Introduction to structural equation modeling using the sem command

Gustavo Sanchez

Senior Econometrician
StataCorp LP

Mexico City, Mexico

Outline

Structural Equation Models (SEM):

- Applications, concepts and components
- Examples
  - Mediation Model
  - Measurement Models
  - SEM Model
  - Other Models

Outline

Structural Equation Models (SEM):

- Applications, concepts and components
- Examples
  - Mediation Model
  - Measurement Models
  - SEM Model
  - Other Models

Outline

Structural Equation Models (SEM):

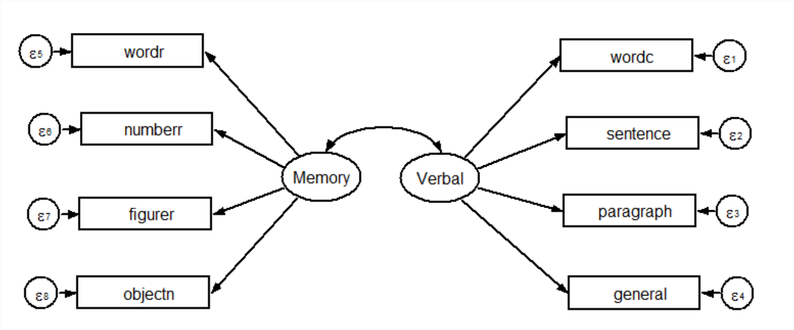
- Applications, concepts and components
- Examples
  - Mediation Model
  - Measurement Models
  - SEM Model
  - Other Models

SEM: *Applications*

- Psychology (e.g. Behavioral analysis, depression)
- Sociology (e.g. Social network, work environment)
- Marketing (e.g. Consumer satisfaction, new products development)
- Academic research (e.g. Analysis of learning abilities)
- Medicine (e.g. Sleep disorders, population health services)
- And more

SEM: *Applications*

- Example: Path diagram for a SEM model



SEM: *Applications*

Models

- Linear regression
- ANOVA
- Multivariate regression
- Simultaneous equation models
- Path analysis
- Simultaneous equation models
- Mediation analysis
- Confirmatory factor analysis
- Reliability estimation
- Full structural equations models
- Multiple indicators and multiple causes (MIMIC)
- Latent growth curve
- Multiple group models

SEM: *Concepts: SEM*

- "**Structural equation modeling** was developed by geneticists (Wright 1921) and economists (Haavelmo 1943; Koopmans 1950, 1953) so that qualitative cause-effect information could be combined with statistical data to provide quantitative assessment of cause-effect relationships among variables of interest" Pearl (2000).
- "**SEM** is a class of statistical techniques used for estimating and testing hypotheses on causal relationships among a set of associated variables"
- "A significant number of models can be expressed as particular cases of structural equation models."

SEM: *Concepts: SEM*

- "**Structural equation modeling** was developed by geneticists (Wright 1921) and economists (Haavelmo 1943; Koopmans 1950, 1953) so that qualitative cause-effect information could be combined with statistical data to provide quantitative assessment of cause-effect relationships among variables of interest" Pearl (2000).
- "**SEM** is a class of statistical techniques used for estimating and testing hypotheses on causal relationships among a set of associated variables"
- "A significant number of models can be expressed as particular cases of structural equation models."

SEM: *Concepts: SEM*

- "**Structural equation modeling** was developed by geneticists (Wright 1921) and economists (Haavelmo 1943; Koopmans 1950, 1953) so that qualitative cause-effect information could be combined with statistical data to provide quantitative assessment of cause-effect relationships among variables of interest" Pearl (2000).
- "**SEM** is a class of statistical techniques used for estimating and testing hypotheses on causal relationships among a set of associated variables"
- "A significant number of models can be expressed as particular cases of structural equation models."

SEM: *Concepts and components: Types of variables*

Observed and Latent

- "A variable is **observed** if it is a variable in your dataset"
- "A variable is **latent** if it is not observed. It is not in your dataset but you wish it were"
- Errors are a special case of latent variables

Exogenous and Endogenous

- "A variable, observed or latent, is **exogenous** (determined outside the system) if paths only originate from it (no path points to it)"
- "A variable, observed or latent, is **endogenous** (determined by the system) if any path points to it."

SEM: *Concepts and components: Types of variables*

Observed and Latent

- "A variable is **observed** if it is a variable in your dataset"
- "A variable is **latent** if it is not observed. It is not in your dataset but you wish it were"
- Errors are a special case of latent variables

Exogenous and Endogenous

- "A variable, observed or latent, is **exogenous** (determined outside the system) if paths only originate from it (no path points to it)"
- "A variable, observed or latent, is **endogenous** (determined by the system) if any path points to it."

SEM: *Concepts and components: Types of variables*

Observed and Latent

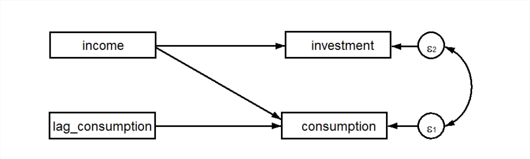
- "A variable is **observed** if it is a variable in your dataset"
- "A variable is **latent** if it is not observed. It is not in your dataset but you wish it were"
- Errors are a special case of latent variables

Exogenous and Endogenous

- "A variable, observed or latent, is **exogenous** (determined outside the system) if paths only originate from it (no path points to it)"
- "A variable, observed or latent, is **endogenous** (determined by the system) if any path points to it."

