Structural Equation Models

David Barron

Trinity Term 2018

Introduction

- Structural-equation models (SEMs) are multiple-equation regression models in which the response variable in one regression equation can appear as an explanatory variable in another equation. Indeed, two variables in an SEM can even effect one-another reciprocally, either directly, or indirectly through a "feedback" loop.
- Structural-equation models can include variables that are not measured directly, but rather indirectly through their effects (called indicators).
- Unmeasured variables are variously termed latent variables, constructs, or factors.

Confirmatory factor analysis

Confirmatory factor analysis

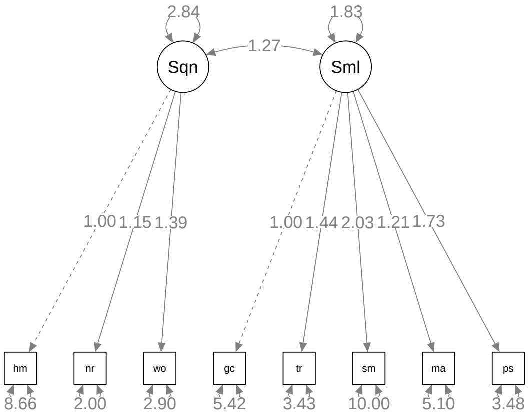
- Confirmatory factor analysis, as the name implies, involves specifying a theoretically motivated model of relationships among variables and factors and carrying out statistical tests to confirm that this model provides an adequate fit to the observed data.
- We can use different assumptions to those standard in exploratory factor analysis.
- Is a special case of Structural Equation Models.
- You can also think of CFA as being "one half" of SEMs, what is often called the *measurement model*.

CFA example

The Kaufman Assessment Battery for Children is an individually administeredd cognitive ability test for children. The eight items are claimed to be measures of two factors: sequential processing and simultaneous processing. The former require correct recall of auditory stimuli (nr: Number Recall, wo: Word Order) or visual stimuli (hm: Hand Movements) in a particular order. The latter are intended to measure more holistic, less order-dependent reasoning (gc: Gestalt Closure, tr: Triangles, sm: Spatial Memory, ma: Matrix Analogies, ps: Photo Series). We will perform a CFA using these tests on 200 children aged 10. The covariance matrix is:

```
      hm    nr   wo    gc   tr    sm   ma   ps
hm 11.56 3.182 3.45 1.928 2.94  5.71 3.71 3.98
nr  3.18 5.760 4.66 0.713 1.75  2.92 2.15 2.09
wo  3.45 4.663 8.41 1.253 2.27  3.41 2.44 3.22
gc  1.93 0.713 1.25 7.290 2.77  3.40 2.34 3.40
tr  2.94 1.750 2.27 2.770 7.29  5.33 3.18 4.70
sm  5.71 2.923 3.41 3.402 5.33 17.64 4.82 6.43
ma  3.71 2.150 2.44 2.344 3.18  4.82 7.84 3.53
ps  3.98 2.088 3.22 3.402 4.70  6.43 3.53 9.00
```

CFA with two factors



Interpretation

- Unlike EFA, it is conventional not to assume variance of unobserved factors = 1, but rather to constrain one regression parameter from each factor to = 1. [Additional homework exercise: try constraining variances of unobserved factors to be 1 and freeing all regression parameters. Model fit should be the same.]
- Can do hypothesis tests on regression parameters; all of these are statistically significant.
- Can obtain standardized results if we prefer; these are a closer equivalent to the EFA loadings.
- Also no need to assume no correlation between unobserved factors.
- But it is conventional to assume no correlation between error terms (although this can be relaxed).

Goodness of fit

- CFA (and more generally SEM) is fit using maximum likelihood estimation, so we obtain the log-likelihood that enable us to compare the relative fit of nested models. Here the log likelihood is -3779.041.
- We can compare that to the log likelihood that would be obtained from a model that reproduces the data perfectly. THis is -3759.878. From these two figures, we can calculate the likelihood ratio χ^2, which in this case is 38.325 with 19 degrees of freedom.
- The number of degrees of freedom is the number of estimated parameters fewer in the estimated model than there are observed moments in the data. There are always $k(k+1)/2$ observed second-order moments (ie, variances and covariances), where k is the number of observed variables, so in this case that is $8 \times 9 \div 2 = 36$. There are 17 estimated parameters: 6 regression parameters, 8 variances of measured variables, 2 variances and 1 covariance of unobserved factors. That gives $36 - 17 = 19$ degrees of freedom.
- Gives p-value of 0.005. That means we would have to conclude this model does not fit the data.

