Advanced Experimental Design
Psych 464
Jeffrey D. Leitzel, Ph.D.

Topic 4
Confirmatory Factor Analysis (CFA)

1

Outline/Overview

- Readings
- EFA vs. CFA
- Isolating True Score variability
- Specialized analyses=specialized software
- Estimation techniques
- Running CFA in Stata
- Postestimation – goodness of fit, residuals, modification indices
- Example – CFA of Rosenberg Self-Esteem Scale

2

Readings

- Pg. 11-57 in Acock book.
  - do the examples
- Stata SEM manual
  - pg. 7-15, in Intro 2
  - Intro 5,
    - single factor measurement models
    - multiple factor measurement models
    - CFA models
    - higher order CFA models

3

EFA vs. CFA

- EFA each indicator is associated with all factors.
  - No restrictions on loadings
- CFA determine whether the number of factors and the loadings conform based on theory
- Path models treated exogenous variables as though measured without error,
- Examine reliability and validity and if acceptable, use the scores in statistical analyses-traditional techniques do not adjust for measurement error in any way

4

Psychometric perspective

- Any measure we use consists of two components, traditional techniques do not separate the components.
- Observed Score = True Score + Error
- Error = noise, can obscure or attenuate the relationship between variables
- CFA allows us to estimate true score components
- Latent variables are thought to be "cleansed" of measurement error

5

SEM Software

- CFA and Structural Equation Modeling programs
- Commercial programs
  - LISREL - Karl Joreskog
  - M-Plus - Bengt Muthen
  - EQS - Peter Bentler
  - AMOS - Jim Arbuckle
- Free options
  - Mx - Mike Neale
  - R has a SEM package and LAVAAN

6

CFA diagrams/setting up in Stata

- Latent variable (factor) is large oval
- Observed variables are squares or rectangles
- Arrows point from the latent variable to the observed variables, indicates that the latent variable is responsible for the individual's level on the observed variable
- Each observed variable has an error term
- Run in Stata using the SEM builder or the SEM command (can also use GSEM command)

7

Setting up CFA in Stata

- Have to set the scale of the latent variable
  - first indicator for a factor used as reference indicator, unstandardized loading is set to 1.0.
  - not an issue with standardized solution
- Latents must start with a capital letter
- Estimation methods
  - Maximum Likelihood [+VCE(robust)]
  - Asymptotic Distribution Free
  - Maximum Likelihood with Missing Values
- Typically use Maximum Likelihood

8

The SEM command

- For a one factor model, takes the form:

SEM (Latent-> item1 … item_n), method(ml) standardized

- Main model test is the Chi-Squared statistic
  - Test works the opposite of what you have learned
  - The Chi-Squared is test the discrepancy between the observed and model-implied covariance matrices
- Chi-Square is very sensitive to sample size
- Use fit indices to assess model fit

9

Post estimation

- Goodness of fit [estat gof, stats(all)]
- Residuals [estat residuals]
- Model implied covariance matrix
  [estat framework, fitted]
- Modification indices [estat mindices]

- After running the model, we may make modifications to improve fit, must be reasonable

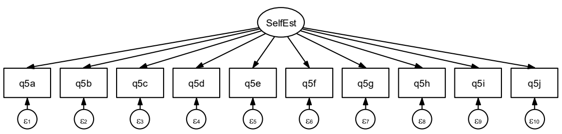
10

Fit indices

- RMSEA (Root mean square error of approximation) –Hu and Bentler (1999) suggest <.06, Browne and Cudeck (1993) suggest <.05=good fit, between .05-.08=adequate fit and >.1=poor fit
- pclose corresponds to a test of RMSEA < .05
- AIC and BIC useful for comparing models
- CFI (comparative fit index) and TLI (Tucker-Lewis index) - incremental fit indices, want values greater than .95
- SRMR (Standardized Root Mean Square Residual) absolute measure of fit - standardized difference between the observed correlations and the predicted correlations. <.08=good fit
- CD(coefficient of determination) closer to 1=better fit

11

CFA of Rosenberg S.E. Scale



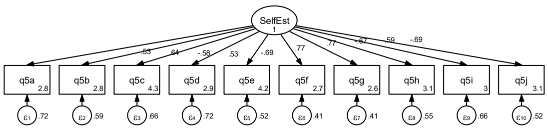
12

Estimate the model

- Typically will be interested in standardized solution
- Intercepts can be ignored
- Will then look at modification indices

13

CFA of Rosenberg S.E. Scale (cont.)



$\chi^2(35)=734.41$, p > .0000, RMSEA=.12, pclose=.000, AIC=25440.66, BIC=25596.61, CFI=.87, TLI=.83, SRMR=.06, CD=.89

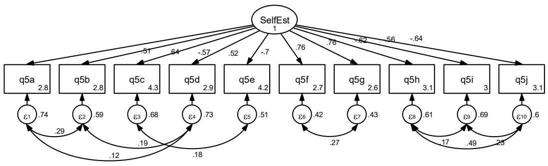
14

SEM Modification indices

```
Modification indices
----------------------------------------------------------------------
                |                                              Standard
                |        MI      df    P>MI           EPC          EPC
----------------+-----------------------------------------------------
cov(e.q5a,e.q5b)|    114.681      1    0.00      .0761988      .316855
cov(e.q5a,e.q5d)|     15.982      1    0.00      .0309309     .1153873
cov(e.q5a,e.q5h)|     29.379      1    0.00      .0543851     .1619546
cov(e.q5a,e.q5i)|     16.094      1    0.00      .0439562     .1169391
cov(e.q5a,e.q5j)|     12.978      1    0.00      .0360964     .1083898
cov(e.q5b,e.q5d)|     49.679      1    0.00       .051317     .2086486
cov(e.q5b,e.q5f)|      8.579      1    0.00      .0209402     .0957948
<11 rows omitted>
cov(e.q5f,e.q5g)|    138.332      1    0.00      .0949114     .4241064
cov(e.q5f,e.q5h)|     24.015      1    0.00      .0496578     .1626854
cov(e.q5f,e.q5i)|      8.902      1    0.00       .032594     .0953951
cov(e.q5f,e.q5j)|     56.081      1    0.00      .0760592     .2512606
cov(e.q5g,e.q5h)|     10.550      1    0.00      .0340917     .1080806
cov(e.q5g,e.q5j)|     11.149      1    0.00      .0351283     .1122968
cov(e.q5h,e.q5i)|     18.570      1    0.00      .0627778     .1303576
cov(e.q5h,e.q5j)|    313.261      1    0.00      .2373183     .5562199
cov(e.q5i,e.q5j)|     46.065      1    0.00      .0987791     .2068269
----------------------------------------------------------------------
EPC = expected parameter change
```

15

CFA of Rosenberg S.E. Scale (cont.)



$\chi^2(27)=81.88$, p > .000, RMSEA=.04, pclose=.97, AIC=24804.12, BIC=25001.66, CFI=.99, TLI=.98, SRMR=.02, CD=.85

16