ Kurtosis	Minimum/ Maximum	% with Min/Max	Percentiles 20%/60%	Percentiles 40%/80%	Median
NEG6	0.618	1.275	0.000	52.43%	0.000	0.000	0.000
NEG26	0.581	2.006	4.000	0.75%	1.000	1.000	0.000
NEG30	0.599	1.205	0.000	55.84%	0.000	0.000	0.000
	274.000	0.605	1.106	4.000	0.36%	1.000	1.000
NEG30	0.668	1.027	0.000	49.64%	0.000	0.000	1.000
	274.000	0.601	0.804	4.000	0.36%	1.000	1.000

NEG35		0.462	1.817	0.000	66.18%	0.000	0.000	0.000
	275.000	0.561	3.636	4.000	0.73%	0.000	1.000	
NEG11		0.470	1.345	0.000	62.78%	0.000	0.000	0.000
	266.000	0.467	1.250	3.000	1.13%	0.000	1.000	
NEG12		0.431	1.312	0.000	65.54%	0.000	0.000	0.000
	267.000	0.425	0.822	3.000	0.37%	0.000	1.000	
NEG13		0.838	0.667	0.000	40.60%	0.000	0.000	1.000
	266.000	0.684	-0.133	4.000	0.38%	1.000	2.000	
NEG14		0.865	0.933	0.000	41.57%	0.000	0.000	1.000
	267.000	0.821	0.536	4.000	1.12%	1.000	2.000	
NEG16		0.547	1.706	0.000	58.43%	0.000	0.000	0.000
	267.000	0.600	3.682	4.000	1.12%	1.000	1.000	
NEG17		0.376	2.119	0.000	71.57%	0.000	0.000	0.000
	197.000	0.468	5.185	4.000	0.51%	0.000	1.000	
NEG19		0.982	0.837	0.000	34.55%	0.000	1.000	1.000
	275.000	0.854	0.447	4.000	1.45%	1.000	2.000	
NEG20		0.662	0.942	0.000	51.64%	0.000	0.000	0.000
	275.000	0.624	0.050	3.000	2.18%	1.000	1.000	

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 39

Loglikelihood

H0 Value	-2884.221
H0 Scaling Correction Factor for MLR	1.5834
H1 Value	-2806.647
H1 Scaling Correction Factor for MLR	1.5080

Information Criteria

Akaike (AIC)	5846.443
Bayesian (BIC)	5987.497
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	5863.835

Chi-Square Test of Model Fit

Value	106.973*
Degrees of Freedom	51
P-Value	0.0000
Scaling Correction Factor for MLR	1.4504

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.063
90 Percent C.I.	0.046 0.080
Probability RMSEA <= .05	0.095

CFI/TLI

CFI	0.948
TLI	0.932

Chi-Square Test of Model Fit for the Baseline Model

Value	1137.407
Degrees of Freedom	66
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value 0.049

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
HOSTILE BY				
NEG6	1.000	0.000	999.000	999.000
NEG26	1.256	0.124	10.154	0.000
NEG30	1.204	0.122	9.839	0.000
NEG35	1.015	0.096	10.571	0.000
BADADV BY				
NEG11	1.000	0.000	999.000	999.000
NEG12	1.017	0.133	7.643	0.000
NEG13	1.375	0.146	9.421	0.000
NEG14	1.520	0.151	10.047	0.000
DEMANDS BY				
NEG16	1.000	0.000	999.000	999.000
NEG17	1.024	0.113	9.059	0.000
NEG19	1.019	0.178	5.712	0.000
NEG20	0.987	0.159	6.222	0.000
BADADV WITH HOSTILE				
	0.194	0.038	5.100	0.000
DEMANDS WITH HOSTILE				
	0.256	0.069	3.696	0.000
BADADV	0.208	0.042	4.995	0.000
Intercepts				
NEG6	0.622	0.047	13.336	0.000
NEG26	0.597	0.047	12.719	0.000
NEG30	0.666	0.047	14.252	0.000
NEG35	0.462	0.045	10.222	0.000
NEG11	0.473	0.042	11.288	0.000
NEG12	0.434	0.040	10.856	0.000
NEG13	0.842	0.051	16.657	0.000
NEG14	0.870	0.055	15.684	0.000
NEG16	0.551	0.048	11.591	0.000
NEG17	0.395	0.047	8.362	0.000
NEG19	0.982	0.056	17.616	0.000
NEG20	0.662	0.048	13.896	0.000
Variances				
HOSTILE	0.282	0.070	4.051	0.000
BADADV	0.240	0.045	5.295	0.000
DEMANDS	0.312	0.083	3.783	0.000
Residual Variances				
NEG6	0.302	0.035	8.685	0.000
NEG26	0.159	0.031	5.064	0.000
NEG30	0.191	0.028	6.839	0.000
NEG35	0.271	0.058	4.651	0.000
NEG11	0.228	0.030	7.493	0.000
NEG12	0.179	0.032	5.538	0.000
NEG13	0.233	0.034	6.785	0.000
NEG14	0.271	0.044	6.109	0.000
NEG16	0.291	0.046	6.290	0.000
NEG17	0.165	0.034	4.901	0.000
NEG19	0.530	0.074	7.182	0.000
NEG20	0.320	0.047	6.770	0.000