Other GoF statistics

- **RMSEA** This is actually a "badness of fit" statistic, so we want values close to 0. Measures size of discrepency from *close fit*, defined as $\hat{\Delta}_{model} = \max(0, \chi^2_{model} - df_{model})$. This is then standardized, to give the RMSEA statistic:

$$\hat{\epsilon} = \sqrt{\frac{\hat{\Delta}_{model}}{df_{model}(N-1)}}.$$

- **CFI** Ranges from 0 to 1. Compares estimated model to null model, using the same $\hat{\Delta}$ as before:

$$\mathrm{CFI} = 1 - \frac{\hat{\Delta}_{model}}{\hat{\Delta}_{null}}.$$

  Commonly cited rule is that CFI > 0.95, implying a fit that is 95% better than the null model. The TLI and NNFI are variants on this index.

- **SRMR** Standardised root mean square residual, where the residual is the difference between observed and fitted correlation matrices. Rule of thumb is that SRMR > 0.10 may indicate a poor fit, but it is a good idea to look at the residual matrix itself.

GoF statistics

```
             npar                fmin               chisq
           17.000               0.096              38.325
               df              pvalue      baseline.chisq
           19.000               0.005             498.336
      baseline.df     baseline.pvalue                 cfi
           28.000               0.000               0.959
              tli                nnfi                 rfi
            0.939               0.939               0.887
              nfi                pnfi                 ifi
            0.923               0.626               0.960
              rni                logl   unrestricted.logl
            0.959           -3779.041           -3759.878
              aic                 bic              ntotal
         7592.082            7648.153             200.000
             bic2               rmsea      rmsea.ci.lower
         7594.295               0.071               0.038
   rmsea.ci.upper        rmsea.pvalue                 rmr
            0.104               0.132               0.767
       rmr_nomean                srmr        srmr_bentler
            0.767               0.072               0.072
srmr_bentler_nomean       srmr_bollen  srmr_bollen_nomean
            0.072               0.072               0.072
       srmr_mplus   srmr_mplus_nomean               cn_05
            0.072               0.072             158.306
            cn_01                 gfi                agfi
          189.864               0.956               0.917
             pgfi                 mfi                ecvi
            0.505               0.953               0.362
```

GoF rules of thumb

Statistic	Rule of thumb	Comments
χ^2	Not significant (ie, p-value $>$ .05)	Influenced by sample size.
CFI	More than 0.93	Compares model to the independence model. Relatively insensitive to sample size, but biased. Must lie between 0 and 1
RMSEA	Less than .08, ideally less than .05	Has no upper bound, so hard to interpret.
TLI	Greater than .9 or .95	Compares to null model, but controls for complexity. Relatively insensitive to sample size.
AIC		Only useful for comparing models. Controls for model complexity.
BIC		Similar to AIC, with greater penalty for complexity

Not clear which measure is "best", so good idea to look at more than one. There are many others that you might see used in articles, but these are the most common. In the example, most show inadequate fit.

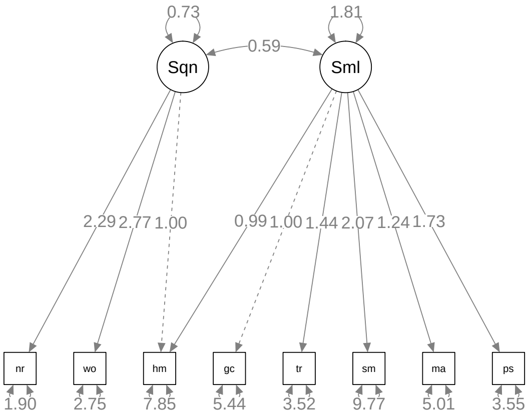
Modification indices

<div class="kable-table">

	lhs	op	rhs	mi	epc	sepc.lv	sepc.all	sepc.nox
25	Simultan	=~	hm	20.10	1.054	1.428	0.421	0.421
35	nr	~~	wo	20.10	4.741	4.741	0.685	0.685
26	Simultan	=~	nr	7.01	-0.510	-0.691	-0.289	-0.289
29	hm	~~	wo	7.01	-1.746	-1.746	-0.178	-0.178
32	hm	~~	sm	4.85	1.609	1.609	0.113	0.113
33	hm	~~	ma	3.80	0.995	0.995	0.105	0.105
23	Sequent	=~	ma	3.25	0.269	0.454	0.162	0.162
40	nr	~~	ps	3.15	-0.502	-0.502	-0.070	-0.070
20	Sequent	=~	gc	2.90	-0.254	-0.429	-0.159	-0.159
55	ma	~~	ps	2.73	-0.733	-0.733	-0.088	-0.088

