



# Linearni strukturni model

Multivariatna analiza

# Linearni strukturni modeli (SEM)

- Ang. (Linear) structural equation models (SEM)
- Multivariatna metoda, katere cilj je dobiti vpogled v odnose med več spremenljivkami
- Lahko jo smatramo kot kombinacijo in razširitev linearne/ih regresij(e) in faktorske analize.

# SEM se uporablja za:

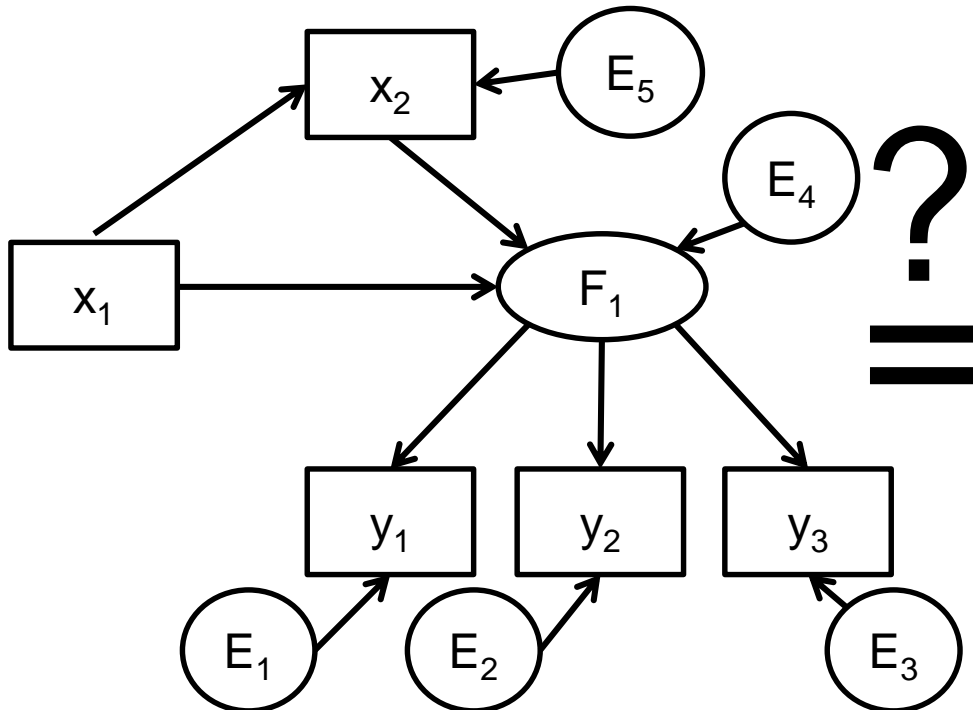
- Validacijo (preverjanje) merskih inštrumentov, merjenje merske napake.
- Ocenjevanje linearnih odnosov med večjim številom spremenljivk, kjer je lahko odvisnih spremenljivk več.
- Zavračanje vzročnih teorij.
- ...

# SEM – Linearni strukturni **MODEL**



?

=



	$x_1$	$x_2$	$y_1$	$y_2$	$y_3$
$x_1$	0.96	0.80	0.51	0.52	0.50
$x_2$	0.80	1.04	0.64	0.62	0.63
$y_1$	0.51	0.64	0.96	0.61	0.61
$y_2$	0.52	0.62	0.61	0.96	0.61
$y_3$	0.50	0.63	0.61	0.61	0.94

# Kaj je SEM?

- Multivariatna metoda, katere cilj je dobiti vpogled v odnose med več *spremenljivkami*
- Ne modeliramo le ene enačbe, ampak *sistem enačb*:

- ☐ Regresija:  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

---

$$Y_1 = \lambda_{11} F_1 + \epsilon_1$$

$$Y_2 = \lambda_{21} F_1 + \epsilon_2$$

- ☐ SEM:  $Y_3 = \lambda_{31} F_1 + \epsilon_3$

$$F_1 = \beta_1 X_1 + \beta_2 X_2 + \epsilon_4$$

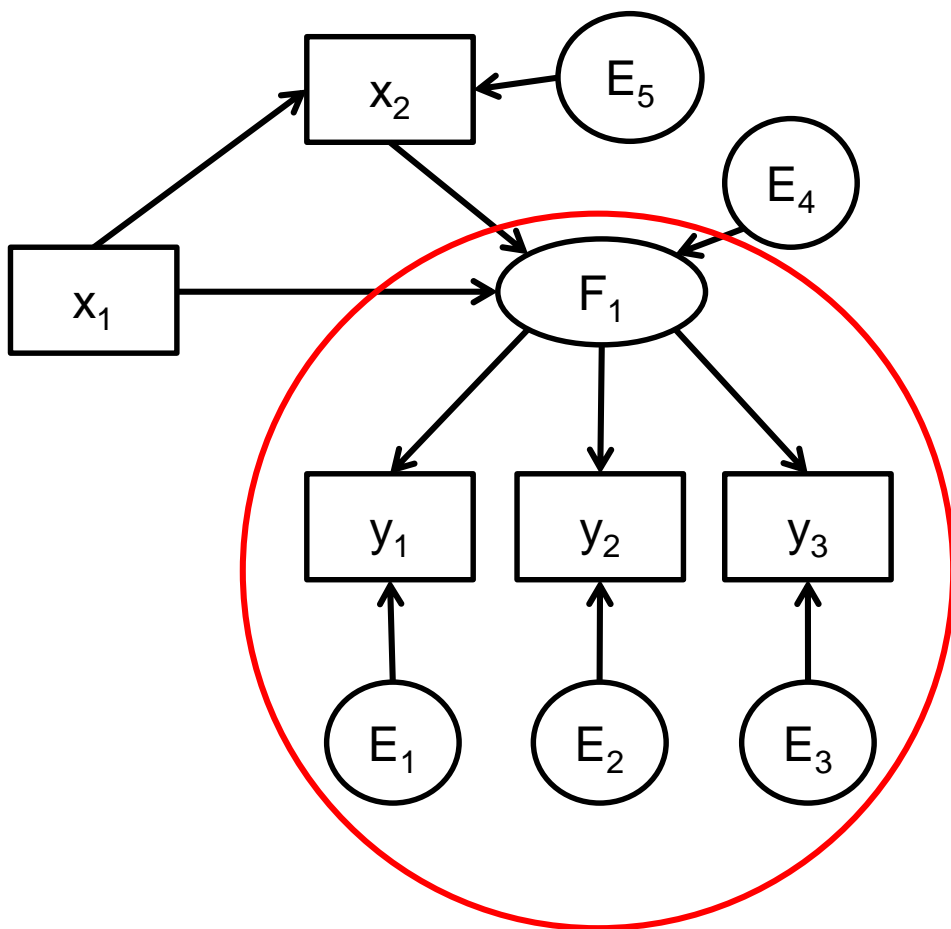
$$X_2 = \beta_3 X_1 + \epsilon_5$$

# Prednosti SEM

- Ocenjevanje skritih (latentnih) spremenljivk
  - Pogosto raziskovalce zanimajo koncepti, ki jih ni mogoče direktno meriti  
Npr.: vrednote, stališča, socialno-ekonomski status, inteligenca
  - „Obvoz“: Merimo jih posredno preko večjega števila indikatorjev, ki vključujejo tudi merske napake (slučajne in ne-slučajne)
  - SEM omogoča oceniti odnose med skritimi spremenljivkami (namesto ocenjevanja odnosov med nezanesljivimi indikatorji)

# Prednosti SEM

- Ocenjevanje skritih (latentnih) spremenljivk



$$Y_1 = \lambda_{11}F_1 + \epsilon_1$$

$$Y_2 = \lambda_{21}F_1 + \epsilon_2$$

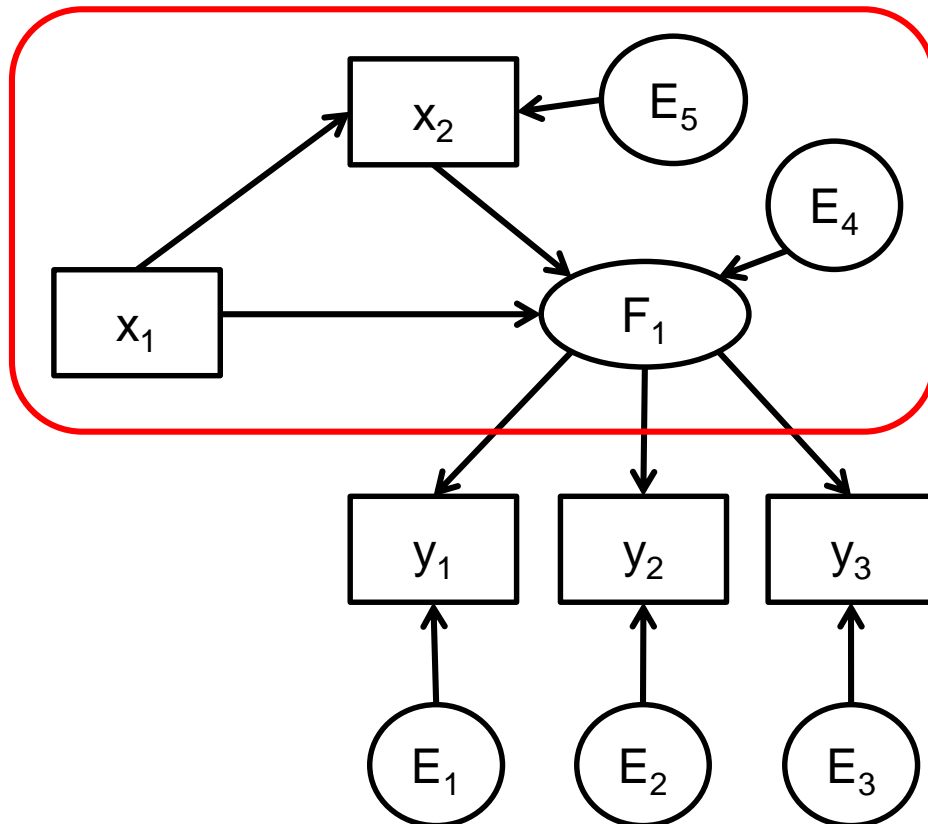
$$Y_3 = \lambda_{31}F_1 + \epsilon_3$$

$$F_1 = \beta_1X_1 + \beta_2X_2 + \epsilon_4$$

$$X_2 = \beta_3X_1 + \epsilon_5$$

# Prednosti SEM

- Ocenjevanje neposrednih in posrednih učinkov: ocenimo lahko, ali odnosi potekajo preko tretje spremenljivke



$$Y_1 = \lambda_{11}F_1 + \epsilon_1$$

$$Y_2 = \lambda_{21}F_1 + \epsilon_2$$

$$Y_3 = \lambda_{31}F_1 + \epsilon_3$$

$$F_1 = \beta_1 X_1 + \beta_2 X_2 + \epsilon_4$$

$$X_2 = \beta_3 X_1 + \epsilon_5$$



# Diagrami poti (ang. path diagrams)

## ■ Grafična notacija za SEM

### □ Tipi spremenljivk:

 Skrita oz. latentna spremenljivka

 Merljiva oz. manifestna spremenljivka

 Stohastični člen / Sprecifični faktor

# Diagrami poti

## ■ Grafična notacija za SEM

### □ Tipi odnosov/učinkov

————→ Enosmerni učinek

↔ Korelacija

↔ Obojestranski (dvosmerni) učinek/odnos

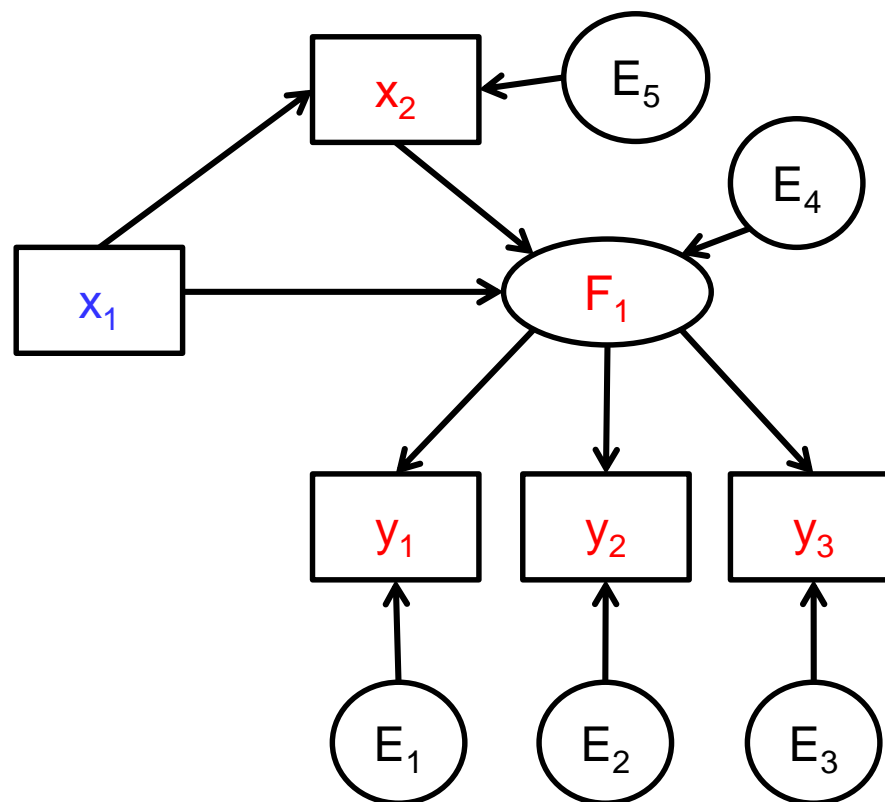
Ni povezave/učinka

# Diagrami poti

## ■ Spremenljivke:

□ **eksogene** (*niso pojasnjene z modelom*)

□ **enodogene** (*so pojasnjene z modelom*)



# Diagrami poti

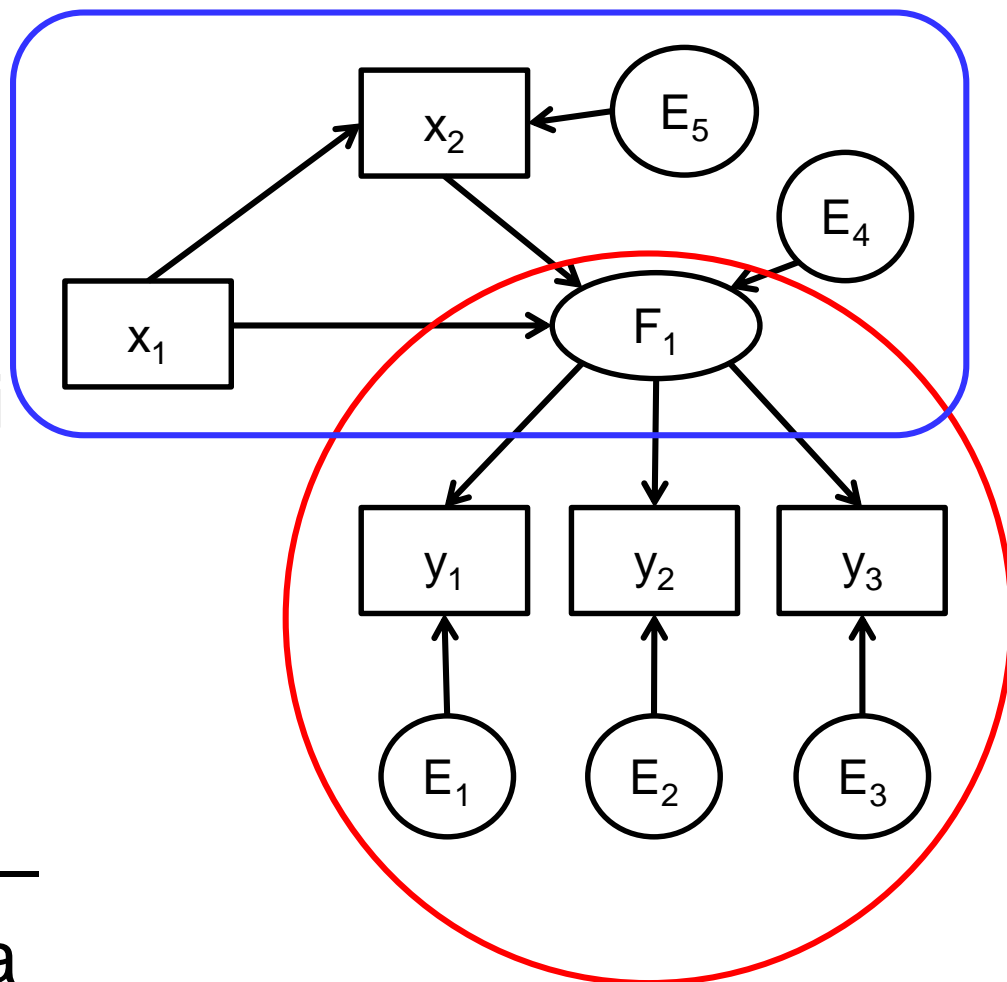
## ■ Deli modela:

### □ **Merski model**

(odnosi med indikatorji/merjenimi spremenljivkami in konstrukti/skritimi spremenljivkami)

### □ **Strukturni model**

(učinki med sprem. – vse razen merskega modela)