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.147E-02
 (ratio of smallest to largest eigenvalue)

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
HOSTILE BY				
NEG6	0.695	0.053	13.177	0.000
NEG26	0.858	0.030	28.153	0.000
NEG30	0.826	0.031	26.764	0.000
NEG35	0.720	0.052	13.744	0.000
BADADV BY				
NEG11	0.716	0.042	16.853	0.000
NEG12	0.763	0.047	16.230	0.000
NEG13	0.813	0.033	24.748	0.000
NEG14	0.820	0.034	24.213	0.000
DEMANDS BY				
NEG16	0.719	0.057	12.549	0.000
NEG17	0.815	0.043	18.847	0.000
NEG19	0.616	0.059	10.504	0.000
NEG20	0.698	0.053	13.111	0.000
BADADV WITH HOSTILE				
	0.745	0.049	15.336	0.000
DEMANDS WITH HOSTILE				
	0.863	0.052	16.624	0.000
	0.758	0.055	13.833	0.000
Intercepts				
NEG6	0.814	0.048	16.878	0.000
NEG26	0.767	0.044	17.428	0.000
NEG30	0.860	0.048	17.879	0.000
NEG35	0.616	0.039	15.857	0.000
NEG11	0.691	0.041	16.670	0.000
NEG12	0.664	0.041	16.360	0.000
NEG13	1.016	0.055	18.586	0.000
NEG14	0.957	0.051	18.913	0.000
NEG16	0.709	0.043	16.352	0.000
NEG17	0.562	0.043	13.205	0.000
NEG19	1.062	0.054	19.685	0.000
NEG20	0.838	0.046	18.401	0.000
Variances				
HOSTILE	1.000	0.000	999.000	999.000
BADADV	1.000	0.000	999.000	999.000
DEMANDS	1.000	0.000	999.000	999.000
Residual Variances				
NEG6	0.517	0.073	7.046	0.000
NEG26	0.263	0.052	5.027	0.000
NEG30	0.318	0.051	6.238	0.000
NEG35	0.482	0.075	6.398	0.000
NEG11	0.487	0.061	8.013	0.000
NEG12	0.418	0.072	5.829	0.000
NEG13	0.339	0.053	6.342	0.000
NEG14	0.328	0.056	5.908	0.000
NEG16	0.483	0.082	5.858	0.000
NEG17	0.335	0.071	4.747	0.000
NEG19	0.621	0.072	8.593	0.000
NEG20	0.513	0.074	6.898	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
NEG6	0.483	0.073	6.588	0.000
NEG26	0.737	0.052	14.076	0.000
NEG30	0.682	0.051	13.382	0.000

NEG35	0.518	0.075	6.872	0.000
NEG11	0.513	0.061	8.427	0.000
NEG12	0.582	0.072	8.115	0.000
NEG13	0.661	0.053	12.374	0.000
NEG14	0.672	0.056	12.106	0.000
NEG16	0.517	0.082	6.275	0.000
NEG17	0.665	0.071	9.424	0.000
NEG19	0.379	0.072	5.252	0.000
NEG20	0.487	0.074	6.555	0.000

Beginning Time: 13:09:04
 Ending Time: 13:09:04
 Elapsed Time: 00:00:00

MUTHEN & MUTHEN
 3463 Stoner Ave.
 Los Angeles, CA 90066

Tel: (310) 391-9971
 Fax: (310) 391-8971
 Web: www.StatModel.com
 Support: Support@StatModel.com

Copyright (c) 1998-2023 Muthen & Muthen

lavaan

```
> library(lavaan)
>
> ## Missing data estimation example
>
> misssdat = read.table("c:/jason/plusplus/semclass/missing.dat", header=FALSE)
> names(misssdat) = c("neg6","neg26","neg30","neg35",
+ "neg11","neg12","neg13","neg14",
+ "neg16","neg17","neg19","neg20")
> misssdat[misssdat == -99] <- NA
>
> model = '
+ hostile =~ neg6 + neg26 + neg30 + neg35
+ badadv =~ neg11 + neg12 + neg13 + neg14
+ demands =~ neg16 + neg17 + neg19 + neg20
+
>
> #Yuan-Bentler robust estimates for nonnormal missing data
> fit = sem(model, data = misssdat, missing = "fiml", estimator="mlr")
> summary(fit,fit.measures=TRUE, rsquare=TRUE, standardized=TRUE)
lavaan 0.6-18 ended normally after 51 iterations
```

Estimator	ML
Optimization method	NLMINB
Number of model parameters	39
Number of observations	275
Number of missing patterns	7

Model Test User Model:		
Test Statistic	Standard	Scaled
Degrees of freedom	155.149	106.973
P-value (Chi-square)	51	51
Scaling correction factor	0.000	0.000
Yuan-Bentler correction (Mplus variant)		1.450