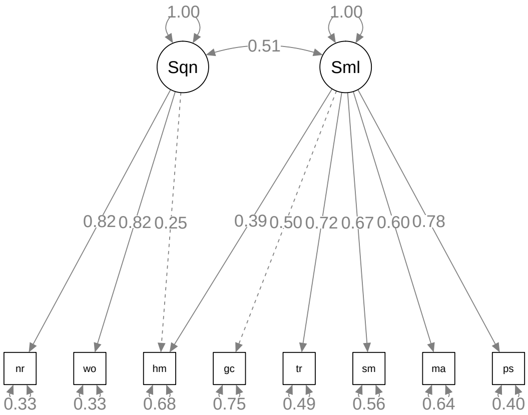
These show the effect of freeing a constrained parameter. We may then choose to modify our model accordingly.

Modified model




```
 chisq     df pvalue    cfi  rmsea    gfi    tli
18.108 18.000  0.449  1.000  0.005  0.977  1.000
```

Standardized results

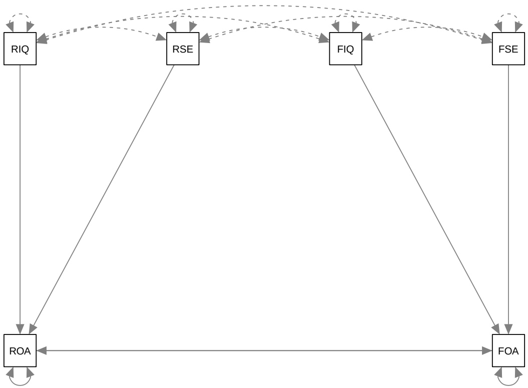


Path models

Path models

Path models have only observed variables, but they differ from the regression models we've seen so far in that they allow us to model indirect and reciprocal effects as well as direct effects. Models with reciprocal effects are known as *nonrecursive* models. The following example is Duncan, Haller, and Portes's (nonrecursive) peer-influences model. It is based on a sample of Michigan high school students. It is an example of a general class of peer influence models that acknowledge that if I am influencing my peers (e.g., my best friend), then he or she could be influencing me.

Path diagram



Duncan, Haller, and Portes's (nonrecursive) peer-influences model: RIQ: respondent's IQ; RSE: respondent's family SES; FSE: best friend's family SES; FIQ: best friend's IQ; ROA: respondent's occupational aspiration; FOA: best friend's occupational aspiration.

Results

```
lavaan (0.5-23.1097) converged normally after  24 iterations

  Number of observations                           329

  Estimator                                         ML
  Minimum Function Test Statistic                2.820
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.244

Parameter Estimates:

  Information                                 Expected
  Standard Errors                             Standard

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  ROccAsp ~
    RIQ               0.237    0.053    4.480    0.000
    RSES              0.176    0.047    3.728    0.000
    FOccAsp           0.398    0.104    3.816    0.000
  FOccAsp ~
    FIQ               0.311    0.056    5.598    0.000
    FSES              0.219    0.047    4.689    0.000
    ROccAsp           0.422    0.131    3.215    0.001

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
 .ROccAsp ~~
   .FOccAsp          -0.494    0.136   -3.634    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .ROccAsp           0.790    0.074   10.749    0.000
   .FOccAsp           0.715    0.086    8.272    0.000


  cfi rmsea  srmr
0.997 0.035 0.013
```