# Odnos med grafičnim modelom in matematično notacijo

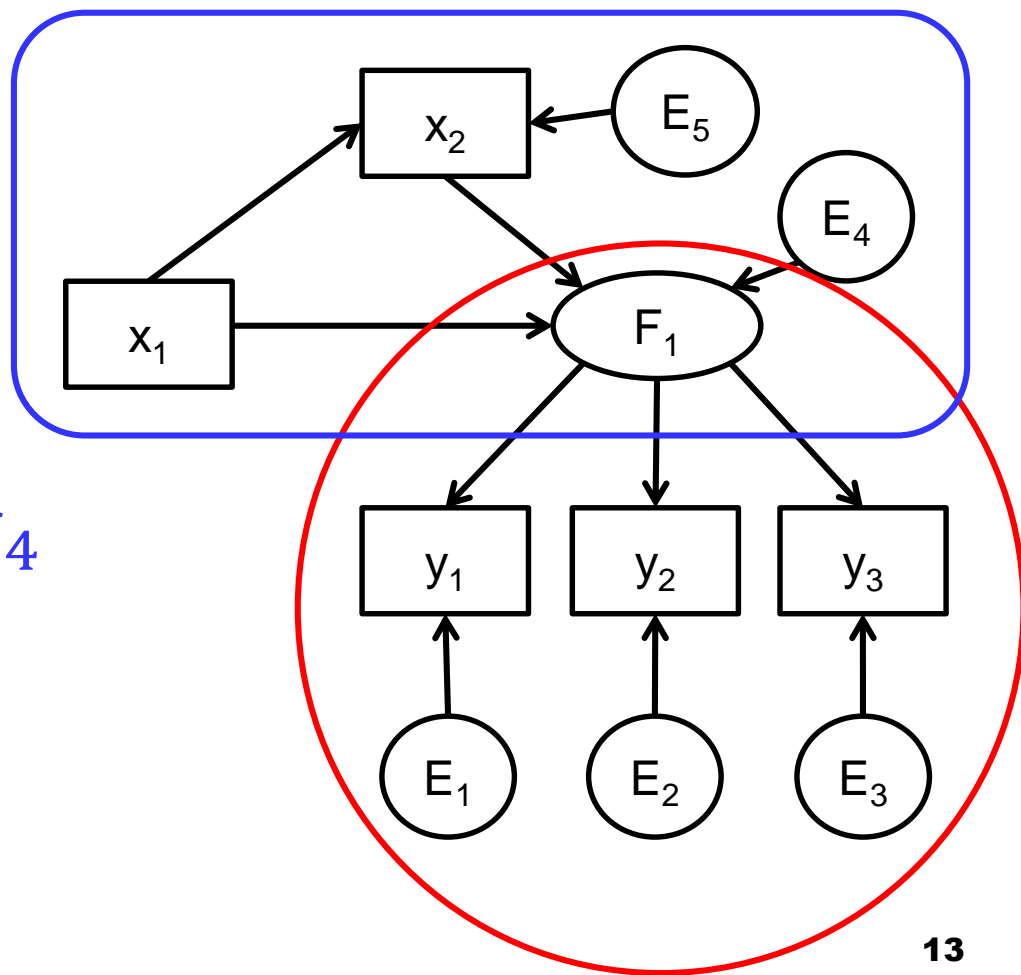
$$Y_1 = \lambda_{11}F_1 + \epsilon_1$$

$$Y_2 = \lambda_{21}F_1 + \epsilon_2$$

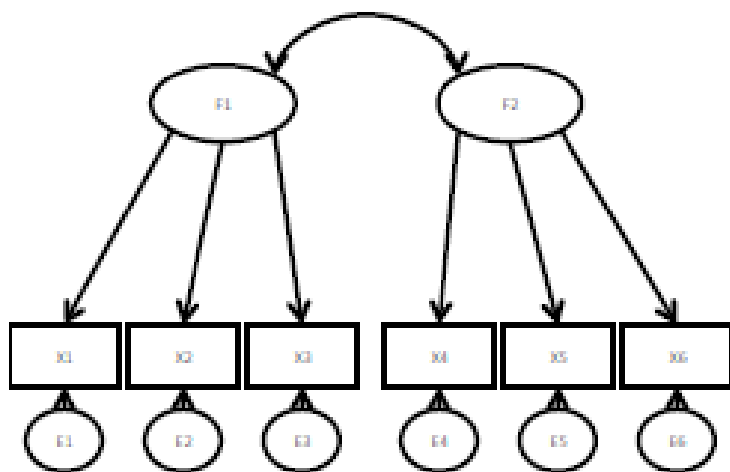
$$Y_3 = \lambda_{31}F_1 + \epsilon_3$$

$$F_1 = \beta_1X_1 + \beta_2X_2 + \epsilon_4$$

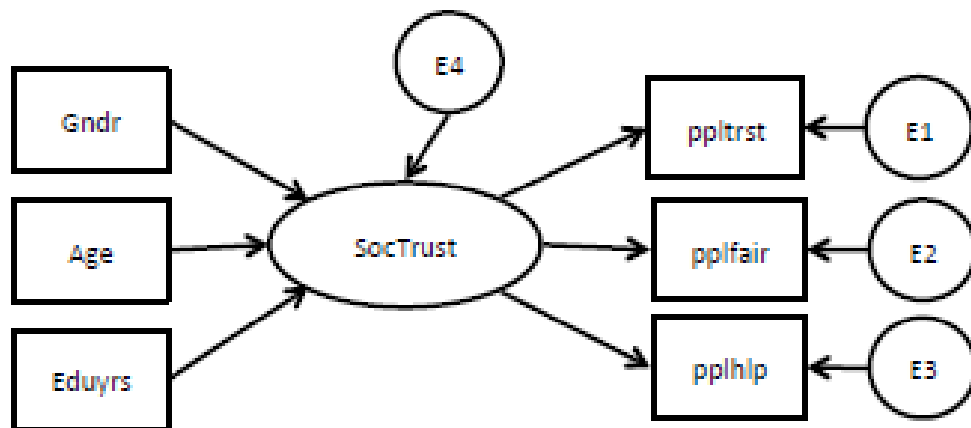
$$X_2 = \beta_3X_1 + \epsilon_5$$



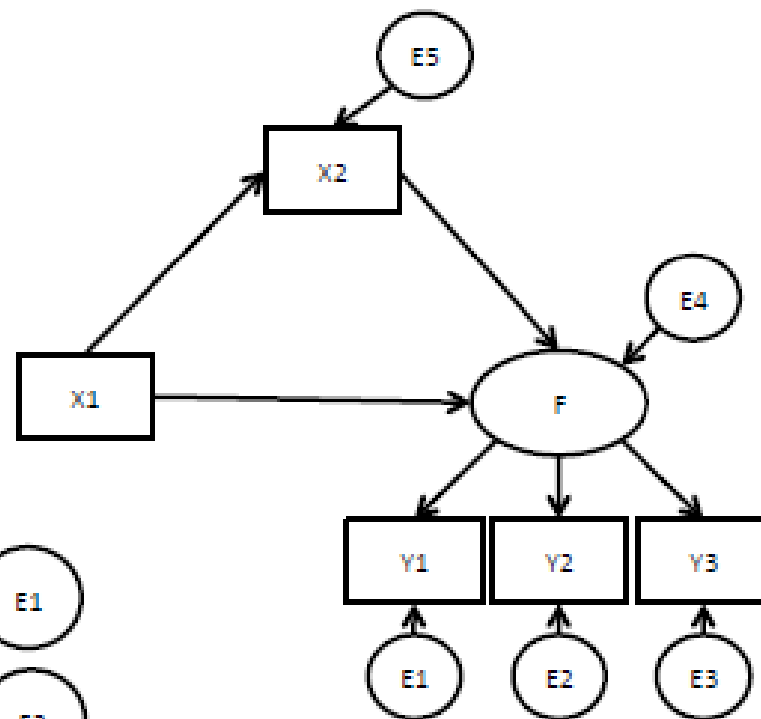
# Vrste SEM modelov



**CFA**

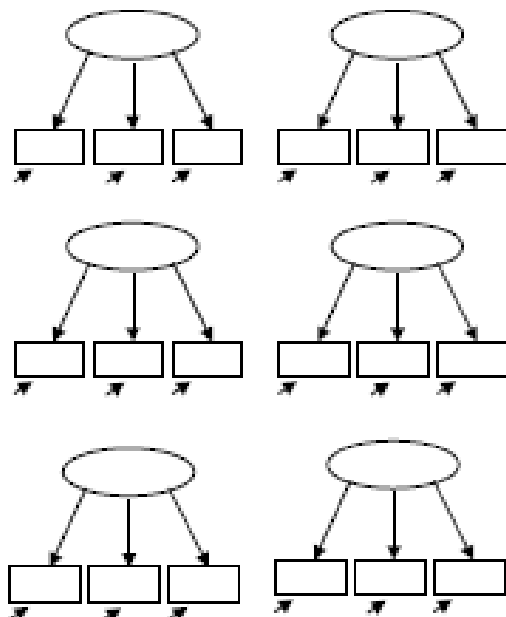


**MIMIC**

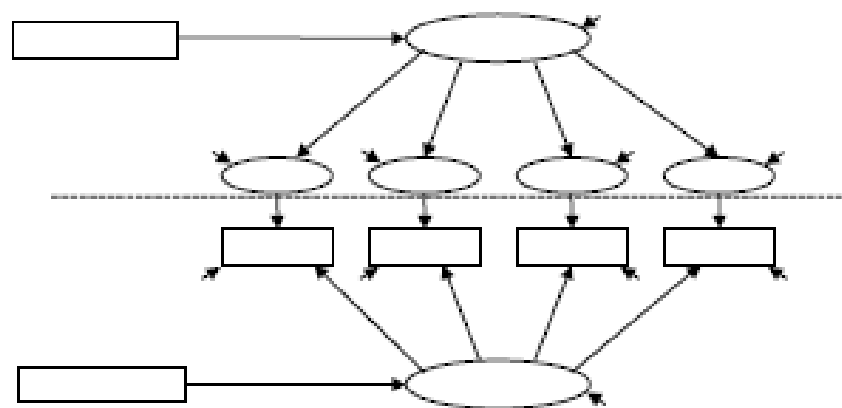


**FULL SEM**

# Vrste SEM modelov



**Multigroup  
SEM**



**Multilevel  
SEM**

in še več ...

# Konfirmatorna faktorska analiza (ang. CFA)

- CFA: način predstavitve skritih konstruktov, ki so merjeni z več indikatorji, ki vključujejo merske napake

$$\begin{array}{rcccl} X_1 & = & T & + & E_1 \\ X_2 & = & T & + & E_2 \end{array}$$

$$\text{Meritev} = \text{Prava vrednost} + \text{Napaka}$$

- Cilj: s skritimi konstrukti pojasniti odnose med indikatorji

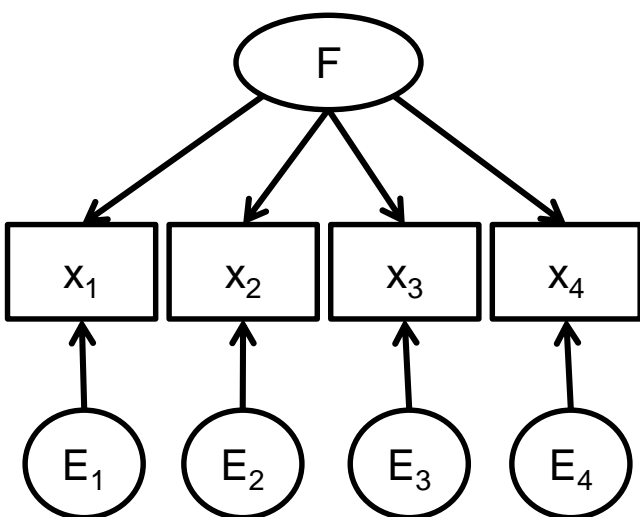


# Konfirmatorna faktorska analiza (ang. CFA)

- CFA vs. likertove lestvice (vsote ali povprečja)
  - CFA testira predpostavke, ki se jih predpostavi pri uporabi likertovih lestvic
- CFA vs. Eksploratorna faktorska analiza (EFA)
  - CFA testira model, ki ga postavi uporabnik → testiranje domnev
  - EFA „išče“ pravi model

# Notacija CFA modela

Diagram



Enačbe

$$X_1 = \lambda_{11}F_1 + \epsilon_1$$

$$X_2 = \lambda_{21}F_1 + \epsilon_2$$

$$X_3 = \lambda_{31}F_1 + \epsilon_3$$

$$X_4 = \lambda_{41}F_1 + \epsilon_4$$

Matrike

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} F + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix}$$

$$\mathbf{X} = \boldsymbol{\lambda}\mathbf{F} + \boldsymbol{\epsilon}$$

# Parametri CFA modela:

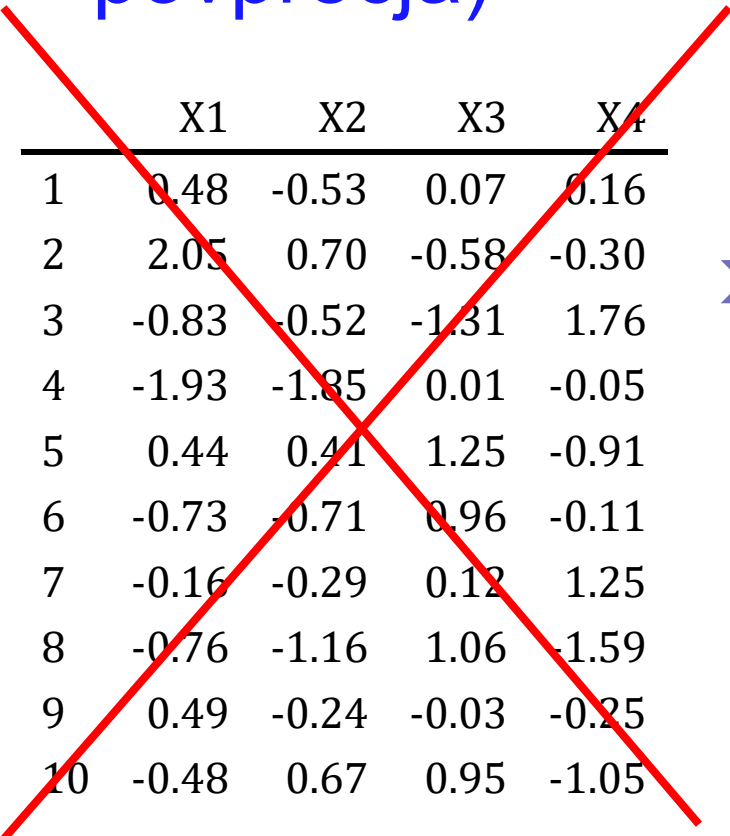
- Faktorske uteži: odnosi med skritimi konstrukti in merjenimi indikatorji —  $\lambda$
- Variance (in včasih izbrane kovariance) specifičnih faktorjev (nepojasnjena varianca):  $Var(\epsilon_i)$ ,  $Cov(\epsilon_i, \epsilon_j)$
- Kovariance in variance skritih konstruktov —  $Var(F_i)$ ,  $Cov(F_i, F_j)$

# Parametri CFA modela – matrična notacija

- Matrika faktorskih uteži: — matrika  $\Lambda$
- Matrika (ko)varianc specifičnih faktorjev — matrika  $\Theta$
- Kovariančna matrika skritih faktorjev — matrika  $\Psi$

# Ocenjevanje modela

- Ne modeliramo originalnih podatkov, ampak **modeliramo kovariance** (in povprečja)



	X1	X2	X3	X4
1	0.48	-0.53	0.07	0.16
2	2.05	0.70	-0.58	-0.30
3	-0.83	-0.52	-1.31	1.76
4	-1.93	-1.85	0.01	-0.05
5	0.44	0.41	1.25	-0.91
6	-0.73	-0.71	0.96	-0.11
7	-0.16	-0.29	0.12	1.25
8	-0.76	-1.16	1.06	-1.59
9	0.49	-0.24	-0.03	-0.25
10	-0.48	0.67	0.95	-1.05



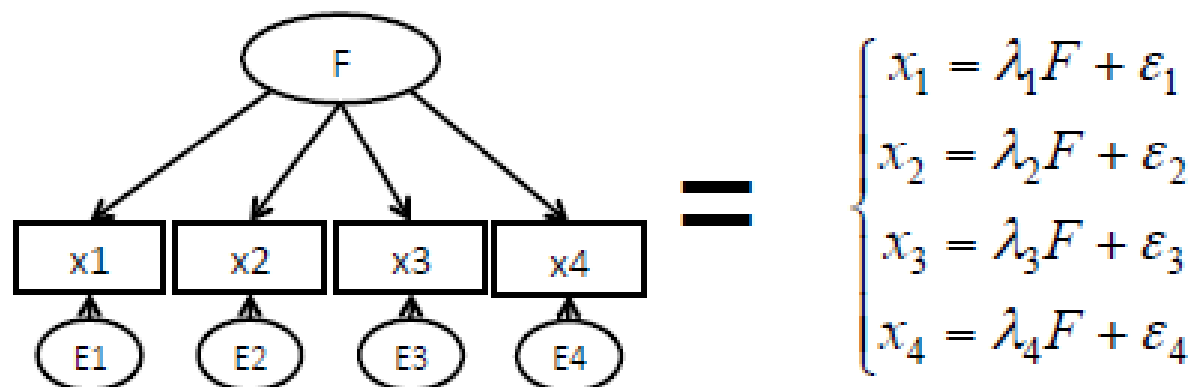
	X1	X2	X3	X4
X1	1.16	0.65	-0.13	-0.11
X2	0.65	0.65	0.04	-0.11
X3	-0.13	0.04	0.66	-0.62
X4	-0.11	-0.11	-0.62	1.02

# Ocenjevanje modela

- Vsaka kombinacija parametrov implicira določeno kovariančno matriko (in povprečja) = reproducirana kovariančna matrika (in povprečja) oz. kovariančna matrika (in povprečja) na podlagi modela
- Paramtere ocenimo tako, da se lete čim bolj prilegajo dejanskim kovariancam (in povprečjem)
- Običajno: Metoda največjega verjetja

# Ocenjevanje modela

Npr.



implicira...