Model Test Baseline Model:		
Test statistic	1758.226	1137.409
Degrees of freedom	66	66
P-value	0.000	0.000
Scaling correction factor		1.546

User Model versus Baseline Model:		
Comparative Fit Index (CFI)	0.938	0.948
Tucker-Lewis Index (TLI)	0.920	0.932
Robust Comparative Fit Index (CFI)		0.950
Robust Tucker-Lewis Index (TLI)		0.935

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-2884.221	-2884.221
Scaling correction factor for the MLR correction		1.583
Loglikelihood unrestricted model (H1)	-2806.647	-2806.647
Scaling correction factor for the MLR correction		1.508
Akaike (AIC)	5846.443	5846.443
Bayesian (BIC)	5987.497	5987.497
Sample-size adjusted Bayesian (SABIC)	5863.835	5863.835

Root Mean Square Error of Approximation:

RMSEA	0.086	0.063
90 Percent confidence interval - lower	0.071	0.049
90 Percent confidence interval - upper	0.102	0.077
P-value H_0: RMSEA <= 0.050	0.000	0.060
P-value H_0: RMSEA >= 0.080	0.756	0.022
Robust RMSEA		0.079
90 Percent confidence interval - lower		0.058
90 Percent confidence interval - upper		0.101
P-value H_0: Robust RMSEA <= 0.050		0.015
P-value H_0: Robust RMSEA >= 0.080		0.503

Standardized Root Mean Square Residual:

SRMR	0.049	0.049
------	-------	-------

Parameter Estimates:

Standard errors Information based on Observed information based on	Sandwich Observed Hessian
--	---------------------------------

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
hostile =~						
neg6	1.000				0.531	0.695
neg26	1.256	0.124	10.154	0.000	0.668	0.858
neg30	1.204	0.122	9.840	0.000	0.640	0.826
neg35	1.015	0.096	10.571	0.000	0.539	0.720
badadv =~						
neg11	1.000				0.490	0.716
neg12	1.017	0.133	7.643	0.000	0.499	0.763
neg13	1.375	0.146	9.421	0.000	0.674	0.813
neg14	1.520	0.151	10.047	0.000	0.745	0.820
demands =~						
neg16	1.000				0.559	0.719
neg17	1.024	0.113	9.059	0.000	0.572	0.815
neg19	1.019	0.178	5.711	0.000	0.569	0.616
neg20	0.987	0.159	6.222	0.000	0.551	0.698

Covariances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
hostile ~~ badadv	0.194	0.038	5.100	0.000	0.745	0.745
hostile ~~ demands	0.256	0.069	3.696	0.000	0.863	0.863
badadv ~~ demands	0.208	0.042	4.995	0.000	0.758	0.758

Intercepts:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
.neg6	0.622	0.047	13.336	0.000	0.622	0.814
.neg26	0.597	0.047	12.719	0.000	0.597	0.767
.neg30	0.666	0.047	14.252	0.000	0.666	0.860
.neg35	0.462	0.045	10.222	0.000	0.462	0.616
.neg11	0.473	0.042	11.288	0.000	0.473	0.691
.neg12	0.434	0.040	10.856	0.000	0.434	0.664
.neg13	0.842	0.051	16.657	0.000	0.842	1.016
.neg14	0.870	0.055	15.684	0.000	0.870	0.957
.neg16	0.551	0.048	11.590	0.000	0.551	0.709
.neg17	0.395	0.047	8.362	0.000	0.395	0.562
.neg19	0.982	0.056	17.616	0.000	0.982	1.062
.neg20	0.662	0.048	13.896	0.000	0.662	0.838

Variances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
.neg6	0.302	0.035	8.685	0.000	0.302	0.517
.neg26	0.159	0.031	5.064	0.000	0.159	0.263
.neg30	0.191	0.028	6.839	0.000	0.191	0.318
.neg35	0.271	0.058	4.651	0.000	0.271	0.482
.neg11	0.228	0.030	7.493	0.000	0.228	0.487
.neg12	0.179	0.032	5.538	0.000	0.179	0.418

.neg13	0.233	0.034	6.785	0.000	0.233	0.339
.neg14	0.271	0.044	6.109	0.000	0.271	0.328
.neg16	0.291	0.046	6.290	0.000	0.291	0.483
.neg17	0.165	0.034	4.901	0.000	0.165	0.335
.neg19	0.530	0.074	7.182	0.000	0.530	0.621
.neg20	0.320	0.047	6.770	0.000	0.320	0.513
hostile	0.282	0.070	4.051	0.000	1.000	1.000
badadv	0.240	0.045	5.295	0.000	1.000	1.000
demands	0.312	0.083	3.783	0.000	1.000	1.000

R-Square:

	Estimate
neg6	0.483
neg26	0.737
neg30	0.682
neg35	0.518
neg11	0.513
neg12	0.582
neg13	0.661
neg14	0.672
neg16	0.517
neg17	0.665
neg19	0.379
neg20	0.487