Add covariance between the two error terms

```
lavaan (0.5-23.1097) converged normally after  24 iterations

  Number of observations                           329

  Estimator                                         ML
  Minimum Function Test Statistic                2.820
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.244

Parameter Estimates:

  Information                                 Expected
  Standard Errors                             Standard

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  ROccAsp ~
    RIQ               0.237    0.053    4.480    0.000
    RSES              0.176    0.047    3.728    0.000
    FOccAsp           0.398    0.104    3.816    0.000
  FOccAsp ~
    FIQ               0.311    0.056    5.598    0.000
    FSES              0.219    0.047    4.689    0.000
    ROccAsp           0.422    0.131    3.215    0.001

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
 .ROccAsp ~~
   .FOccAsp          -0.494    0.136   -3.634    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .ROccAsp           0.790    0.074   10.749    0.000
   .FOccAsp           0.715    0.086    8.272    0.000


  cfi rmsea  srmr
0.997 0.035 0.013
```

Structural equation models

Types of variables

Several classes of variables appear in SEMs:

- Endogenous variables are the response variables of the model.
  - There is one structural equation (regression equation) for each endogenous variable.
  - An endogenous variable may, however, also appear as an explanatory variable in other structural equations.
  - For the kinds of models that we will consider, the endogenous variables are (as in the single-equation linear model) quantitative continuous variables.
- Exogenous variables appear only as explanatory variables in the structural equations.
  - The values of exogenous variable are therefore determined outside of the model (hence the term).

Types of variables 2

- Structural errors (or disturbances) represent the aggregated omitted causes of the endogenous variables, along with measurement error (and possibly intrinsic randomness) in the endogenous variables.
  - There is one error variable for each endogenous variable (and hence for each structural equation).
  - The errors are assumed to have zero expectations and to be independent of (or at least uncorrelated with) the exogenous variables.
  - The errors for different observations are assumed to be independent of one another, but (depending upon the form of the model) different errors for the same observation may be related.

General structural equation model

That is, a structural equation model can contain some or all of the following:

- Exogenous concepts (unobserved);
- Endogenous concepts (unobserved);
- Indicators of exogenous concepts;
- Indicators of endogenous concepts;
- Structural errors;
- Measurement errors;
- Structural parameters
- Covariances

LISREL

There are three basic equations in a SEM. These are shown using the notation that is standard in LISREL, the first and most well-known computer software for analysing these models:

$$
\begin{aligned}
\eta &= \beta\eta + \Gamma\xi + \zeta \\
y &= \Lambda_y\eta + \epsilon \\
x &= \Lambda_x\xi + \delta
\end{aligned}
$$

Meanings

These terms have the following meaning:

- η: Endogenous concepts.
- β: Structural coefficients for the relationships among endogenous concepts.
- ξ: Exogenous concepts.
- Γ: Structural coefficients for the relationships between exogenous and endogenous concepts.
- ζ: Structural errors.
- x and y: Observed exogeneous and endogenous indicators, respectively.
- Λ_y: Structural coefficients relating indicators to endogenous concepts.
- ϵ and δ: Measurement errors.
- Λ_x: Structural coefficients relating indicators to exogenous concepts.

Covariance matrices

In addition, the following covariance matrices are defined:

- Φ: Covariances among the concepts.
- Ψ: Covariances among the structural errors.
- Θ_ϵ: Covariances among the ϵ measurement errors.
- Θ_δ: Covariances among the δ measurement errors.

Assumptions of general SEM

- The measurement errors, δ and ϵ,
  - have expectations of 0;
  - are each multivariately-normally distributed;
  - are independent of each other;
  - are independent of the latent exogenous variables (ξ), latent endogenous variables (η), and structural disturbances (ζ).
- The N observations are independently sampled.
- The latent exogenous variables, ξ, are multivariate normal.
- This assumption is unnecessary for exogenous variables that are measured without error.

Assumptions 2

- The structural disturbances, ζ
  - have expectation 0;
  - are multivariately-normally distributed;
  - are independent of the latent exogenous variables (ξ's).
- Under these assumptions, the observable indicators, x and y, have a multivariate-normal distribution.

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} \sim N_{q+p}(\mathbf{0}, \Sigma)$$

where Σ represents the population covariance matrix of the indicators.

Identification of SEMs

Identification of models with latent variables is a complex problem without a simple general solution.

- A global necessary condition for identification is that the number of free parameters in the model can be no larger than the number of variances and covariances among observed variables,

$$\frac{(k)(k+1)}{2}$$

- This condition is insufficiently restrictive to give us any confidence that a model that meets the condition is identified.
- That is, it is easy to meet this condition and still have an underidentified model.