	<b>x1</b>	<b>x2</b>	<b>x3</b>	<b>x4</b>
<b>x1</b>	$\lambda_1^2 \text{Var}(F) + \text{Var}(E_1)$			
<b>x2</b>	$\lambda_1 \lambda_2 \text{Var}(F)$	$\lambda_2^2 \text{Var}(F) + \text{Var}(E_2)$		
<b>x3</b>	$\lambda_1 \lambda_3 \text{Var}(F)$	$\lambda_2 \lambda_3 \text{Var}(F)$	$\lambda_3^2 \text{Var}(F) + \text{Var}(E_3)$	
<b>x4</b>	$\lambda_1 \lambda_4 \text{Var}(F)$	$\lambda_2 \lambda_4 \text{Var}(F)$	$\lambda_3 \lambda_4 \text{Var}(F)$	$\lambda_4^2 \text{Var}(F) + \text{Var}(E_4)$

Splošno:

$$\Sigma = \Lambda \Psi \Lambda' + \Theta$$

# Ocenjevanje modela

Dejanske kovariance  
(na vzorcu):

	<b>x1</b>	<b>x2</b>	<b>x3</b>	<b>x4</b>
<b>x1</b>	0.653			
<b>x2</b>	0.47	0.695		
<b>x3</b>	0.408	0.406	0.771	
<b>x4</b>	0.446	0.509	0.41	0.688

Rešitev 1

$\lambda_1$	0.5	Var(F)	1
$\lambda_2$	0.5	Var(E1)	0.3
$\lambda_3$	0.5	Var(E2)	0.3
$\lambda_4$	0.5	Var(E3)	0.3
		Var(E4)	0.3

=>

Implicirana kov. matrika 1

	<b>x1</b>	<b>x2</b>	<b>x3</b>	<b>x4</b>
<b>x1</b>	0.550			
<b>x2</b>	0.250	0.550		
<b>x3</b>	0.250	0.250	0.550	
<b>x4</b>	0.250	0.250	0.250	0.550

Rešitev 2

$\lambda_1$	0.653	Var(F)	1
$\lambda_2$	0.719	Var(E1)	0.226
$\lambda_3$	0.588	Var(E2)	0.177
$\lambda_4$	0.697	Var(E3)	0.425
		Var(E4)	0.201

=>

Implicirana kov. matrika 2

	<b>x1</b>	<b>x2</b>	<b>x3</b>	<b>x4</b>
<b>x1</b>	0.652			
<b>x2</b>	0.470	0.694		
<b>x3</b>	0.384	0.423	0.771	
<b>x4</b>	0.455	0.501	0.410	0.687



# Evalvacija (preverjanje) modela

- Gledamo odstopanja med dejanskimi in impliciranimi kovariančnimi matrikami
- $\chi^2$  test:
  - Testira, ali predpostavljena linearna struktura drži na populaciji
  - Če je  $\chi^2$  vrednost **statistično značilna** pri izbrani stopnji tveganja → **model zavrnemo**

# Evalvacija (preverjanje) modela

## ■ $\chi^2$ test (nadaljevanje):

- Stopinje prostosti: Če imamo  $k$  merjenih spremenljivk in linearni model s  $t$  neodvisnih parametrov, ki jih je potrebno oceniti, potem:

$$df = \frac{k(k + 1)}{2} - t$$

- Vendar: test je občutljiv za velike vzorce (in odstopanja od normalnosti)
- $\chi^2$  test razlike za gnezdene modele (ko je en model podmodel drugega)

# Evalvacija (preverjanje) modela

## Alternativni indeksi prileganja:

- Root Mean Squared Error of Approximation (RMSEA):

$$RMSEA = \sqrt{\frac{\chi^2 - df}{(N - 1) \times df}}$$

- Razpon: 0 - 1; sprejemljivo, če  $RMSEA < 0.05$  (kateri pravijo  $< 0.06$  - Hu & Bentler, 1999)
- Napaka „približka“ na populaciji – oceni, v kakšni meri se model razmeroma dobro prilega dejanskim vrednostmi na populaciji
- Občutljiv na število ocenjenih parametrov
- Lahko izračunamo intervale zaupanja

# Evalvacija (preverjanje) modela

Alternativni indeksi prileganja:

## ■ Comparative Fit Index (CFI):

$$CFI = 1 - \frac{\chi_{est}^2 - df_{est}}{\chi_{null}^2 - df_{null}}$$

- Primerja ocenjeni model glede na null model, kjer so spremenljivke neodvisne (vse korelacije 0)
- Razpon: 0 – 1
- Sprejemljivo če  $> .95$

# Evalvacija (preverjanje) modela

Alternativni indeksi prileganja:

## ■ Tucker-Lewis Index (TLI):

$$TLI = \frac{\chi_{null}^2/df_{null} - \chi_{est}^2/df_{est}}{\chi_{null}^2/df_{null} - 1}$$

- Podoben kot CFI, vendar kontrolira za kompleksnost modela (število parametrov)
- Razpon: 0 - ? (ni normiran, vrednosti nad 1 so redke)
- Sprejemljivo če  $> .95$

# Evalvacija (preverjanje) modela

Alternativni indeksi prileganja:

- Koren standardiziranega kvadriranega odklona - Standardized root mean squared residual (SRMR):
  - Poenostavljeno je to koren povprečne kvadrirane razlike med ocenjenimi in modelskimi korelacijami. Upošteva tudi variance.
  - Novejša različica upošteva tudi povprečja.
  - Sprejemljivo če  $\leq 0,08$ .

# Evalvacija (preverjanje) modela

- Prej omenjene mere so mere **globalnega prileganja** (prileganja modela kot celote)
- Relevantne so tudi mere lokalnega (ne)prileganja—Modification indices (MI):
  - Na voljo za vse omejene in fiksne parametre (tudi tiste, ki so 0 – jih ni v modelu)
  - Vrednost je pričakovano zmanjšanje  $\chi^2$  statistike, če se obravnavani parameter sprosti
  - $\chi^2$  testi z eno stopinjo prostosti
  - Preveriti je potrebno tudi pričakovano spremembo parametra - expected parameter change (EPC)!

# Omejitve in identifikacija

## ■ Tri vrste parametrov:

- ☐ Prosti parametri

- ☐ Fiksirani parametri (nastavljeni na določeno konstanto)

- ☐ Parametri z omejitvami (parametri, ki so nastavljeni, da se enaki nekemu drugemu parametru)

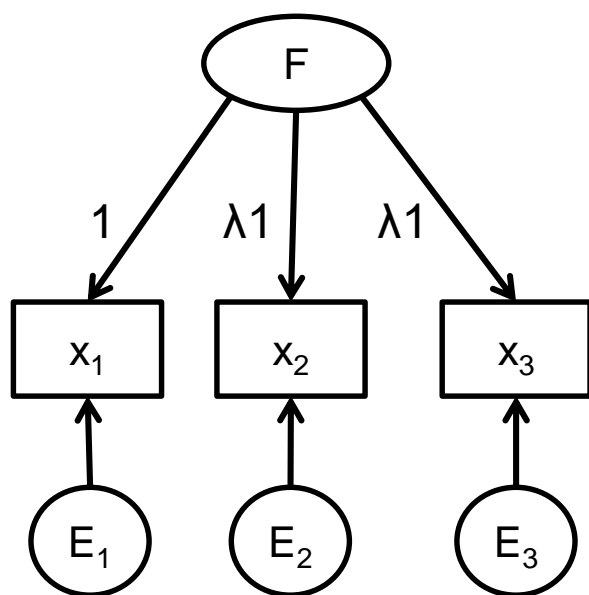
**Omejitve:** odražanje pozvanjanje problema (pred analizo podatkov)



# Omejitve in identifikacija

- Vloge omejitev:
  - Identifikacija modela
  - Statistične predpostavke modela
  - Predstavitev teoretičnih predpostavk

# Omejitve in identifikacija



## Primeri omejitev

- Identifikacija modela: prava utež pri faktorju = 1
- Statistične predpostavke:  $\text{Cov}(E_1, E_2) = 0$
- Teoretične predpostavke: utež 2 = utež 3

# Omejitve in identifikacija

- Ali je dovolj informacij, da lahko ocenimo vse proste parametre?
  - Je model indetifikabilen: Ali vzorčne variance in kovariance vključujejo dovolj informacij, da se oceni vse proste parametre
- Pogoji 1: Skritim (latentnim) spremenljivka je potrebno nastaviti lestvico...:
  - ...z nastavitvijo faktorske uteži pri enem indikatorju na 1 (i.e. „marker“ indikator) ALI
  - ...z nastavitvijo variance skrite (latentne) spremenljivke

# Omejitve in identifikacija

## ■ Pogoji 2: Statistična identifikacija:

- Število parametrov, ki jih ocenjujemo ( $t$ ), ne sme preseči števila informacij, ki so na voljo  $[k(k+1)/2]$
- $=$  število prostostnih stopenj ne sme biti negativno
- V smislu enačb: število neznank ne sme biti večje kot število enačb.

# Omejitve in identifikacija

- **Nezadostno identificirani modeli:**
  - Manj informacij kot ocenjenih parametrov;  $df < 0$
  - Neskončno število rešitev
  - Model je bolj kompleksen kot struktura podatkov
- **Ravno identificirani modeli:**
  - Toliko informacij kot ocenjenih parametrov;  $df = 0$
  - Unikatna rešitev, ki se popolnoma prilega
  - Model je tako kompleksen kot podatki
- **Nad identificirani modeli:**
  - Več informacij kot ocenjenih parametrov;  $df > 0$
  - Popolno prileganje ni možno
  - Model je manj kompleksen kot struktura podatkov

# Example 1 - CFA

- Podatki so bili zbrani v okviru raziskave Kakovost merjenja egocentričnih socialnih omrežij (Ferligoj in drugi, 2000) leta 2000. Vzorec vsebuje 1033 prebivalcev Ljubljane. Analiza je bila narejena na 631 prebivalcih, ki so bili osebno intervjuvani.
- Za konformatorno faktorsko analizo smo izbrali spremenljivke, ki merijo Ekstravertiranost in Emocionalno stabilnost (osebnostne lastnosti iz inventarja „Big Five“, International Personality Item Pool).

- Spremenljivke merijo dve dimenziji (zaradi omejitev prostora sem izbral podvzorec vseh spremenljivk).

## **Ekstrovertivnost**

EXTE Na zabavah se pomenkujem z mnogo ljudmi vseh vrst.

EXTLR Sem redkobeseden. \*

EXTNR V navzocnosti neznanih oseb sem molčeč. \*

EXTOR Imam malo povedati. \*

EXTRR Zadržujem se v ozadju.

## **Emocionalna stabilnost**

EMOCDR Zlahka me kaj vrže iz tira. \*

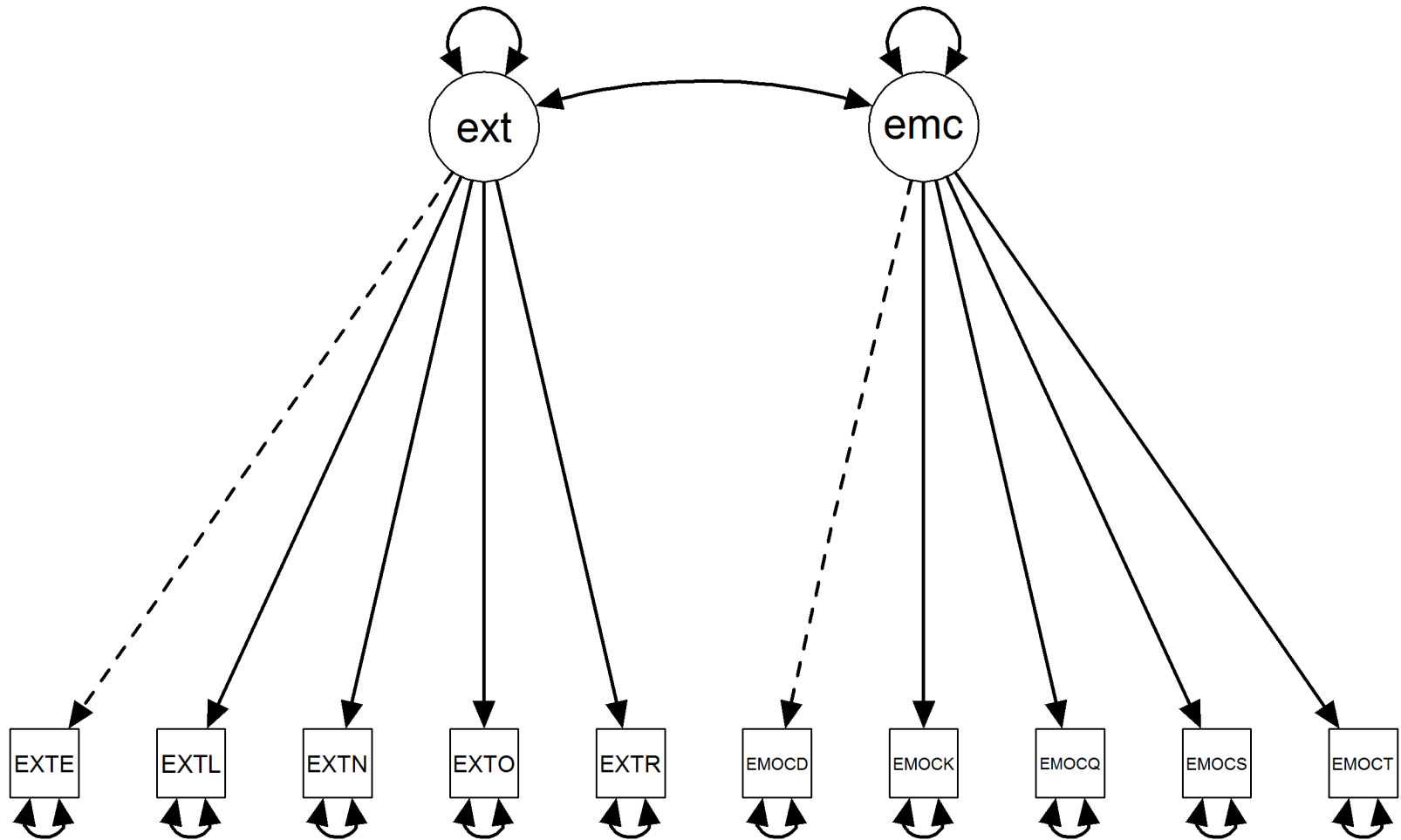
EMOCKR Sem zaskrbljene narave. \*

EMOCQR Moje razpoloženje se pogosto menja. \*

EMOCSR Pogosto sem potr. \*

EMOCTR Zlahka se me poloti napetost. \*

# Primer 1 - model





# Korelacije

	EXTE	EXTLR	EXTNR	EXTOR	EXTRR	EMOCDR	EMOCKR	EMOCQR	EMOCSR	EMOCTR
EXTE Na zabavah se pomenkujem z mnogo ljudmi vseh vrst.		,313	,334	,235	,285	,102	,106	,055	,143	,084
EXTLR Sem redkobeseden. *	,313		,365	,469	,408	-,013	,134	,041	,147	,063
EXTNR V navzočnosti neznanih oseb sem molč	,334	,365		,425	,459	,131	,250	,108	,160	,195
EXTOR Imam malo povedati. *	,235	,469	,425		,444	,084	,174	,121	,143	,139
EXTRR Zadržujem se v ozadju.	,285	,408	,459	,444		,142	,276	,164	,272	,215
EMOCDR Zlahka me kaj vrže iz tira. *	,102	-,013	,131	,084	,142		,294	,369	,337	,431
EMOCKR Sem zaskrbljene narave. *	,106	,134	,250	,174	,276	,294		,250	,350	,391
EMOCQR Moje razpoloženje se pogosto menja.	,055	,041	,108	,121	,164	,369	,250		,439	,442
EMOCSR Pogosto sem potr. *	,143	,147	,160	,143	,272	,337	,350	,439		,379
EMOCTR Zlahka se me poloti napetost. *	,084	,063	,195	,139	,215	,431	,391	,442	,379	

# Primer 1 – faktorske uteži

Latent Variables:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
ext =~						
EXTE	1.000				0.599	0.437
EXTLR	1.494	0.168	8.878	0.000	0.894	0.618
EXTNR	1.739	0.190	9.148	0.000	1.041	0.657
EXTOR	1.542	0.173	8.894	0.000	0.923	0.670
EXTRR	1.734	0.190	9.117	0.000	1.038	0.697
emoc =~						
EMOCDR	1.000				0.909	0.571
EMOCKR	0.955	0.097	9.796	0.000	0.868	0.545
EMOCQR	1.009	0.095	10.664	0.000	0.917	0.605
EMOCSR	0.879	0.083	10.593	0.000	0.799	0.626
EMOCTR	1.103	0.095	11.599	0.000	1.003	0.694

# Primer 1 – kovariance in variance

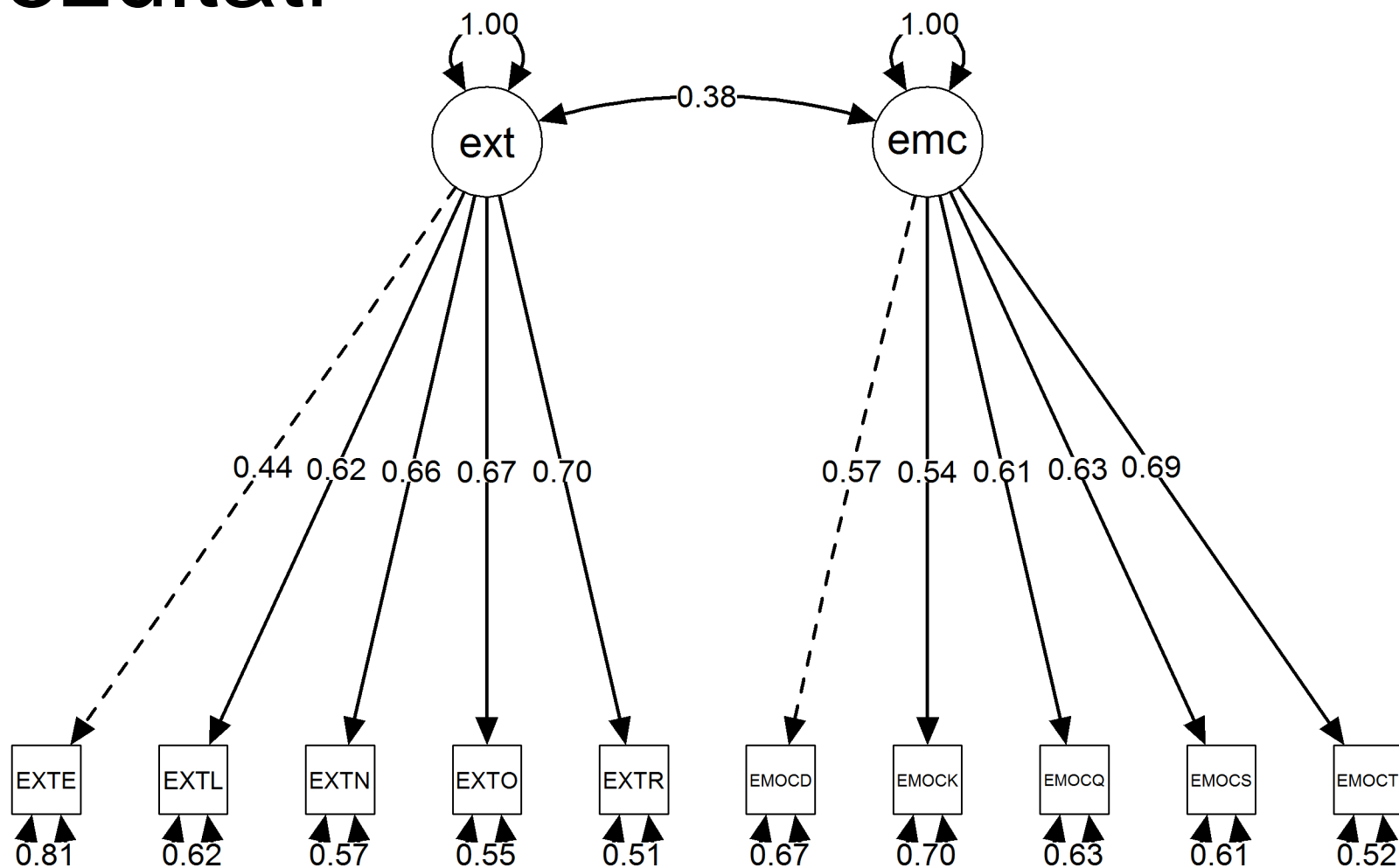
Covariances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
ext ~~						
emoc	0.206	0.037	5.506	0.000	0.379	0.379

Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EXTE	1.516	0.092	16.417	0.000	1.516	0.809
EXTLR	1.297	0.090	14.373	0.000	1.297	0.619
EXTNR	1.426	0.104	13.678	0.000	1.426	0.568
EXTOR	1.043	0.078	13.322	0.000	1.043	0.550
EXTRR	1.140	0.090	12.656	0.000	1.140	0.514
EMOCDR	1.711	0.114	14.980	0.000	1.711	0.674
EMOCKR	1.783	0.117	15.229	0.000	1.783	0.703
EMOCQR	1.453	0.101	14.354	0.000	1.453	0.633
EMOCSR	0.992	0.072	13.796	0.000	0.992	0.608
EMOCTR	1.081	0.089	12.217	0.000	1.081	0.518
ext	0.358	0.071	5.022	0.000	1.000	1.000
emoc	0.826	0.122	6.773	0.000	1.000	1.000

# Primer 1 – standardizirani rezultati



# Primer 1 – Ocenjevanje prileganja modela: dejanske korelacij in korelacije na podlagi modela

Dejanske	EXTE	EXTLR	EXTNR	EXTOR	EXTRR	EMOCDR	EMOCKR	EMOCQR	EMOCSR	EMOCTR
EXTE Na zabavah se pomenkujem z mnogo ljudmi vseh vrst.		,313	,334	,235	,285	,102	,106	,055	,143	,084
EXTLR Sem redkobeseden. *	,313		,365	,469	,408	-,013	,134	,041	,147	,063
EXTNR V navzocnosti neznanih oseb sem molčeč.	,334	,365		,425	,459	,131	,250	,108	,160	,195
EXTOR Imam malo povedati. *	,235	,469	,425		,444	,084	,174	,121	,143	,139
EXTRR Zadržujem se v ozadju.	,285	,408	,459	,444		,142	,276	,164	,272	,215
EMOCDR Zlahka me kaj vrže iz tira. *	,102	-,013	,131	,084	,142		,294	,369	,337	,431
EMOCKR Sem zaskrbljene narave. *	,106	,134	,250	,174	,276	,294		,250	,350	,391
EMOCQR Moje razpoloženje se pogosto menja. *	,055	,041	,108	,121	,164	,369	,250		,439	,442
EMOCSR Pogosto sem potr. *	,143	,147	,160	,143	,272	,337	,350	,439		,379
EMOCTR Zlahka se me poloti napetost. *	,084	,063	,195	,139	,215	,431	,391	,442	,379	
Na podlagi modela	EXTE	EXTLR	EXTNR	EXTOR	EXTRR	EMOCDR	EMOCKR	EMOCQR	EMOCSR	EMOCTR
EXTE Na zabavah se pomenkujem z mnogo ljudmi vseh vrst.		,270	,287	,293	,305	,095	,090	,100	,104	,115
EXTLR Sem redkobeseden. *	,270		,406	,414	,430	,133	,127	,142	,146	,162
EXTNR V navzocnosti neznanih oseb sem molčeč.	,287	,406		,441	,458	,142	,136	,151	,156	,173
EXTOR Imam malo povedati. *	,293	,414	,441		,467	,145	,138	,154	,159	,176
EXTRR Zadržujem se v ozadju.	,305	,430	,458	,467		,151	,144	,160	,165	,183
EMOCDR Zlahka me kaj vrže iz tira. *	,095	,133	,142	,145	,151		,311	,346	,357	,396
EMOCKR Sem zaskrbljene narave. *	,090	,127	,136	,138	,144	,311		,330	,341	,378
EMOCQR Moje razpoloženje se pogosto menja. *	,100	,142	,151	,154	,160	,346	,330		,379	,420
EMOCSR Pogosto sem potr. *	,104	,146	,156	,159	,165	,357	,341	,379		,434
EMOCTR Zlahka se me poloti napetost. *	,115	,162	,173	,176	,183	,396	,378	,420	,434	

# Primer 1 – Indeks prileganja

Number of observations	631
Number of missing patterns	8

Estimator	ML
Minimum Function Test Statistic	101.830
Degrees of freedom	34
P-value (Chi-square)	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	0.950
Tucker-Lewis Index (TLI)	0.934

Root Mean Square Error of Approximation:

RMSEA	0.056
90 Percent Confidence Interval	0.044 0.069
P-value RMSEA $\leq$ 0.05	0.194

Standardized Root Mean Square Residual:

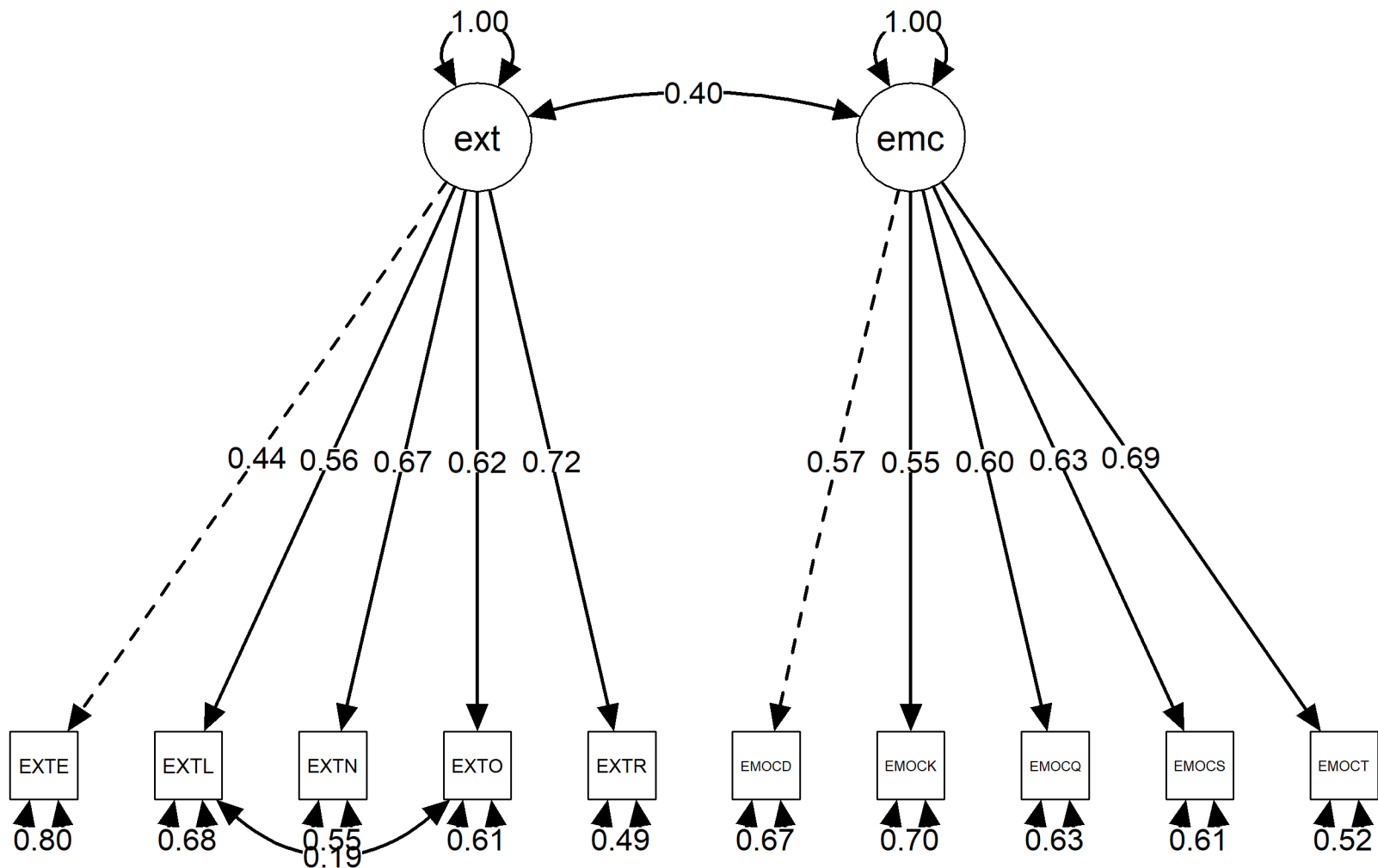
SRMR	0.045
------	-------

# Primer 1 – modification indices

	lhs	op	rhs	mi	epc	sepc.lv	sepc.all	sepc.nox
56	EXTLR	~~	EXTOR	14.775	0.257	0.257	0.129	0.129
85	EMOCKR	~~	EMOCQR	14.389	-0.310	-0.310	-0.128	-0.128
37	ext	=~	EMOCKR	13.147	0.448	0.268	0.168	0.168
42	emoc	=~	EXTLR	12.936	-0.259	-0.236	-0.163	-0.163
45	emoc	=~	EXTRR	12.915	0.261	0.237	0.159	0.159
90	EMOCSR	~~	EMOCTR	11.906	-0.232	-0.232	-0.126	-0.126
88	EMOCQR	~~	EMOCSR	7.400	0.181	0.181	0.094	0.094
58	EXTLR	~~	EMOCDR	7.121	-0.182	-0.182	-0.079	-0.079
48	EXTE	~~	EXTOR	7.120	-0.166	-0.166	-0.088	-0.088
79	EXTRR	~~	EMOCSR	6.421	0.132	0.132	0.069	0.069
38	ext	=~	EMOCQR	5.436	-0.269	-0.161	-0.106	-0.106
66	EXTNR	~~	EMOCKR	4.868	0.163	0.163	0.065	0.065
77	EXTRR	~~	EMOCKR	4.839	0.149	0.149	0.063	0.063
55	EXTLR	~~	EXTNR	4.641	-0.165	-0.165	-0.072	-0.072
39	ext	=~	EMOCSR	4.413	0.203	0.122	0.095	0.095
47	EXTE	~~	EXTNR	4.306	0.149	0.149	0.069	0.069
36	ext	=~	EMOCDR	4.289	-0.254	-0.152	-0.095	-0.095
84	EMOCDR	~~	EMOCTR	3.708	0.155	0.155	0.068	0.068

Vsebinsko  
smiselno,  
zato  
dodamo v  
model

# Primer 1 – standardizirani rezultati po popravku





# Primer 1 – Indeks prileganja popopravku

Estimator	ML	
Minimum Function Test Statistic	87.966	
Degrees of freedom	33	
P-value (Chi-square)	0.000	
User model versus baseline model:		
Comparative Fit Index (CFI)	0.960	
Tucker-Lewis Index (TLI)	0.945	
Root Mean Square Error of Approximation:		
RMSEA	0.051	
90 Percent Confidence Interval	0.039	0.065
P-value RMSEA <= 0.05	0.409	
Standardized Root Mean Square Residual:		
SRMR	0.042	

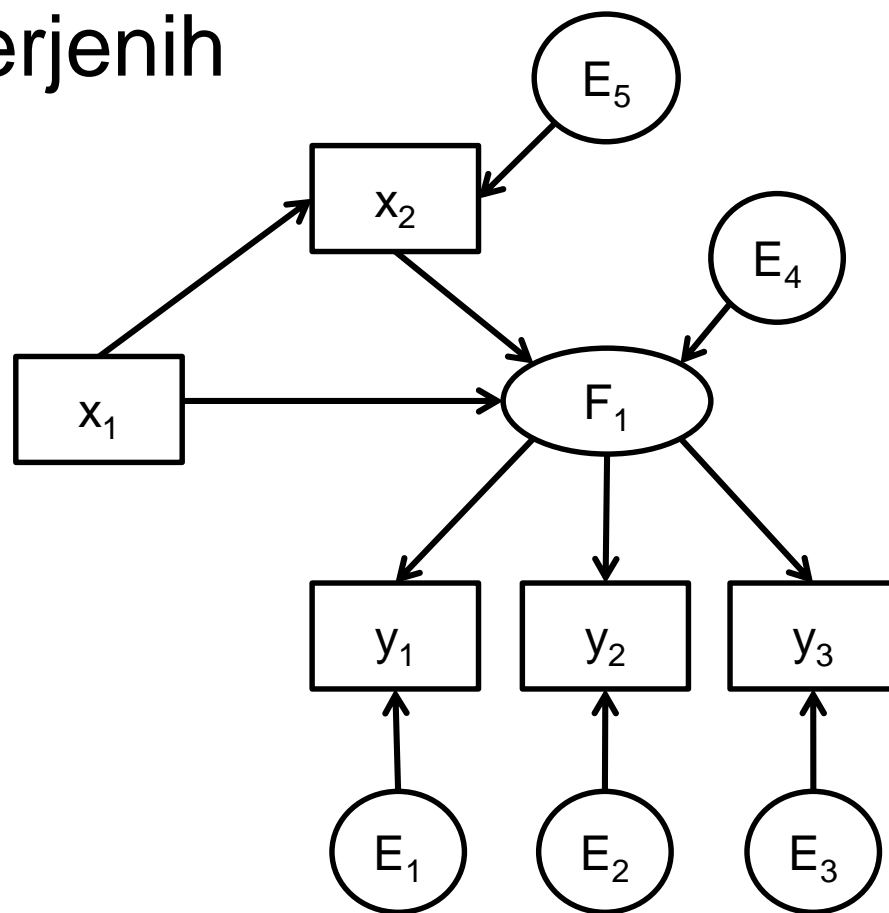
# Primer 1 – modification indices po popravku

	lhs	op	rhs	mi	epc	sepc.lv	sepc.all	sepc.nox
85	EMOCKR	~~	EMOCQR	14.453	-0.310	-0.310	-0.128	-0.128
38	ext	=~	EMOCKR	13.875	0.469	0.284	0.178	0.178
90	EMOCSR	~~	EMOCTR	11.785	-0.229	-0.229	-0.124	-0.124
46	emoc	=~	EXTRR	8.869	0.229	0.208	0.140	0.140
43	emoc	=~	EXTLR	8.801	-0.214	-0.195	-0.134	-0.134
47	EXTE	~~	EXTLR	7.717	0.180	0.180	0.091	0.091
88	EMOCQR	~~	EMOCSR	7.496	0.182	0.182	0.094	0.094
58	EXTLR	~~	EMOCDR	6.369	-0.170	-0.170	-0.074	-0.074
79	EXTRR	~~	EMOCSR	6.226	0.130	0.130	0.068	0.068
39	ext	=~	EMOCQR	5.870	-0.285	-0.172	-0.114	-0.114
49	EXTE	~~	EXTOR	4.851	-0.134	-0.134	-0.071	-0.071
66	EXTNR	~~	EMOCKR	4.263	0.152	0.152	0.060	0.060
40	ext	=~	EMOCSR	4.192	0.202	0.122	0.096	0.096
77	EXTRR	~~	EMOCKR	4.168	0.138	0.138	0.058	0.058
37	ext	=~	EMOCDR	3.983	-0.249	-0.151	-0.095	-0.095

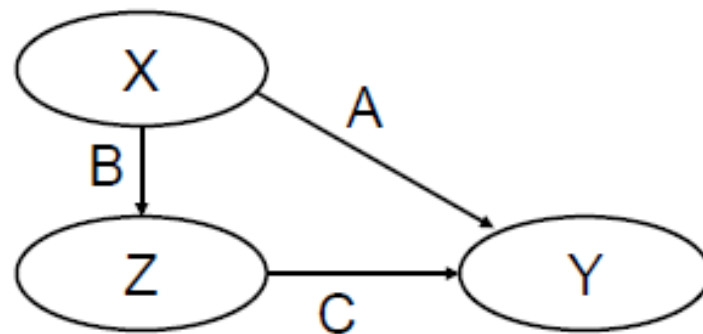
Sicer  
precej  
veliki, a  
težko  
najdemo  
utemeljitev

# Poln SEM model

- Kakršna koli kombinacija skritih faktorjev in indikatorjev (merjenih spremenljivk)
- Strukturne „poti“



# Poln SEM model



učinka učinek spremenljivke X na Y poteka (tudi) preko spremenljivke Z:

- Neposredni učinek:  $A$
- Posredni učinek:  $B * C$
- Skupni učinek:  $A + B * C$
- Polna vs. delna mediacija (samo posredni učinek ali posredni in neposredni)

# Ocenjevanje

- Klasično se za ocenjevanje parametrov uporablja Metoda največjega verjetja (ML – maximum likelihood)
- GLS (Generalized least squares) Asimptotično enakovredna ML. Temelji na istih predpostavkah (več o tem kasneje). Običajno se raje uporablja ML. Ne bomo gledali podrobneje.
- WLS – (Weighted least squares) Metoda uteženih (ali navadnih) najmanjših kvadartov.

# Ocenjevanje – ML (maximum likelihood)

- Najbolj klasična metoda
- Predpostavlja multivariatno normalno porazdelitev.
- Ta omogoča tudi uporabo FIML (Full information maximum likelihood) pristopa za obravnavnavo manjkajočih vrednosti
- Obstaja več „robustnih“ možnosti za oceno standardnih napak in testnih statistik.

# Ocenjevanje - ML

Vse ML različice (tudi robustne) optimizirajo:

$$F_{ML} = \ln|\Sigma(\hat{\Theta})| + tr\left(\mathbf{S}\Sigma^{-1}(\hat{\Theta})\right) - \ln(\mathbf{S}) - m$$

kjer je

- $m$  število spremenljivk
- $\mathbf{S}$  empirična kovariančna matrika. Običajno se SS/SCP deli z  $n$  in ne z  $n - 1$  (v Lavaanu delimo z  $n - 1$ , če zahtevamo `likelihood = "wishart"`).
- $\Sigma(\hat{\Theta})$  kovariančna matrika na podlagi modela
- $\Theta$  vektor parametrov.

