

SEM

Notes

- The hypothetical path models in SEM are used for hypothesis testing, and consist of endogenous/dependent/latent variables and exogenous/independent/observed variables.
- Unlike in a regression analysis, in a SEM/path analysis model, in addition to direct effect there is also indirect effect of an Independent Variable (IV), via an mediating variable, on the Dependent Variable (DV)
- Exogenous variables are measured by single items, and are not influenced by others.
- Endogenous variables are defined in terms of >1 exogenous variables which are thought be the cause of the variance in the endogenous ones.
- The strength and direction of the associations are measured in terms of the route coefficients and their signs.
- Interpretation: a standadised path coefficient of -0.2 is read as, *ceteris paribus*, when the exogenous var increases by 1 SD from its mean, the endogenous var is expected to decrease by $.2$ SD from its own mean; a path coefficient of $.15$ is read as, *ceteris paribus*, when the exogenous var increases by 1 SD from its mean, the endogenous var is expected to increase by 0.15 SD from its own mean.

Example 1

Data

General Social Survey 2016.

Hypotheses

1. Job satisfaction is positively correlated with Psychological well being.
2. Social networking is positively correlated with Psychological well being.
3. Social networking is positively correlated with Job satisfaction.

Indicators

JS = Job satisfaction

rsecjob = job is secure, rintjob = job is interesting, rhlpoths = helps others in job

SN= Social networking

use of: twitter whatsapp linkedin

Psychological well-being

cesd1 = how much time felt depressed in past wk, cesd2 = how much time sleep was restless in past wk, cesd3 = how much time felt happy in past wk

Equation

```
sem (SN -> twitter whatsapp linkedin ) (JS -> rsecjob rintjob  
rhlpths) (WB -> cesd1 cesd2 cesd3) (JS -> WB ) (SN -> JS WB) , latent(SN  
JS WB ) nocapslatent cformat(%9.2f)
```

Results

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
-----+-----						

Structural						
JS <-						
	SN	0.29	0.13	2.23	0.026	0.04
						0.55
-----+-----						

WB <-						
	JS	0.46	0.15	2.99	0.003	0.16
						0.77
	SN	0.14	0.20	0.71	0.480	-0.25
						0.52
-----+-----						

Measurement						
twitter <-						
	SN	1.00	(constrained)			
	_cons	1.79	0.02	95.73	0.000	1.75
						1.83
-----+-----						

whatsapp <-						
	SN	0.41	0.12	3.48	0.001	0.18
						0.64
	_cons	1.85	0.02	114.24	0.000	1.82
						1.89


```

-----+-----
-----
cesd2 <-      |
      WB |      0.68      0.12      5.80      0.000      0.45
0.92
      _cons |      2.02      0.04      50.26      0.000      1.94
2.10
-----+-----

```

```

-----+-----
-----
cesd3 <-      |
      WB |      -0.81      0.13      -6.36      0.000      -1.06
-0.56
      _cons |      3.04      0.03      92.62      0.000      2.97
3.10
-----+-----

```

```

-----+-----
-----
var(e.twitter)|      0.12      0.02      0.09
0.16
var(e.whatsapp)|      0.12      0.01      0.10
0.13
var(e.linkedin)|      0.13      0.04      0.08
0.22
var(e.rsecjob)|      1.01      0.07      0.88
1.15
var(e.rintjob)|      0.33      0.15      0.13
0.83
var(e.rhlpths)|      0.54      0.05      0.45
0.65
var(e.cesd1)|      0.10      0.04      0.05
0.21
var(e.cesd2)|      0.64      0.05      0.56
0.74
-----+-----

```

var(e.cesd3)	0.34	0.03	0.28
0.41			
var(e.JS)	0.07	0.03	0.02
0.18			
var(e.WB)	0.24	0.04	0.17
0.35			
var(SN)	0.04	0.02	0.02
0.09			

 LR test of model vs. saturated: $\chi^2(24) = 50.68$, Prob > $\chi^2 = 0.0012$

Interpretation

Hypothesis 1 is supported by the findings: Job satisfaction is positively correlated with Psychological well being

($\beta = 0.29$, $p = 0.026$)

Hypothesis 2 is supported by the findings: Social networking is positively correlated with Job satisfaction

($\beta = 0.46$, $p = 0.003$)

Hypothesis 3 is NOT supported by the findings: Social networking is not correlated with Psychological well being

($\beta = 0.14$, $p = 0.480$)

Goodness of fit statistics

```
. estat gof, stats(all)
```

```
-----  
--  
Fit statistic      |      Value  Description  
-----+-----  
--  
Likelihood ratio  |  
    chi2_ms(24) |    50.680   model vs. saturated  
    p > chi2 |    0.001  
    chi2_bs(36) |   434.097   baseline vs. saturated  
    p > chi2 |    0.000  
-----+-----  
--  
Population error  |  
    RMSEA |    0.048   Root mean squared error of  
approximation  
    90% CI, lower bound |    0.030  
    upper bound |    0.067  
    pclose |    0.525   Probability RMSEA <= 0.05  
-----+-----  
--  
Information criteria |  
    AIC |   8117.664   Akaike's information criterion  
    BIC |   8242.437   Bayesian information criterion  
-----+-----  
--  
Baseline comparison |  
    CFI |    0.933   Comparative fit index
```

```

          TLI |      0.899   Tucker-Lewis index
-----+-----
--
Size of residuals |
          SRMR |      0.044   Standardized root mean squared
residual
          CD |      0.550   Coefficient of determination
-----
--

```

Interpretation

Chi2 $p < .05$; the hypothesis of perfect fit is rejected (sensitive to sample size)

RMSEA $< .05$ indicates a good fit. (The closer to 0 the better)

CFI $> .90$ indicates good fit; compares the fit of a target model to the fit of a null model.