Useful rule

A useful rule that sometimes helps is that a model is identified if:

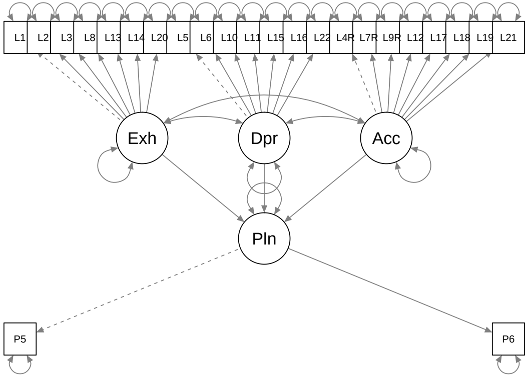
1. all of the measurement errors in the model are uncorrelated with one another;
2. there are at least two unique indicators for each latent variable, or if there is only one indicator for a latent variable, it is measured without error;
3. the structural sub model would be identified were it an observed variable model

Estimation

The variances and covariances of the observed variables (Σ) are functions of the parameters of the SEM (β, Γ, Λ_x, Λ_y, Φ, Θ_δ , Θ_ϵ, and Ψ).

- In any particular model, there will be restrictions on many of the elements of the parameter matrices.
- Most commonly, these restrictions are exclusions: certain parameters are prespecified to be 0.
- The Λ matrices (or the Ψ matrix) must contain normalizing restrictions to set the metrics of the latent variables.
- If the restrictions on the model are sufficient to identify it, then MLEs of the parameters can be found.

SEM example


L1
L2
L3
L8
L13
L14
L20
L5
L6
L10
L11
L15
L16
L22
L4R
L7R
L9R
L12
L17
L18
L19
L21
Exh
Dpr
Acc
Pln
P5
P6

Model explained

- There are three exogenous latent variables (ξ); these are the three components of burnout as measured by the Maslach Burnout Inventory.
- Each exogenous latent variable has a number of exogenous indicators (x), 22 in all.
- Each indicator has a path with a coefficient (Λ_x). One per latent variable is fixed to 1, so there are $22 - 3 = 19$ free parameters.
- Each of the exogenous indicators has an error (δ).
- There is one endogenous latent variable (η), the career plans of a nurse.
- There is a path between each exogenous latent variable and the endogenous latent variable (Γ)

Continued

- The endogenous latent variable has a structural error (ζ).
- The endogenous latent variable has two indicators (y).
- The endogenous indicators have error terms (ϵ).
- Each indicator has a path with a coefficient (Λ_y), one of which will be set to one, so there is 1 free parameter.
- There are 3 covariances between the exogenous concepts and 3 variances (Φ), so 6 free parameters.
- Covariances among the δ measurement errors are zero (that it, Θ_δ is a diagonal matrix with 22 free parameters).
- Covariances among the ϵ measurement errors are zero (that it, Θ_ϵ is a diagonal matrix with 2 free parameters).
- There is only one structural error, so Ψ is just a single free parameter.

Specification

There are 24 observed variables, and hence there are $24 \times 25 \div 2 = 300$ observed variances and covariances. There are 54 free parameters to be estimated. Therefore there are $300 - 54 = 246$ degrees of freedom.

```
mas_mod <- ' Exhaust =~ L1 + L2 + L3 + L8 + L13 + L14 + L20
             Depers =~ L5 + L6 + L10 + L11 + L15 + L16 + L22
             Accomp =~ L4R + L7R + L9R + L12R + L17R + L18R + L19R + L21R
             Plans =~ P5 + P6
             Plans ~ Exhaust + Depers + Accomp '
```

Results

<div class="kable-table">

lhs	op	rhs	est	se
Exhaust	=~	L1	1.000	0.000
Exhaust	=~	L2	0.938	0.044
Exhaust	=~	L3	1.214	0.056
Exhaust	=~	L8	1.322	0.055
Exhaust	=~	L13	0.997	0.055
Exhaust	=~	L14	1.011	0.058
Exhaust	=~	L20	0.799	0.053
Depers	=~	L5	1.000	0.000
Depers	=~	L6	0.855	0.070
Depers	=~	L10	1.565	0.091
Depers	=~	L11	1.531	0.089
Depers	=~	L15	0.544	0.045
Depers	=~	L16	0.456	0.043
Depers	=~	L22	0.910	0.082
Accomp	=~	L4R	1.000	0.000
Accomp	=~	L7R	1.435	0.128
Accomp	=~	L9R	1.740	0.150
Accomp	=~	L12R	1.404	0.148
Accomp	=~	L17R	1.561	0.142
Accomp	=~	L18R	1.828	0.167
Accomp	=~	L19R	1.869	0.157
Accomp	=~	L21R	1.585	0.146
Plans	=~	P5	1.000	0.000
Plans	=~	P6	1.571	0.125
Plans	~	Exhaust	-0.151	0.022
Plans	~	Depers	-0.131	0.033
Plans	~	Accomp	-0.138	0.043