# Ocenjevanje – ML - robustne

Robustne različice drugače ocenjujejo standardne napake in nekatere tudi testne statistike. Lavaan podpira (originalni opisi):

- "MLM": (RSE) robust standard errors and a Satorra-Bentler scaled test statistic. For complete data only (FCDO).
- "MLMVS": RSE and a mean- and variance adjusted test statistic (aka the Satterthwaite approach). FCDO.
- "MLMV": RSE and a mean- and variance adjusted test statistic (using a scale-shifted approach). FCDO.
- "MLF": RSE based on the first-order derivatives, and a conventional test statistic. For both complete and incomplete data.
- "MLR": RSE (Huber-White) and a scaled test statistic that is (asymptotically) equal to the Yuan-Bentler test statistic. For both complete and incomplete data.

Različne različice standardnih napak in testnih statistik se da dobiti tudi z izbirami parametrov `se` in `test`.



# Ocenjevanje - WLS

Drugi pristop temelji na metodi uteženih najmanjših kvadratov. Optimiziramo:

$$F_{WLS} = [s - \sigma(\hat{\Theta})]' \mathbf{W}^{-1} [s - \sigma(\hat{\Theta})]$$

Metode se razlikujejo glede na to, kako določimo matriko uteži  $\mathbf{W}$ .

$s$  je vektor unikatnih elementov iz  $\mathbf{S}$

$\sigma(\hat{\Theta})$  je vektor unikatnih elementov iz  $\Sigma(\hat{\Theta})$

# Ocenjevanje - WLS

Metode:

- WLS/ADF (klasično):  $\mathbf{W}$  je običajno asimptotična kovariacijska matrika med elementi kovariančne matrike. Načeloma lahko izberemo tudi kakšno drugo matriko.
- DWLS: Podobno kot zgoraj, le da so vsi izven-diagonalni elementi 0
- ULS:  $\mathbf{W}$  je enotska matrika -  $\mathbf{I}$ . V bistvu je

$$F_{ULS} = \sum_{i \leq j} [\mathbf{s} - \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})]_{ij}^2$$

# Ocenjevanje – WLS

- Te metode se še posebej priporočajo oz. uporabljajo pri ordinarnih in binarnih sprem.
- V teh primerih se namesto navadnih korelacij/kovarianc priporoča uporabo „polychoric“ korelacij (več o tem pri ordinalnih indikatorjih)
- Posledično mnogi viri pri opisu teh metod predpostavljajo te korelacije, čeprav to ni nujno (`lavaan` lahko uporablja ene ali druge)

# Ocenjevanje – WLS - robustno

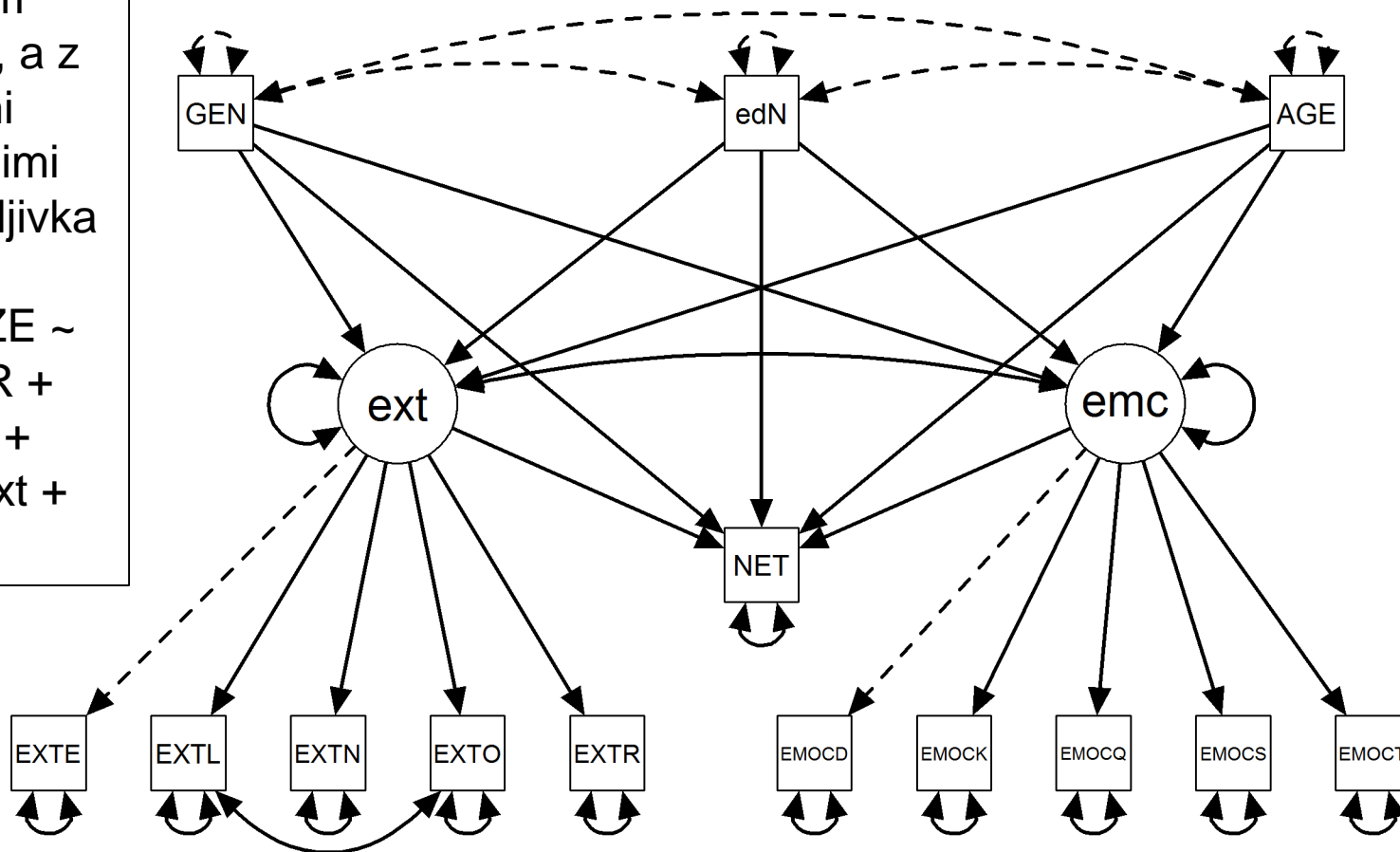
Za cenilki DWLS in ULS, `lavaan` ponuja tudi 'robustne' različice: WLSM, WLSMVS, WLSMV, ULSM, ULSMVS, ULSMV. Pri robustnih različicah WLS, se uporabi diagonalna matrika uteži pri ocenjevanju, polna matrika uteži pa pri popravkih standardnih napak in pri računanju testne statistike.

Različne različice standardnih napak in testnih statistik se da dobiti tudi z izbirami parametrov `se in test`.

# Polni SEM primer – primer 2

Predvsem  
regresija, a z  
latentnimi  
neodvisnimi  
spremenljivka  
mi:

$\text{NET\_SIZE} \sim$   
 $\text{GENDER} +$   
 $\text{eduNum} +$   
 $\text{AGE} + \text{ext} +$   
 $\text{emc}$



# Primer 2 - ocene

Latent Variables:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
<b>ext =~</b>						
EXTE	1.000				0.586	0.428
EXTLR	1.386	0.166	8.364	0.000	0.813	0.561
EXTNR	1.763	0.197	8.940	0.000	1.034	0.652
EXTOR	1.474	0.172	8.552	0.000	0.864	0.628
EXTRR	1.881	0.213	8.849	0.000	1.103	0.741
<b>emoc =~</b>						
EMOCDR	1.000				0.914	0.574
EMOCKR	0.954	0.096	9.918	0.000	0.872	0.548
EMOCQR	1.005	0.093	10.780	0.000	0.919	0.607
EMOCSR	0.884	0.082	10.787	0.000	0.808	0.633
EMOCTR	1.077	0.092	11.684	0.000	0.985	0.682

# Primer 2 - ocene

GENDER:

1- moški

2 - ženske

Regressions:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
ext ~						
GENDER	0.084	0.052	1.595	0.111	0.143	0.070
eduNum	0.071	0.020	3.570	0.000	0.121	0.162
AGE	-0.011	0.002	-6.675	0.000	-0.019	-0.366
emoc ~						
GENDER	-0.247	0.084	-2.953	0.003	-0.270	-0.133
eduNum	0.201	0.033	6.172	0.000	0.220	0.294
AGE	-0.005	0.002	-2.158	0.031	-0.005	-0.097
NET_SIZE ~						
GENDER	0.541	0.249	2.176	0.030	0.541	0.089
eduNum	0.169	0.095	1.783	0.075	0.169	0.075
AGE	0.003	0.007	0.494	0.621	0.003	0.022
ext	0.545	0.300	1.814	0.070	0.319	0.106
emoc	-0.182	0.185	-0.983	0.326	-0.166	-0.055

Če bi uporabili vse ext/emoc spremenljivke, se ta *p*-vrednost zmanjša na 0.013.

# Primer 2 - ocene

## Covariances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EXTLR ~~ EXTOR	0.246	0.068	3.627	0.000	0.246	0.192
ext ~~ emoc	0.182	0.034	5.417	0.000	0.392	0.392

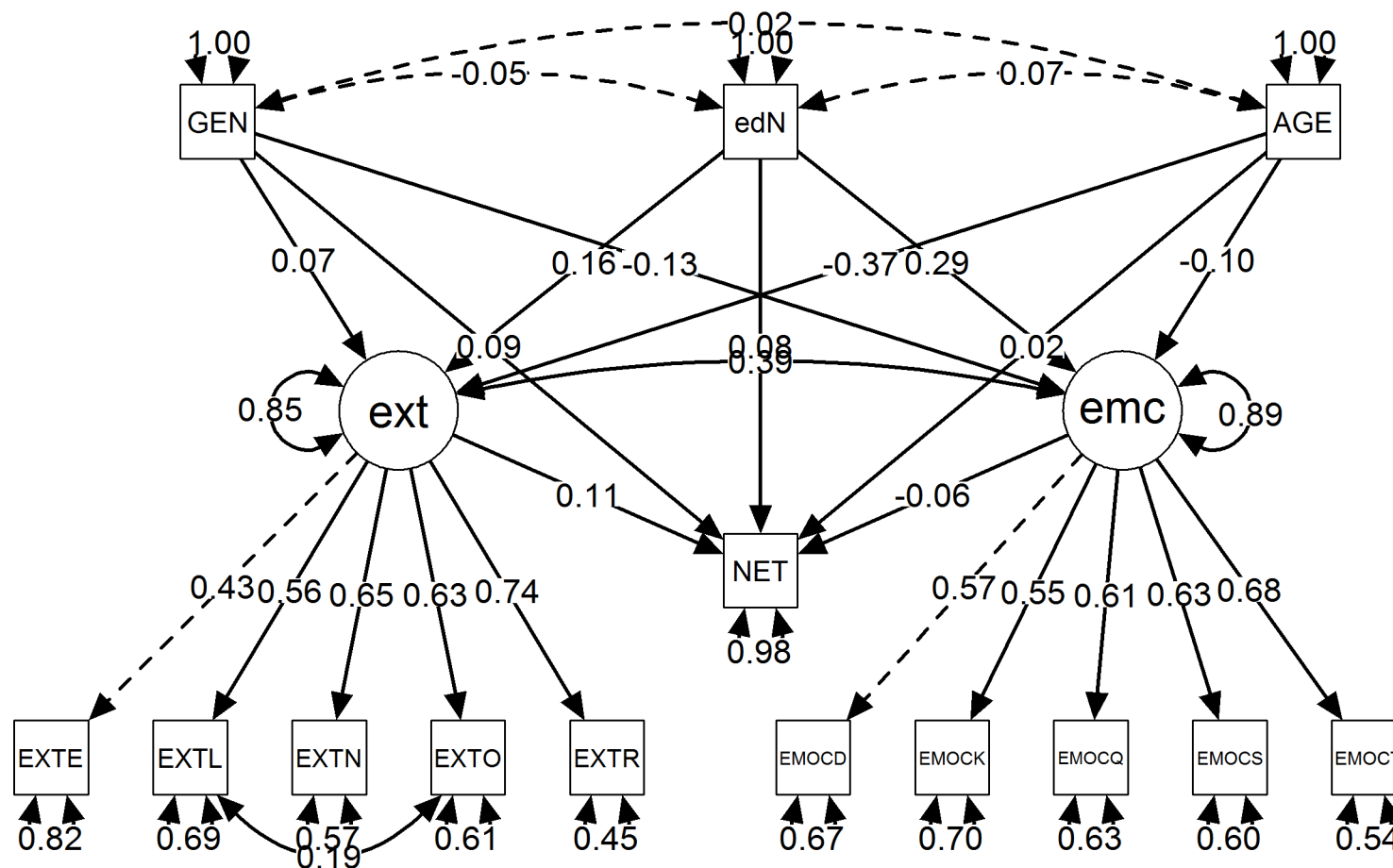
## Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EXTE	1.531	0.093	16.422	0.000	1.531	0.817
EXTLR	1.437	0.098	14.730	0.000	1.437	0.685
EXTNR	1.442	0.106	13.547	0.000	1.442	0.574
EXTOR	1.148	0.083	13.824	0.000	1.148	0.606
EXTRR	1.000	0.092	10.822	0.000	1.000	0.451
EMOCDR	1.701	0.113	15.024	0.000	1.701	0.671
EMOCKR	1.776	0.116	15.263	0.000	1.776	0.700
EMOCQR	1.450	0.101	14.380	0.000	1.450	0.632
EMOCSR	0.978	0.070	13.867	0.000	0.978	0.600
EMOCTR	1.116	0.087	12.773	0.000	1.116	0.535
NET_SIZE	8.861	0.501	17.677	0.000	8.861	0.976
ext	0.291	0.061	4.800	0.000	0.846	0.846
emoc	0.740	0.109	6.791	0.000	0.886	0.886



# Primer 2 - ocene

GENDER:  
1- moški  
2 - ženske



# Primer 2 – prileganje modela

lavaan (0.5-20) converged normally after 95 iterations

Number of observations	631
------------------------	-----

Number of missing patterns	9
----------------------------	---

Estimator	ML
-----------	----

Minimum Function Test Statistic	198.274
---------------------------------	---------

Degrees of freedom	65
--------------------	----

P-value (Chi-square)	0.000
----------------------	-------

User model versus baseline model:

Comparative Fit Index (CFI)	0.915
-----------------------------	-------

Tucker-Lewis Index (TLI)	0.885
--------------------------	-------

Root Mean Square Error of Approximation:

RMSEA	0.057
-------	-------

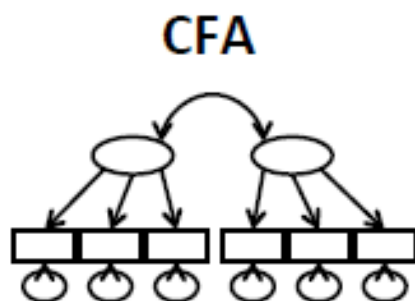
90 Percent Confidence Interval	0.048	0.066
--------------------------------	-------	-------

P-value RMSEA $\leq$ 0.05	0.097
---------------------------	-------

# SEM za več skupin

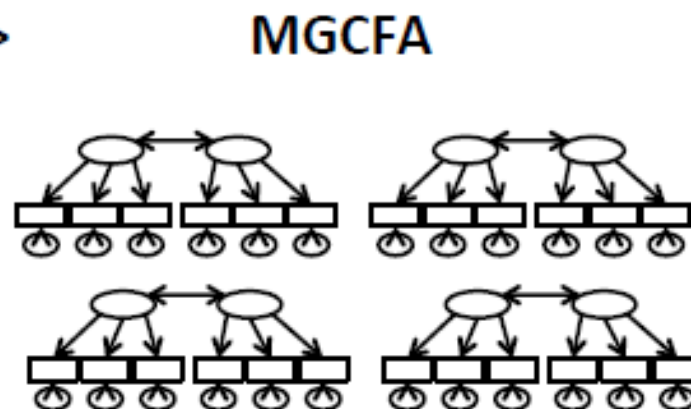
Multiple Group SEM:

- Ocenimo podobne modele za različne skupine – različne modele za skupine ocenimo sočasno



$$\mathbf{X} = \boldsymbol{\lambda}\boldsymbol{\xi} + \boldsymbol{\delta}$$

=>

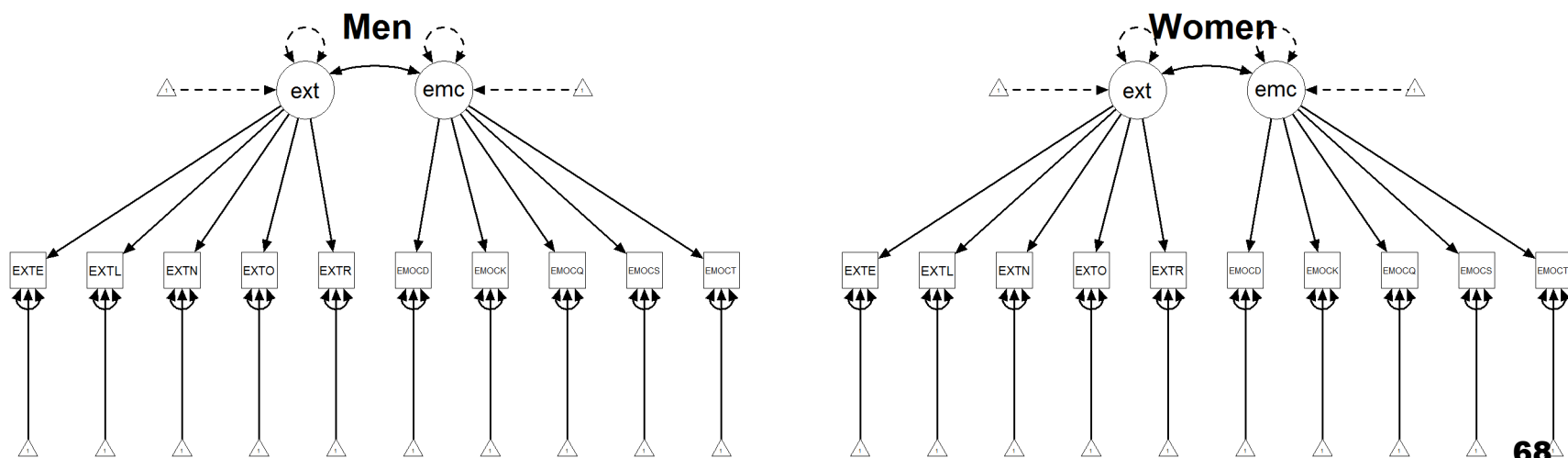


$$\mathbf{X}^g = \boldsymbol{\lambda}^g \boldsymbol{\xi}^g + \boldsymbol{\delta}^g$$

# SEM za več skupin

Prednost: Možnost testiranja, do kakšne mere so si modeli med različnimi skupinami podobni:

- → Ali sta ekstrovertiranost in emocionalna stabilnost enako močno korelirane pri moških in ženskah?



