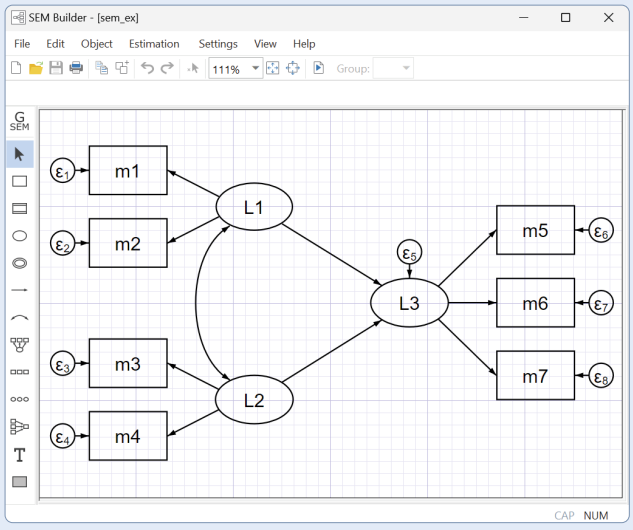
SEM

Structural equation modeling

Stata provides an easy-to-use and comprehensive suite of tools for SEM—everything you need for fitting your model, evaluating model fit, and interpreting results. And all of this is integrated in a complete package for statistics, visualization, data manipulation, and reporting.



▪ Easy model specification
- Path diagram builder
- Intuitive command syntax

▪ Types of models
- Path analysis
- Mediation analysis
- Confirmatory factor analysis (CFA)
- Multiple indicators and multiple causes (MIMIC) models
- Latent growth curve models
- Hierarchical confirmatory factor models
- Multiple-group models
- Models with binary, ordinal, count, nominal, and survival-time outcomes
- Multilevel models
- Latent class analysis (LCA)
- And many more

▪ Interpretation
- Direct, indirect, and total effects
- Standardized and unstandardized estimates

▪ Model fit
- Model x^2
- RMSEA
- CFI
- TLI
- SRMR
- Likelihood-ratio and Wald tests
- Modification indices
- Latent class model statistics

▪ Estimation methods
- Maximum likelihood
- Maximum likelihood with missing values, sometimes called FIML
- Asymptotic distribution free (ADF)

▪ Standard errors
- Satorra–Bentler
- Robust (Huber/White/sandwich estimator)
- Cluster–robust
- Bootstrap
- And more

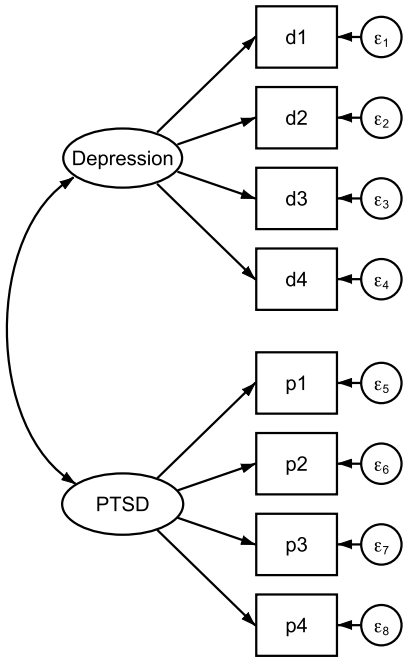
▪ Survey data support
- Sampling weights
- Stratification and poststratification
- Multistage cluster sampling

Use sem to fit linear models

Here we fit a two-factor CFA model with four measurements of depression and four measurements of posttraumatic stress disorder (PTSD). The results are the same whether we use the straightforward command syntax,

```
. sem (Depression -> d1 d2 d3 d4)
      (PTSD -> p1 p2 p3 p4)
```

or draw the path diagram,



Report model fit statistics

Many commands are available for evaluating the fit of our model. For instance,

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

Viewer - view sem2.smcl

view sem2.smcl

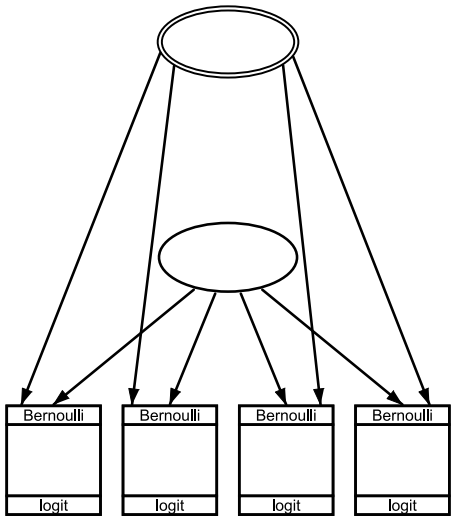
Dialog | Also see | Jump to

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(8)	12.648	model vs. saturated
p > chi2	0.125	
chi2_bs(15)	1164.146	baseline vs. saturated
p > chi2	0.000	
Population error		
RMSEA	0.052	Root mean squared error of approximation
90% CI, lower bound	0.000	
upper bound	0.104	
pclose	0.418	Probability RMSEA <= 0.05
Information criteria		
AIC	11707.791	Akaike's information criterion
BIC	11751.669	Bayesian information criterion
Baseline comparison		
CFI	0.996	Comparative fit index
TLI	0.992	Tucker-Lewis index
Size of residuals		
SRMR	0.020	Standardized root mean squared residual
CD	0.988	Coefficient of determination

CAP NUM INS

Use gsem to fit multilevel models and models with binary, ordinal, count, or survival-time outcomes

More complex models can be fit just as easily. We now fit a multilevel CFA model with binary measurements of mathematical ability for students nested in schools.




```
. gsem (MathAb SchQual[school] -> q1 q2 q3 q4), logit
```

Viewer - view sem3.smcl

view sem3.smcl

Dialog | Also see | Jump to

Log likelihood = -1348.3644

(1) [q1]SchQual[school] = 1
(2) [q2]MathAb = 1

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
q1						
SchQual[school]	1	(constrained)				
MathAb	5.277956	4.995708	1.06	0.291	-4.513451	15.06936
_cons	.0413352	.1770215	0.23	0.815	-.3056206	.3882909
q2						
SchQual[school]	.600067	.3447607	1.74	0.082	-.0756516	1.275786
MathAb	1	(constrained)				
_cons	-.449189	.1165887	-3.85	0.000	-.6776987	-.2206793
q3						
SchQual[school]	.3999959	.3008142	1.33	0.184	-.1895891	.989581
MathAb	1.788696	1.10452	1.62	0.105	-.3761236	3.953516
_cons	.1485335	.1070996	1.39	0.165	-.0613779	.3584449
q4						
SchQual[school]	.5925695	.34909	1.70	0.090	-.0916343	1.276773
MathAb	1.071626	.7310121	1.47	0.143	-.3611311	2.504384
_cons	-.3203425	.1152657	-2.78	0.005	-.5462592	-.0944258
var(SchQual[school])	.2483231	.24206			.0367523	1.677838
var(MathAb)	.1050076	.1133871			.0126501	.8716606

CAP NUM INS