Results, continued

	lhs	op	rhs	est	se
28	L1	~~	L1	1.674	0.086
29	L2	~~	L2	1.005	0.055
30	L3	~~	L3	1.495	0.083
31	L8	~~	L8	0.853	0.061
32	L13	~~	L13	2.231	0.111
33	L14	~~	L14	2.600	0.129
34	L20	~~	L20	2.585	0.124
35	L5	~~	L5	1.475	0.075
36	L6	~~	L6	1.763	0.086
37	L10	~~	L10	0.956	0.068
38	L11	~~	L11	0.963	0.067
39	L15	~~	L15	0.751	0.036
40	L16	~~	L16	0.791	0.038
41	L22	~~	L22	2.738	0.131
42	L4R	~~	L4R	1.165	0.058
43	L7R	~~	L7R	1.291	0.069
44	L9R	~~	L9R	1.519	0.084
45	L12R	~~	L12R	2.725	0.134
46	L17R	~~	L17R	1.690	0.088
47	L18R	~~	L18R	2.434	0.127
48	L19R	~~	L19R	1.387	0.081
49	L21R	~~	L21R	1.880	0.097
50	P5	~~	P5	0.355	0.029
51	P6	~~	P6	0.255	0.062
52	Exhaust	~~	Exhaust	1.562	0.133
53	Depers	~~	Depers	0.747	0.083
54	Accomp	~~	Accomp	0.316	0.047
55	Plans	~~	Plans	0.254	0.027
56	Exhaust	~~	Depers	0.548	0.055
57	Exhaust	~~	Accomp	0.069	0.028
58	Depers	~~	Accomp	0.143	0.023

Interpretation

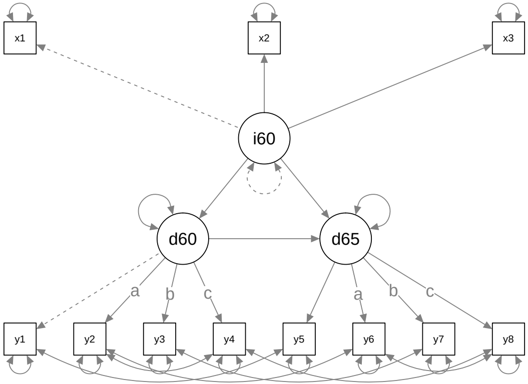
High numbers of the 'Plans' variable mean more likely to stay in nursing, so the interpretation of these results is that each component of burnout reduces the chances that a nurse will remain in nursing. All those parameter estimates are statistically significant.

```
   chisq        df    pvalue       cfi     rmsea      srmr
1301.136   246.000     0.000     0.860     0.067     0.074
```

<div class="kable-table">

	lhs	op	rhs	mi	epc	sepc.lv	sepc.all	sepc.nox
302	L10	~~	L11	124.0	0.786	0.786	0.286	0.286
69	Exhaust	=~	L12R	116.8	0.509	0.636	0.348	0.348
362	L4R	~~	L7R	116.0	0.489	0.489	0.288	0.288
290	L6	~~	L16	104.1	0.413	0.413	0.279	0.279
131	L1	~~	L2	57.2	0.381	0.381	0.137	0.137
86	Depers	=~	L12R	47.0	0.509	0.440	0.241	0.241
95	Accomp	=~	L3	43.1	0.580	0.326	0.167	0.167
158	L2	~~	L20	38.8	-0.368	-0.368	-0.126	-0.126
80	Depers	=~	L13	38.4	0.481	0.416	0.214	0.214
225	L13	~~	L22	34.7	0.500	0.500	0.140	0.140

Political Democracy example



Note

- Data from 75 countries
- Two endogenous concepts (η), Democracy in 1960 and in 1965.
- Each η has four indicators (y); press freedom, freedom of political opposition, fairness of elections, effectiveness of elected legislature
- One exogenous concept (ξ), Industrialisation in 1960.
- This has three indicators (x): GNP per capita, energy consumption per capita, percentage of labour force in industry.
- One β parameter and two Γ parameters.
- This model specifies some correlations between error terms (ie, Λ_ϵ is *not* a diagonal matrix).
- This model constrains some parameters in the measurement model to be equal.
- There are 66 observed moments and 28 parameters to be estimated, so 38 degrees of freedom.