# Merska enakovrednost

- Predno lahko primerjamo vrednosti med skupinami, se moramo prepričati, da so meritve primerljive
- Meritve so lahko neprimerljive ker...
  - ...primerjamo različne koncepte (npr. višina proti teži)
  - ...uporabimo različno mersko lestvico (npr. težav kg proti teži v funtih)

# Merska enakovrednost

- Koncept merske enakovrednost obravnava sledeč problem:
  - ali pri različnih pogojih opazovanja in raziskovanja meritve še vedno merijo isto lastnost (Horn & McArdle 1992)
- Ne-enakovrednost nastopi, kadar razlike v meritvah ne odražajo dejanskih razlik

# Merska enakovrednost

Možni „viri“ merske ne-enakovrednosti (van de Vijver 1998):

- Pristranskost konstrukta – nekateri konstrukti so specifični za določene skupine
- Pristranskost metod – lastnosti skupin lahko vplivajo na tip odgovarjanja: Npr. nagnjenost k strinjanju, nagnjenost k ekstremom
- Pristranskost indikatorjev
  - Prevodi vprašalnikov
  - Specifičen pomen indikatorjev v določenih kontekstih

# Merska enakovrednost

Specifičen pomen indikatorjev v določenih kontekstih

- Npr.: merjenje vernosti v evropski družboslovni raziskavi (European Social Survey – ESS):
  - SUBJEKTIVNA VERNOST: C13: Ne glede na to, ali pripadate kateri od religij ali ne, prosimo ocenite, koliko ste verni? (0 = sploh nisem veren -- 10 = zelo sem veren)
  - OBISK: C14: Če odštejete posebne priložnosti kot so poroke ali pogrebi, kako pogosto obiskujete verske obrede? (1 = vsak dan -- 7 = nikoli)
  - MOLITEV: C15: Če odštejete udeležbo pri verskih obredih, kako pogosto sicer molite, če sploh molite? (1 = vsak dan -- 7 = nikoli)
- Kateri indikator drugače funkcionira v Turčiji?

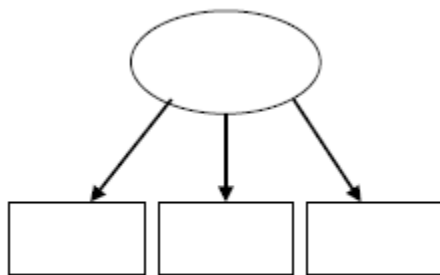


# Testiranje merske enakovrednosti

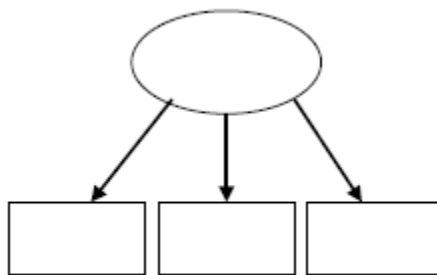
- Merska enakovrednost se lahko testira preko MGCFA
- Obstaja več hierarhičnih nivojev merske enakovrednosti (Steenkamp & Baumgartner 1998)

## 1. Konfiguralna enakovrednost:

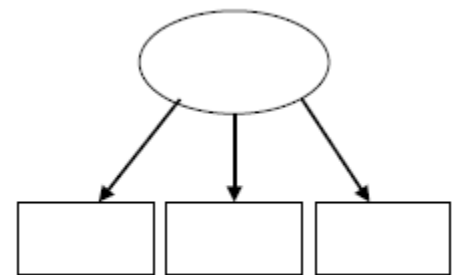
- Identične faktorske strukture med skupinami, a nobenih omejitev, da morajo biti parametri enaki
- Enaki koncepti so merjeni v vseh skupina
- Vendar: vrednosti niso primerljive!



Skupina 1



Skupina 2



Skupina K

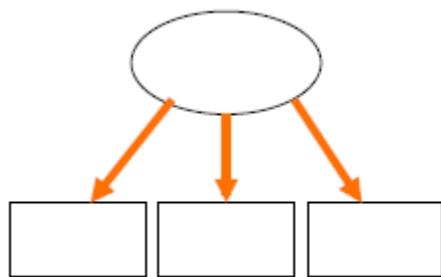
# Testiranje merske enakovrednosti

## 2. Metrična enakovrednost:

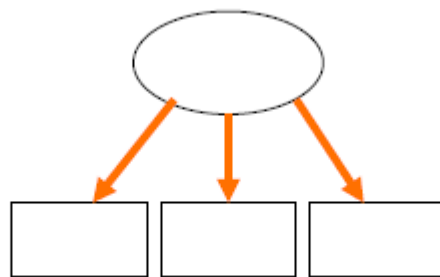
- Enaki „nakloni“ (reg. koeficienti) za vse skupine:

$$\lambda_j^1 = \dots = \lambda_j^G \text{ with } j = 1 \dots J$$

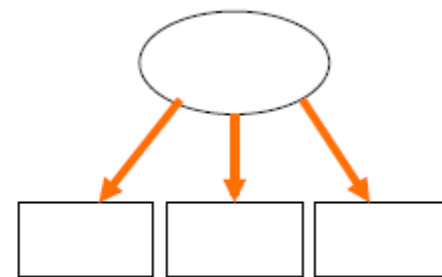
- Povečanje latentne spremenljivke za eno enoto ima enak povem v vseh skupinah
- Statistike, ki temeljijo na razlikah (in ne abs. vrednostih), se lahko primerjajo med skupinami (npr. regresijski koeficienti, kovariacne, korelacije)



Skupina 1



Skupina 2



Skupina K

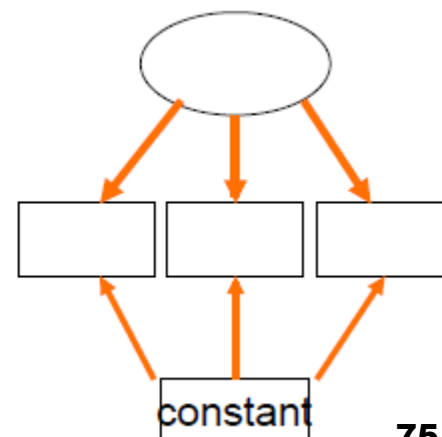
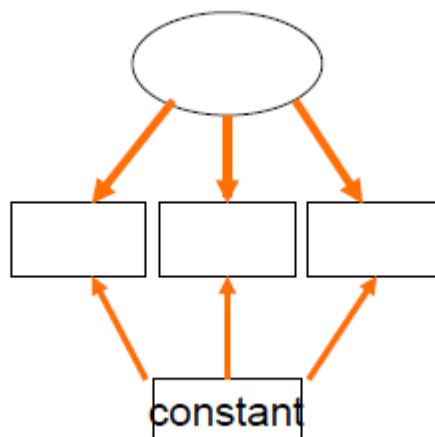
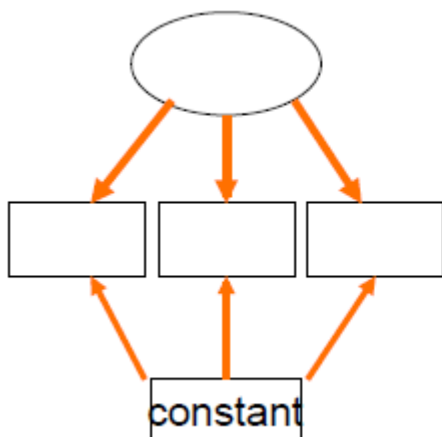
# Testiranje merske enakovrednosti

## 3. Skalarna enakovrednost :

- Enaki nakloni (reg. koeficienti in konstante:

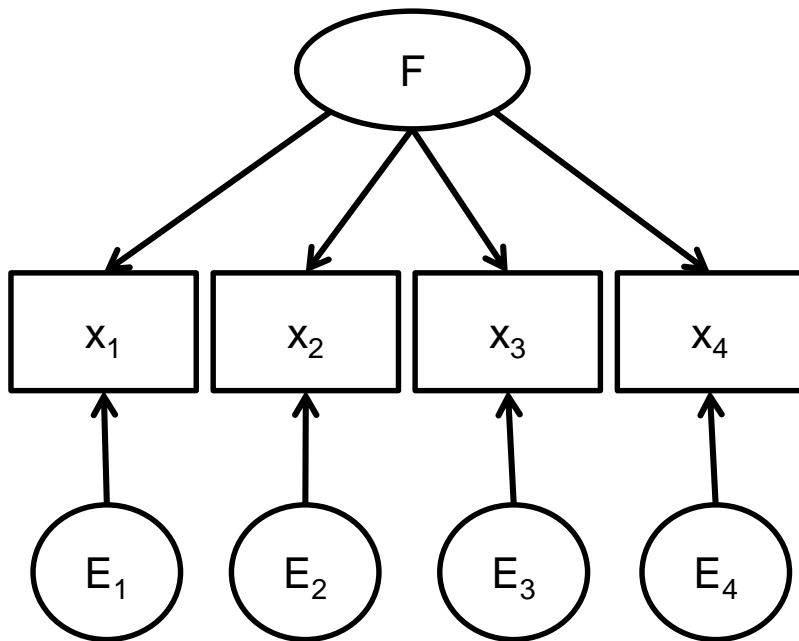
$$\begin{bmatrix} \tau_j^1 \\ \lambda_j^1 \end{bmatrix} = \dots = \begin{bmatrix} \tau_j^G \\ \lambda_j^G \end{bmatrix}$$

- Enote iz različnih skupin bodo pri enakih vrednostih latentnih spremenljivk dale enake odgovore (+ napaka)
- Polna primerljivost vseh vrednosti: primerjave povprečij so možne



# Povprečja

- Za primerjave povprečij, rabimo skalarno enakovrednost
- Uvod v uporabo povprečij v SEM



$$X_1 = \tau_1 + \lambda_{11}F_1 + \epsilon_1$$

$$X_2 = \tau_2 + \lambda_{21}F_1 + \epsilon_2$$

$$X_3 = \tau_3 + \lambda_{31}F_1 + \epsilon_3$$

$$X_4 = \tau_4 + \lambda_{41}F_1 + \epsilon_4$$

$$\mathbf{X} = \boldsymbol{\tau} + \boldsymbol{\lambda}\mathbf{F} + \boldsymbol{\epsilon}$$

# Povprečja

- Novi parametri v modelu:
  - Konstante pri indikatorjih ( $\tau$ ): Napovedna vrednost za  $x$ , če je vrednost latentne spremenljivke 0
  - Povprečja latentnih spremenljiv ( $\kappa$ )
- Novi parametri se lahko uporabijo za izračun povprečne vrednosti indikatorjev:

$$\overline{X_1} = \tau_1 + \lambda \kappa$$

=> Pri ocenjevanju uporabimo nove informacije: dejanska povprečja indikatorjev!

# Povprečja

- Pri CFA za eno skupino, povprečja niso informativna. Zakaj?
  - Problem z identifikacijo: V model z  $n$  indicatorji in  $m$  faktorji moramo vključiti  $n + m$  novih parametrov, pridobimo pa  $n$  novih informacij
  - Potrebujemo dodatne omejitve:  $\kappa_i = 0$
  - Posledica: Povprečja so ravno identificirana → konstante pri indikatorjih = dejanska povprečja
- MGCFA omogoča dodajanje omejitev, ki se nanašajo na več skupin in s tem omogoča rešitev problema z identifikacijo (skalarna enakovrednost je potrebna za primerjavo povprečij)

# Povprečja

- Več možnosti za rešitev problema z identifikacijo pri MGCFA. Najbolj enostavna:
  - Ne računamo dejanskih povprečij latentnih spremenljivk, ampak le relativno glede na referenčno skupino
  - Pri vsaki latentni spremenljivki nastavimo njeno povprečje v prvi skupini na 0; povprečja latenten spremenljivke v drugih skupinah so prosta
  - Za vsaj en indikator pri vsaki latentni spremenljivki omejimo konstante tako, da so pri vseh skupinah enake kot pri prvi skupini (še raje pa več, da zagotovimo skalarno enakovrednost)
- Druge metode za doseganje identifikabilnosti: Little, Slegers & Card (2006) v *Structural equation modeling*, str. 59-72.

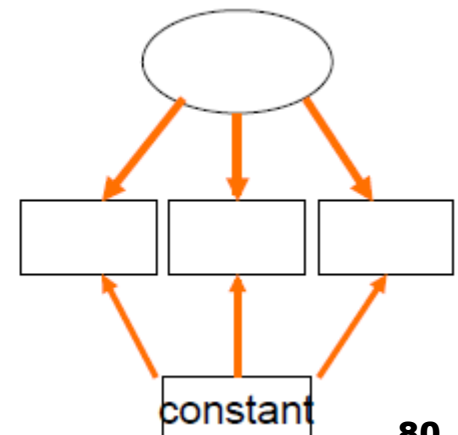
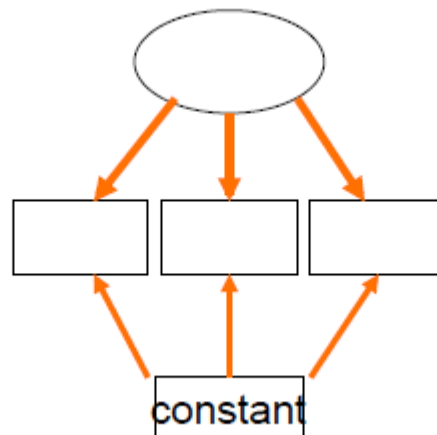
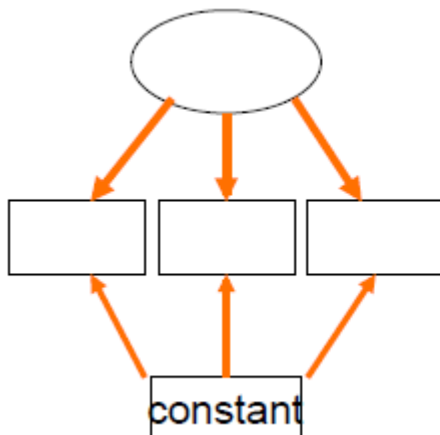
# Testiranje merske enakovrednosti

## 3. Skalarna enakovrednost :

- Enaki nakloni (reg. koeficienti in konstante:

$$\begin{bmatrix} \tau_j^1 \\ \lambda_j^1 \end{bmatrix} = \dots = \begin{bmatrix} \tau_j^G \\ \lambda_j^G \end{bmatrix}$$

- Enote iz različnih skupin bodo pri enakih vrednostih latentnih spremenljivk dale enake odgovore (+ napaka)
- Polna primerljivost vseh vrednosti: primerjave povprečij so možne





# Testiranje merske enakovrednosti

- Delna enakovrednost: primerjave so smiselne, če so parametri za vsaj dva indikatorja (pri obravnavani latentni spremenljivki) omejeni tako, da so enaki med skupinami (referenca Byrne, Shavelson in Muthen, 1989; Vendar se vsi ne strinjajo: De Beuckelaer in Swinnen, 2011)

# Testiranje merske enakovrednosti

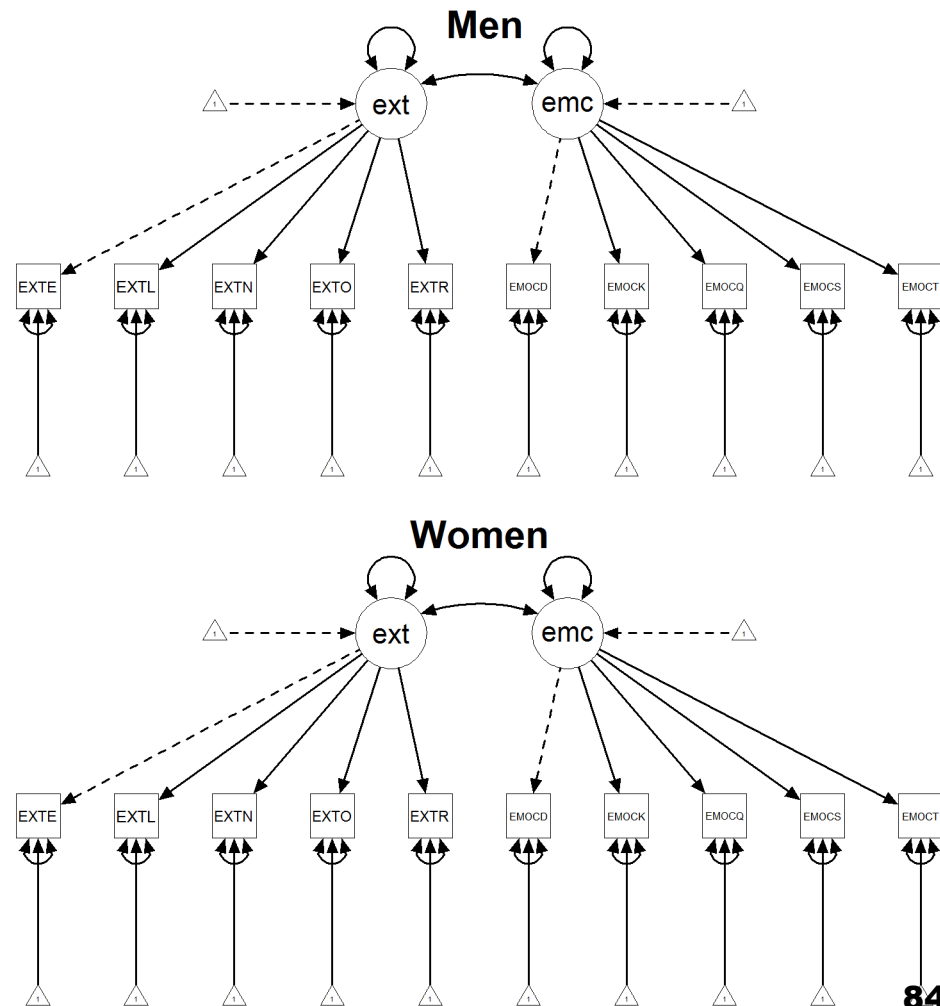
- Nerešen problem: Kako testirati omejitve med skupinami v praksi? Kako se odločiti, kdaj je omejitev kršena?
- Možni pristopi:
  1. **Strogo konformatorni:** Uporabimo omejitve (enakovrednosti); ocenimo splošno prileganje modela:
    - $\chi^2$ -test, RMSEA, CFI, TLI
    - Vendar: Splošni indeksi pogosto niso občutljivi na lokalne „napake“ v specifikaciji

# Testiranje merske enakovrednosti

2. **Alternativni pristop:** Ocenimo več modelov (z (vsemi) omejitvami, brez omejitev, z le nekaterimi omejitvami); primerjaj splošno prileganje modelov
  - $\chi^2$  test za primerjavo modelov (vendar: občutljiv pri velikih vzorcih in pri nenormalnih spremenljivkah)
  - Razlike v alternativnih indeksih prileganja:  $\Delta CFI$ ;  $\Delta RMSEA$  (Cheung in Rensvold 2002; Chen 2007); vendar: mejne vrednosti so arbitrarne ali temeljijo na omejenih simulacijskih študijah
3. **Pristop gradnje modelov:** Oceni model z vsemi omejitvami, identificiraj „lokalna“ neujezanja in popravi model v skladu z njimi
  - Modification indices
  - Vendar: občutljivost testa glede na velikost vzorca

# Primer 3: Testiranje merske enakovrednosti

- Ali lahko primerjamo moške in ženske?



# Primer 3: Testiranje merske enakovrednosti

## Chi Square Difference Test

	Df	AIC	BIC	Chisq	Chisq diff	Df diff	Pr(>Chisq)	
fit.configural	68	21390	21666	150.30				
fit.loadings	76	21394	21634	169.61	19.312	8	0.0132785	*
fit.intercepts	84	21410	21614	201.78	32.167	8	8.693e-05	***
fit.means	86	21424	21619	219.46	17.685	2	0.0001445	***

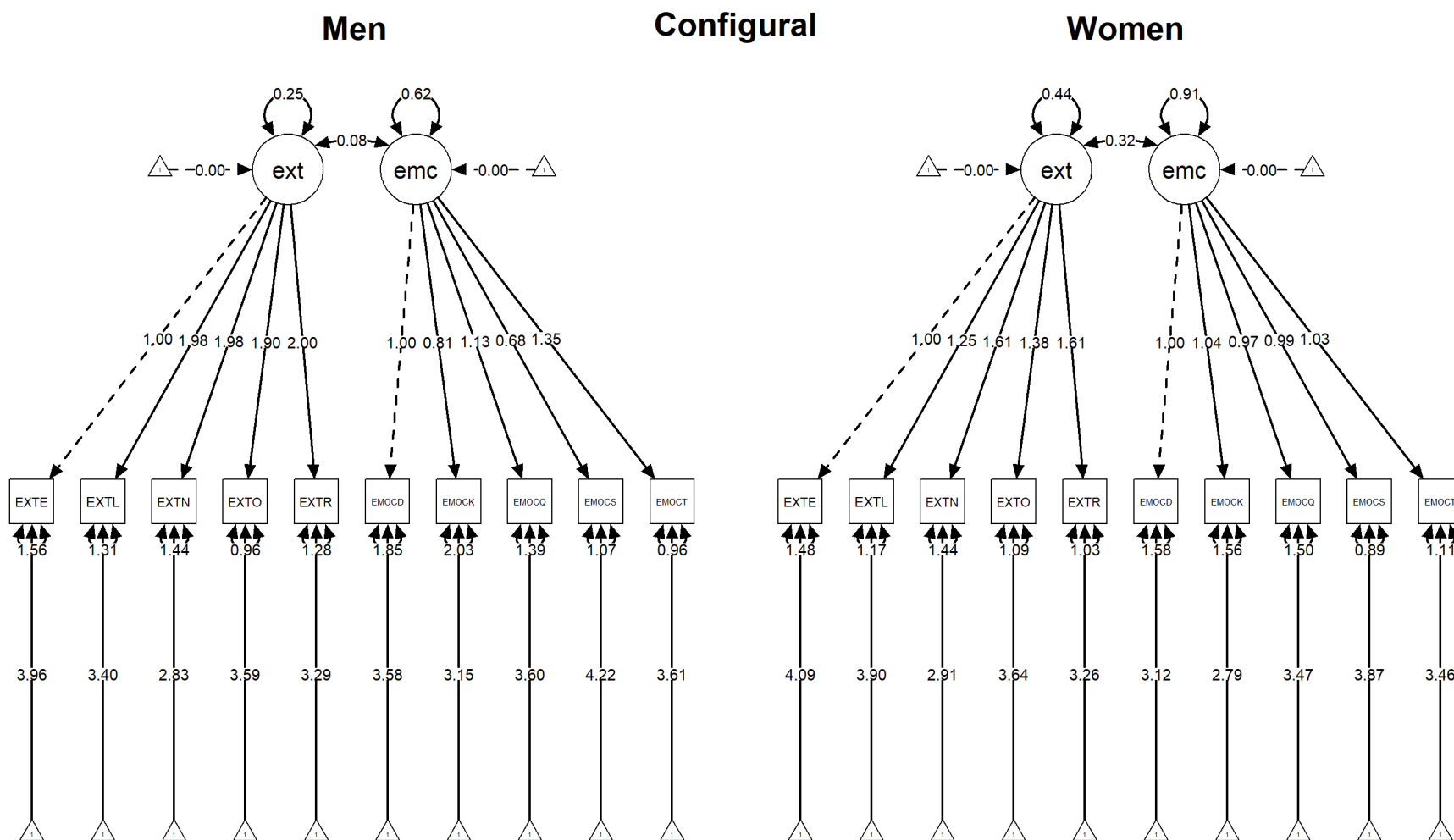
---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

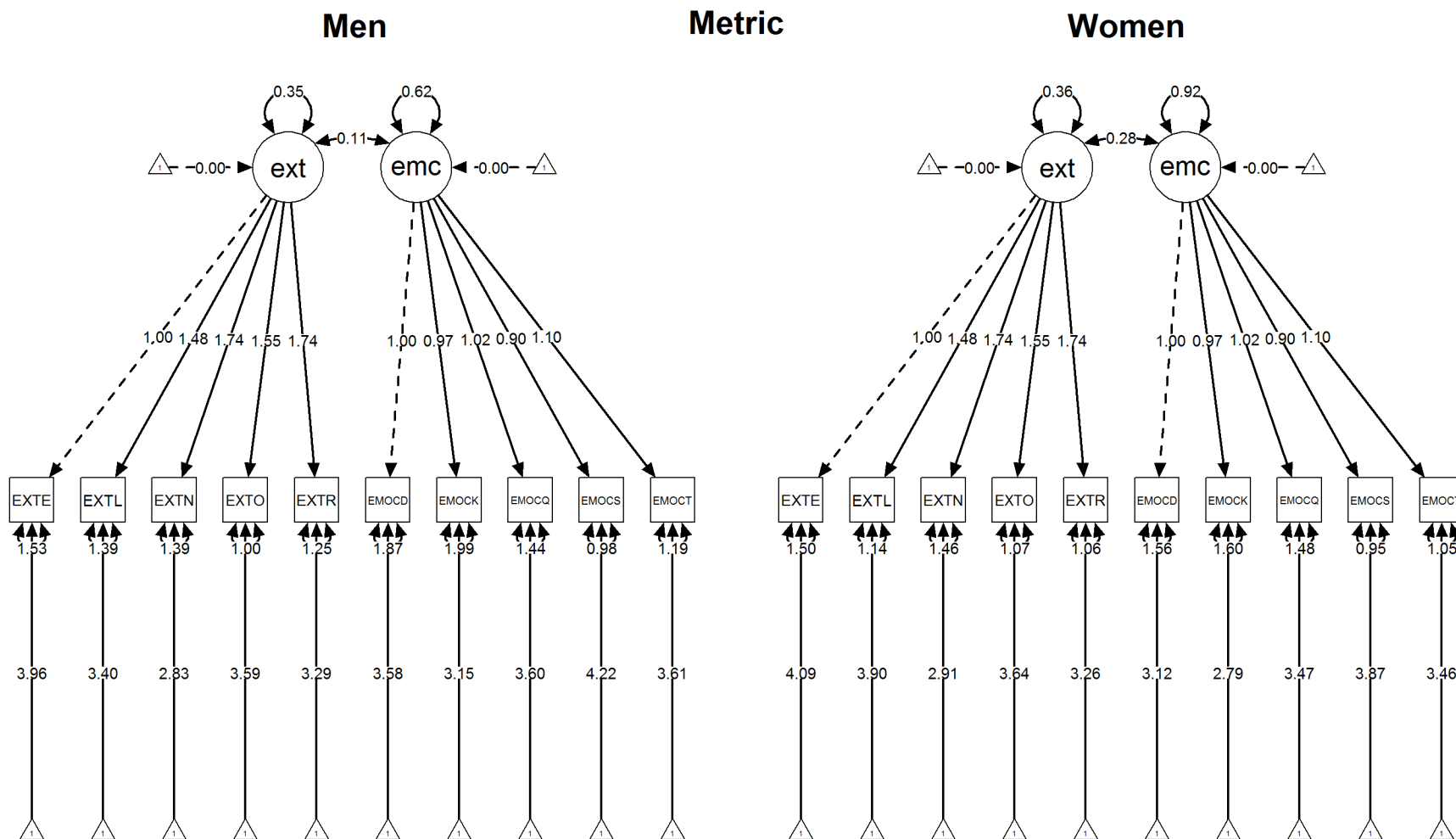
## Fit measures:

	cfi	rmsea	cfi.delta	rmsea.delta
fit.configural	0.941	0.062	NA	NA
fit.loadings	0.933	0.062	0.008	0.001
fit.intercepts	0.916	0.067	0.017	0.004
fit.means	0.905	0.070	0.011	0.003

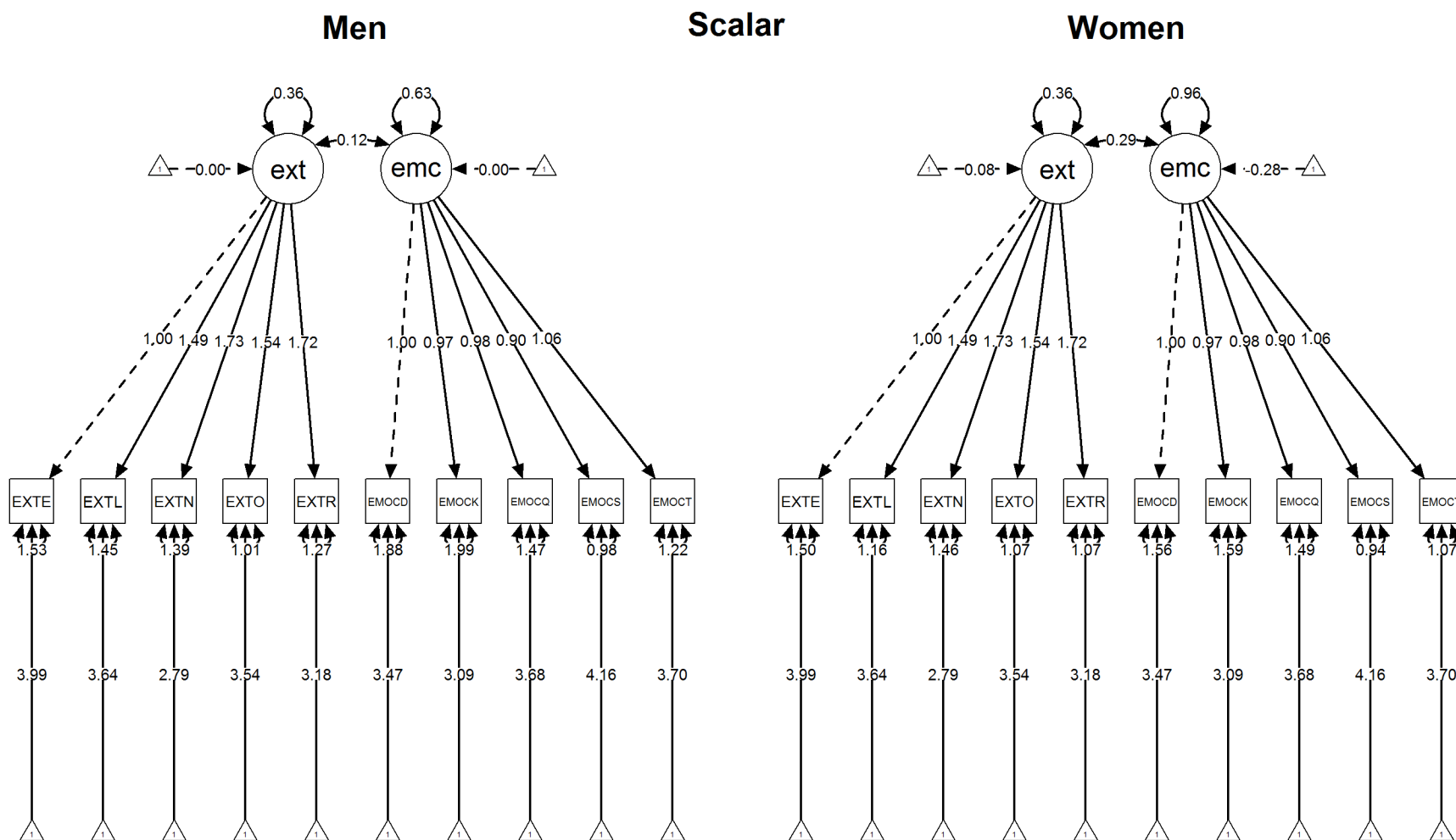
# Primer 3: Testiranje merske enakovrednosti



# Primer 3: Testiranje merske enakovrednosti



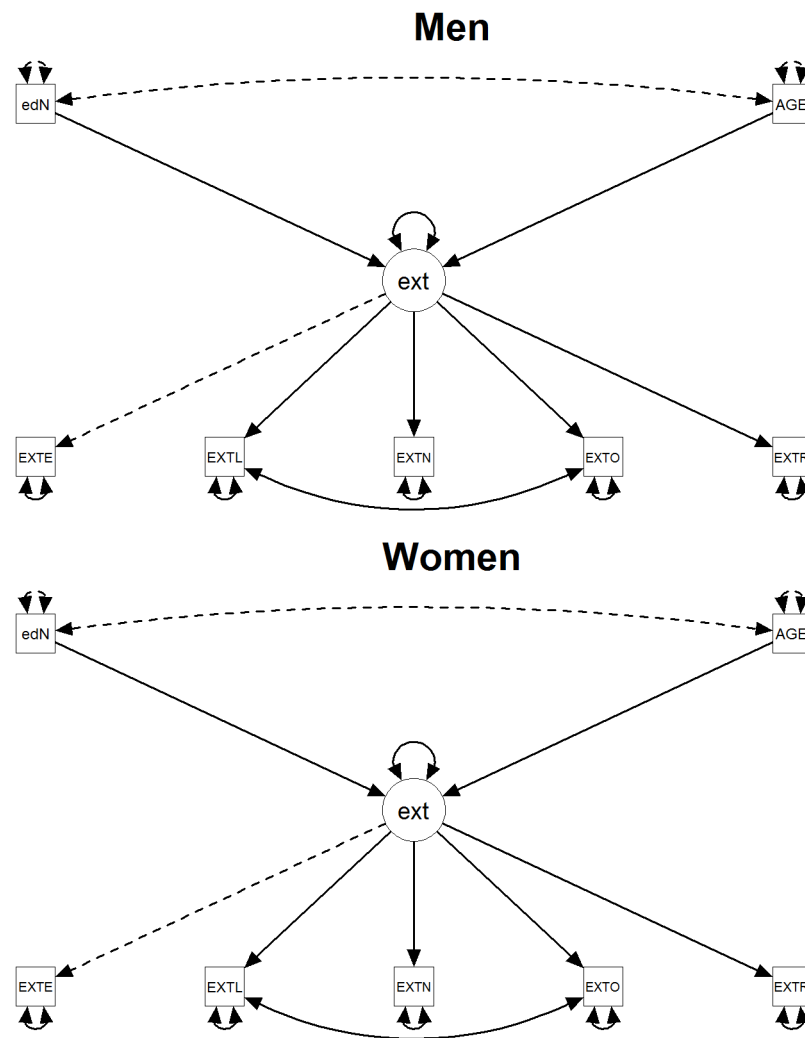
# Primer 3: Testiranje merske enakovrednosti





# SEM za več skupin

- MG model lahko razširimo tako, da vključuje tudi strukturni del
- Npr.: Je učinek starosti in izobrazbe na ekstrovertiranost enak pri obeh spolih



# Predpostavke

## Predpostavke CFA modela:

- Linearni odnosi med indikatorji in konstrukt (skritimi spremenljivkami)
- Merske napake( $\epsilon$ -ji)
  - Imajo pričakovano vrednost 0 -  $E(\epsilon)=0$
  - Imajo konstantno varianco (homoskedastičnost)
  - So nekorelirani –  $Cor(\epsilon_i, \epsilon_j) = 0$
  - Niso povezani (korelirani) s skritimi spremenljivkami -  $Cor(F, \epsilon) = 0$
- Multivariatna normalnost

# Predpostavke - Multivariatna normalnost

Preverjanje ni enostavno!

- Obstajajo testi (Alva in Estrada 2009; Cox in Small 1978; Cox in Wermuth 1994; Székely in Rizzo 2005), nekateri sicer niso dostopni v standardnih statističnih paketih (e.g. Cox in Small, 1978).