Results

<div class="kable-table">

lhs	op	rhs	label	est	se
ind60	=~	x1		1.000	0.000
ind60	=~	x2		2.007	0.079
ind60	=~	x3		1.681	0.095
dem60	=~	y1		1.000	0.000
dem60	=~	y2	a	1.152	0.136
dem60	=~	y3	b	1.140	0.122
dem60	=~	y4	c	1.214	0.119
dem65	=~	y5		0.938	0.106
dem65	=~	y6	a	1.152	0.136
dem65	=~	y7	b	1.140	0.122
dem65	=~	y8	c	1.214	0.119
dem60	~	ind60		1.381	0.269
dem65	~	ind60		0.579	0.179
dem65	~	dem60		0.886	0.086
ind60	~~	ind60		1.000	0.000
y1	~~	y5		0.589	0.353
y2	~~	y4		1.488	0.690
y2	~~	y6		2.179	0.736
y3	~~	y7		0.733	0.607
y4	~~	y8		0.362	0.441
y6	~~	y8		1.341	0.579
x1	~~	x1		0.080	0.021
x2	~~	x2		0.134	0.067
x3	~~	x3		0.465	0.089
y1	~~	y1		1.823	0.436
y2	~~	y2		7.630	1.367
y3	~~	y3		4.952	0.950
y4	~~	y4		3.269	0.721
y5	~~	y5		2.334	0.477
y6	~~	y6		4.939	0.921
y7	~~	y7		3.530	0.707
y8	~~	y8		3.272	0.703
dem60	~~	dem60		4.006	0.901
dem65	~~	dem65		0.193	0.250

Goodness of fit

```
              npar               fmin              chisq
            28.000              0.368             55.269
                df             pvalue     baseline.chisq
            38.000              0.035            730.654
       baseline.df    baseline.pvalue                cfi
            55.000              0.000              0.974
               tli               nnfi                rfi
             0.963              0.963              0.891
               nfi               pnfi                ifi
             0.924              0.639              0.975
               rni               logl  unrestricted.logl
             0.974          -1556.363          -1528.728
               aic                bic             ntotal
          3168.725           3233.615             75.000
              bic2              rmsea     rmsea.ci.lower
          3145.366              0.078              0.022
    rmsea.ci.upper       rmsea.pvalue                rmr
             0.120              0.159              1.555
        rmr_nomean               srmr        srmr_bentler
             1.555              0.367              0.367
srmr_bentler_nomean       srmr_bollen  srmr_bollen_nomean
             0.367              0.100              0.100
        srmr_mplus  srmr_mplus_nomean              cn_05
             0.219              0.219             73.442
             cn_01                gfi               agfi
            83.997              0.900              0.827
              pgfi                mfi               ecvi
             0.518              0.891              1.484
```

Extras

Definitions of GoF statistics

Statistic	Definition	Criterion
GFI	$1 - [\chi^2_{model}/\chi^2_{null}]$	> 0.9
NFI	$[\chi^2_{null} - \chi^2_{model}]/\chi^2_{null}$	> 0.95
RFI	$1 - [(\chi^2_{model}/df_{model})/(\chi^2_{null}/df_{null})]$	
IFI	$(\chi^2_{null} - \chi^2_{model})/(\chi^2_{null} - df_{model})$	
TLI	$[(\chi^2_{null}/df_{null}) - (\chi^2_{model}/df_{model})]/[(\chi^2_{null}/df_{null}) - 1]$	
CFI	$1 - [(\chi^2_{model} - df_{model})/(\chi^2_{null} - df_{null})]$	> 0.95
Model AIC	$\chi^2_{model} + 2q$(number of free parameters)	
Null AIC	$\chi^2_{null} + 2q$(number of free parameters)	
RMSEA	$\sqrt{[\chi^2_{model} - df_{model}]/[(N-1)df_{model}]}$	< 0.07