# Predpostavke - Multivariatna normalnost

- Primeri dostopnih multivariatnih testov:
  - Mardia-in test multivariatne asimetrije in sploščensoti (Mardia, 1970)
  - Razšitve Shapiro-Wilk-sovega na multivariatno normalnost (Domanski, 1998; Royston, 1983)
  - Test multivariatne normalnosti na podlagi E-statistike (Székely in Rizzo 2005)
- Enostavni pristop: Testiramo univariatno normalnost

# Predpostavke - Multivariatna normalnost

V praksi: predpostavka je pogosto (skoraj vedno) kršena

- Ordinalne (diskretne) spremenljivke – merske lestvice – indikatorji likertovega tipa (kako se strinjate na lestvici od 1 –  $k$ )
- Asimetrične porazdelitve

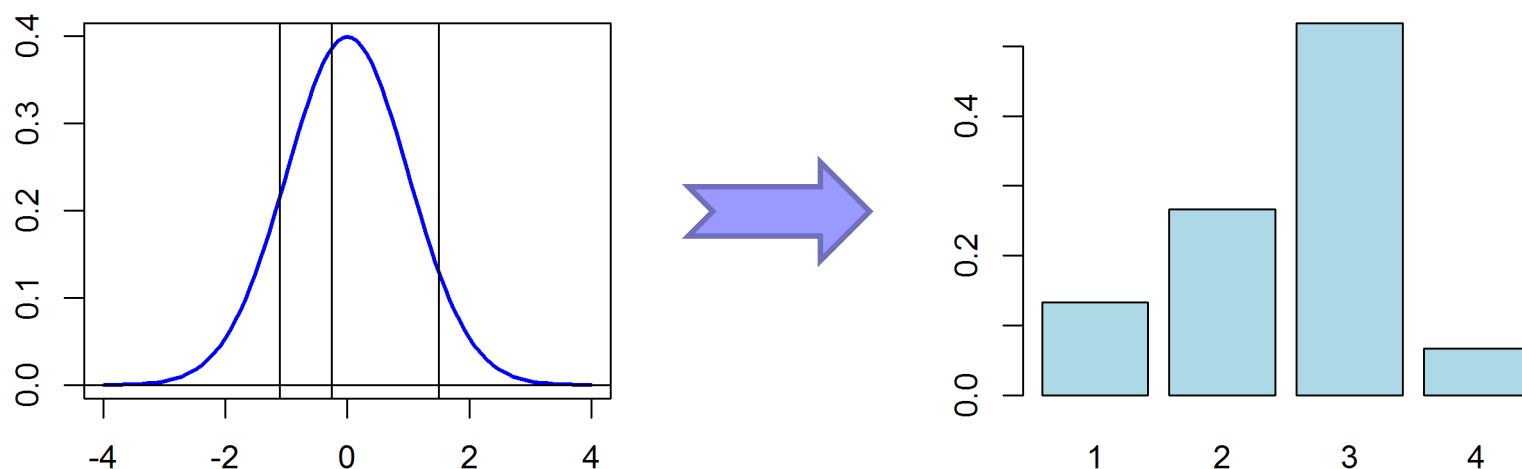
# Ordinalni indikatorji

Kako obravnavati taka odstopanja od normalnosti?

- Jih zanemarimo (ignoriramo)
  - Simulacijske študije so pokazale je, da če je vzorec dovolj velik, imamo vsaj 5 kategorij in podatki niso preveč asimetrični, metoda CFA precej robustna (DiStefano, 2002; Muthén in Kaplan, 1985; West, Finch in Curran, 1995).
- Jih modeliramo!

# Ordinalni indikatorji

- Trik: predpostavimo normalno porazdeljeno latentno spremenljivko  $x^*$ . Merjeno ordinalna (diskretna) spremenljivka  $x$  pa dobimo, ko damo to spremenljivko v razrede.



# Ordinalni indikatorji

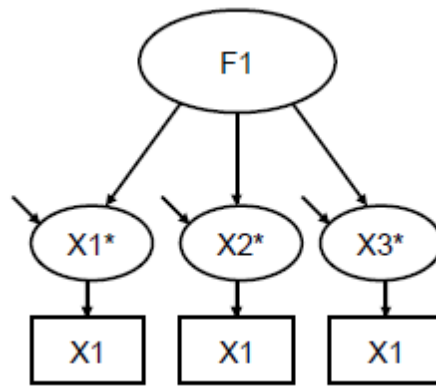
Merjeno spremenljivko  $X_{ij}$  s  $c$  kategorijam dobimo tako, da spremenljivko  $X_{ij}^*$  rekodiramo v razrede s določenimi mejami:

$$X_{ij}^g = m \text{ if } v_{jm}^g < x_{ij}^{g*} \leq v_{j(m+1)}^g$$

kjer velja  $m = 1, \dots, c$  in  $\{v_{j1}^g, v_{j2}^g, \dots, v_{j(c+1)}^g\}$  so meje (parametri) za  $j$ -ti indikator in  $g$ -to skupino (Millsap and Yun--Tein, 2004, 481).



# Ordinalni indikatorji



Novi parametri:  $(c-1)$  mej za vsak indikator  
(na skupino)

# Ordinalni indikatorji

## Problemi z identifikacijo

- Ena rešitev: Nastavi vse konstante na 0
- Meje ocenimo
- Posledica za testiranje merske enakovrednosti: ne moremo več razlikovati med metrično in skalarno enakovrednostjo

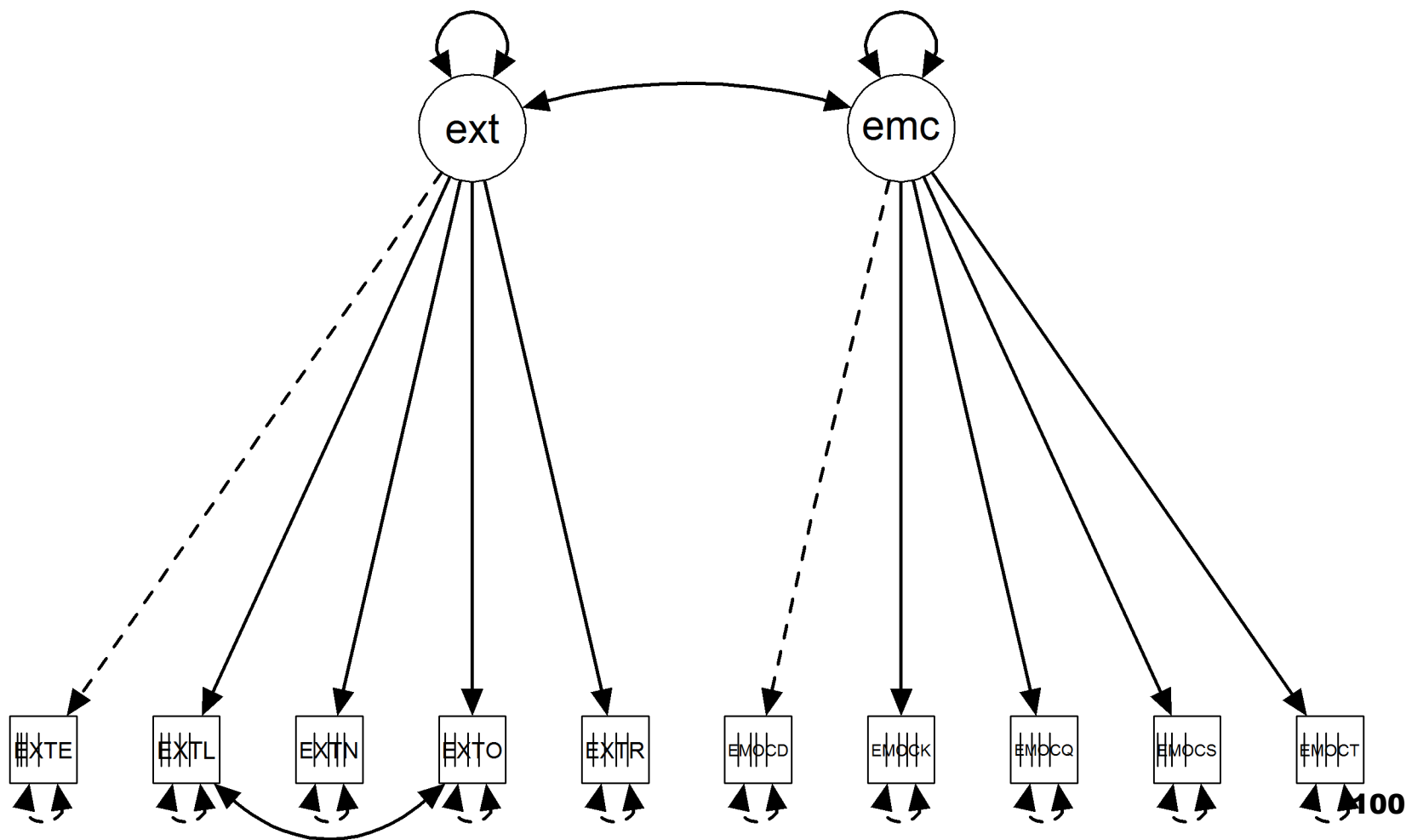


# Ordinalni indikatorji – ocenjevanje

Priporoča se uporabo WLSMV oz. DWLS z robustnimi standardnimi napakami.

Pri tem se za ordinalne spremenljivke uporablja „polychoric“ korelacija.

# Primer 4 – Primer 1 (CFA) z ordinalnimi indikatorji



# Primer 4 – faktorske uteži

Latent Variables:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
ext =~						
EXTE	1.000				0.523	0.523
EXTLR	1.160	0.109	10.672	0.000	0.606	0.606
EXTNR	1.452	0.125	11.589	0.000	0.759	0.759
EXTOR	1.286	0.120	10.760	0.000	0.672	0.672
EXTRR	1.499	0.125	11.956	0.000	0.783	0.783
emoc =~						
EMOCDR	1.000				0.640	0.640
EMOCKR	0.998	0.073	13.726	0.000	0.638	0.638
EMOCQR	1.023	0.070	14.661	0.000	0.654	0.654
EMOCSR	1.156	0.074	15.604	0.000	0.740	0.740
EMOCTR	1.150	0.076	15.200	0.000	0.736	0.736

# Primer 4 – meje

Thresholds:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EXTE t1	-1.237	0.067	-18.402	0.000	-1.237	-1.237
EXTE t2	-0.944	0.059	-15.880	0.000	-0.944	-0.944
EXTE t3	-0.669	0.055	-12.238	0.000	-0.669	-0.669
EXTE t4	-0.183	0.051	-3.610	0.000	-0.183	-0.183
EXTLR t1	-1.162	0.065	-17.903	0.000	-1.162	-1.162
EXTLR t2	-0.710	0.055	-12.852	0.000	-0.710	-0.710
EXTLR t3	-0.253	0.051	-4.971	0.000	-0.253	-0.253
EXTLR t4	0.118	0.051	2.327	0.020	0.118	0.118
EXTNR t1	-0.586	0.054	-10.919	0.000	-0.586	-0.586
EXTNR t2	-0.028	0.050	-0.562	0.574	-0.028	-0.028
EXTNR t3	0.308	0.051	6.010	0.000	0.308	0.308
EXTNR t4	0.600	0.054	11.153	0.000	0.600	0.600
EXTOR t1	-1.329	0.070	-18.878	0.000	-1.329	-1.329
EXTOR t2	-0.700	0.055	-12.699	0.000	-0.700	-0.700
EXTOR t3	-0.171	0.051	-3.369	0.001	-0.171	-0.171
EXTOR t4	0.253	0.051	4.971	0.000	0.253	0.253
EXTRR t1	-0.944	0.059	-15.880	0.000	-0.944	-0.944
EXTRR t2	-0.407	0.052	-7.842	0.000	-0.407	-0.407
EXTRR t3	0.044	0.050	0.883	0.377	0.044	0.044
EXTRR t4	0.488	0.053	9.268	0.000	0.488	0.488

# Primer 4 – meje

Thresholds:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EMOCDR  t1	-0.842	0.057	-14.656	0.000	-0.842	-0.842
EMOCDR  t2	-0.291	0.051	-5.690	0.000	-0.291	-0.291
EMOCDR  t3	-0.073	0.050	-1.445	0.149	-0.073	-0.073
EMOCDR  t4	0.312	0.051	6.090	0.000	0.312	0.312
EMOCKR  t1	-0.586	0.054	-10.919	0.000	-0.586	-0.586
EMOCKR  t2	-0.081	0.050	-1.605	0.108	-0.081	-0.081
EMOCKR  t3	0.233	0.051	4.571	0.000	0.233	0.233
EMOCKR  t4	0.595	0.054	11.075	0.000	0.595	0.595
EMOCQR  t1	-1.057	0.062	-17.035	0.000	-1.057	-1.057
EMOCQR  t2	-0.483	0.053	-9.189	0.000	-0.483	-0.483
EMOCQR  t3	-0.175	0.051	-3.449	0.001	-0.175	-0.175
EMOCQR  t4	0.220	0.051	4.331	0.000	0.220	0.220
EMOCSR  t1	-1.531	0.079	-19.390	0.000	-1.531	-1.531
EMOCSR  t2	-0.931	0.059	-15.739	0.000	-0.931	-0.931
EMOCSR  t3	-0.615	0.054	-11.386	0.000	-0.615	-0.615
EMOCSR  t4	-0.061	0.050	-1.204	0.229	-0.061	-0.061
EMOCTR  t1	-1.203	0.066	-18.187	0.000	-1.203	-1.203
EMOCTR  t2	-0.506	0.053	-9.584	0.000	-0.506	-0.506
EMOCTR  t3	-0.158	0.051	-3.129	0.002	-0.158	-0.158
EMOCTR  t4	0.308	0.051	6.010	0.000	0.308	0.308

# Primer 4 – kovariance in variance

## Covariances:

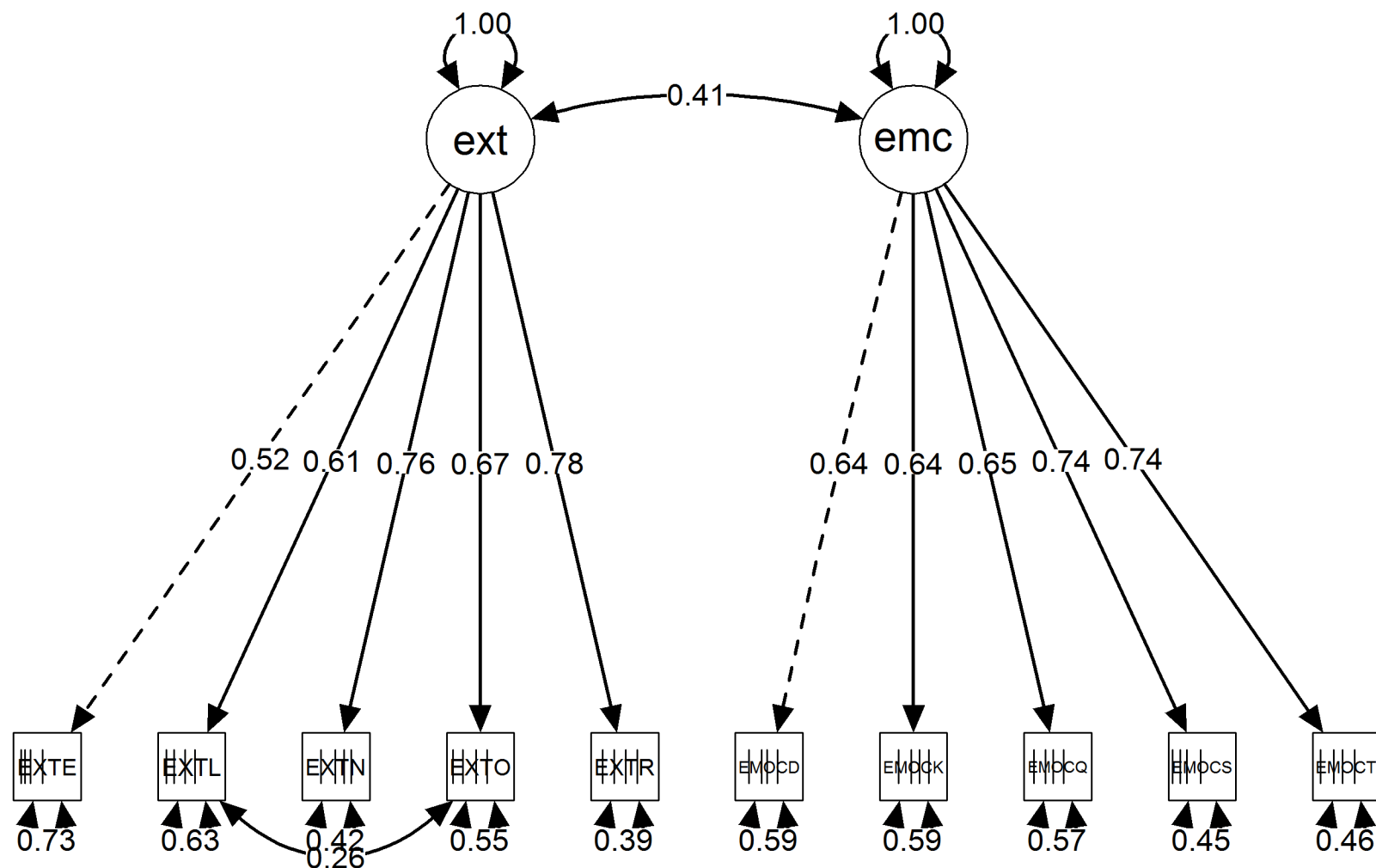
	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EXTLR ~~ EXTOR	0.155	0.034	4.530	0.000	0.155	0.263
ext ~~ emoc	0.138	0.021	6.720	0.000	0.414	0.414

## Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
EXTE	0.727				0.727	0.727
EXTLR	0.633				0.633	0.633
EXTNR	0.424				0.424	0.424
EXTOR	0.548				0.548	0.548
EXTRR	0.387				0.387	0.387
EMOCDR	0.591				0.591	0.591
EMOCKR	0.593				0.593	0.593
EMOCQR	0.572				0.572	0.572
EMOCSR	0.453				0.453	0.453
EMOCTR	0.459				0.459	0.459
ext	0.273	0.044	6.219	0.000	1.000	1.000
emoc	0.409	0.042	9.690	0.000	1.000	1.000



# Primer 4 – standardizirani rezultati



# Primer 4 – mere prileganja

lavaan (0.5-20) converged normally after 21 iterations

	Used	Total
Number of observations	620	631

Estimator	DWLS	Robust
Minimum Function Test Statistic	90.170	120.547
Degrees of freedom	33	33
P-value (Chi-square)	0.000	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	0.986	0.966
Tucker-Lewis Index (TLI)	0.981	0.954

Root Mean Square Error of Approximation:

RMSEA		0.053		0.065
90 Percent Confidence Interval	0.040	0.066	0.053	0.078
P-value RMSEA $\leq$ 0.05		0.338		0.020