SEM: *Concepts and components: Path diagrams*

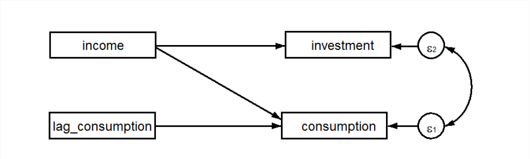
- A Path diagram is a graphical representation of the model
  - Boxes contain observed variables
  - Ovals contain latent variables
  - Circles contain the equation errors
  - Straight arrows represent effects from one variable to another
  - Curved arrows indicate correlation between a pair of variables




```
. sem (income lag_consumption -> consumption)  ///
>     (income -> investment),                  ///
>     cov(e.investment*e.consumption)
```

SEM: *Concepts and components: Path diagrams*

- A Path diagram is a graphical representation of the model
  - Boxes contain observed variables
  - Ovals contain latent variables
  - Circles contain the equation errors
  - Straight arrows represent effects from one variable to another
  - Curved arrows indicate correlation between a pair of variables




```
. sem (income lag_consumption -> consumption)  ///
>     (income -> investment),                  ///
>     cov(e.investment*e.consumption)
```

```
. sem (income lag_consumption -> consumption) (income -> investment), ///
>     cov(e.investment*e.consumption) nolog nodescribe
(1 observations with missing values excluded)
Structural equation model                       Number of obs      =        91
Estimation method  = ml
Log likelihood     = -1932.0358

------------------------------------------------------------------------------------
                   |                 OIM
                   |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
Structural         |
  consumption <-   |
           income  |   .3414577   .0444842     7.68   0.000     .2542703    .4286451
  lag_consumption  |   .6026804   .0531652    11.34   0.000     .4984785    .7068823
            _cons  |   13.01908   2.511575     5.18   0.000     8.096479    17.94167
  -----------------+----------------------------------------------------------------
  investment <-    |
           income  |   .2959361   .0058258    50.80   0.000     .2845177    .3073544
            _cons  |   71.16166   8.916877     7.98   0.000      53.6849    88.63842
-------------------+----------------------------------------------------------------
var(e.consumption) |   108.4106   16.07291                      81.07234    144.9674
 var(e.investment) |   1480.696    219.512                      1107.327    1979.957
-------------------+----------------------------------------------------------------
cov(e.consumption, |
     e.investment) |   30.54975   42.37582     0.72   0.471    -52.50532    113.6048
------------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(1)   =      1.11, Prob > chi2 = 0.2925
```

```
. estat gof
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(1)	1.108	model vs. saturated
p > chi2	0.293	
chi2_bs(5)	1042.569	baseline vs. saturated
p > chi2	0.000	

```
. estat eqgof
```

Equation-level goodness of fit

depvars	Variance fitted	Variance predicted	Variance residual	R-squared	mc	mc2
observed						
consumption	342915.7	342807.3	108.4106	.9996839	.9998419	.9996839
investment	43467.46	41986.77	1480.696	.9659355	.9828202	.9659355
overall				.999688		

mc = correlation between depvar and its prediction
mc2 = mc^2 is the Bentler-Raykov squared multiple correlation coefficient

Example 1: Mediation models

- The explanatory variables may have a direct effect on the outcome and also an indirect effect that is transmitted by a mediator variable
- The traditional mediation analysis was based on a series of linear regressions with no correlated errors (Baron and Kenny (1986))
- With SEM we can fit one simultaneous equation model and get estimates for the indirect and total effects
- The model can be incorporated as part of a larger model

Example 1: Mediation models

- The explanatory variables may have a direct effect on the outcome and also an indirect effect that is transmitted by a mediator variable
- The traditional mediation analysis was based on a series of linear regressions with no correlated errors (Baron and Kenny (1986))
- With SEM we can fit one simultaneous equation model and get estimates for the indirect and total effects
- The model can be incorporated as part of a larger model

Example 1: Mediation models

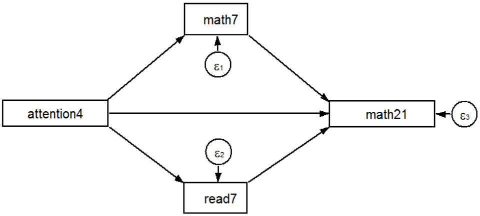
- The explanatory variables may have a direct effect on the outcome and also an indirect effect that is transmitted by a mediator variable
- The traditional mediation analysis was based on a series of linear regressions with no correlated errors (Baron and Kenny (1986))
- With SEM we can fit one simultaneous equation model and get estimates for the indirect and total effects
- The model can be incorporated as part of a larger model

Example 1: Mediation models

- The explanatory variables may have a direct effect on the outcome and also an indirect effect that is transmitted by a mediator variable
- The traditional mediation analysis was based on a series of linear regressions with no correlated errors (Baron and Kenny (1986))
- With SEM we can fit one simultaneous equation model and get estimates for the indirect and total effects
- The model can be incorporated as part of a larger model

Example 1: Mediation models

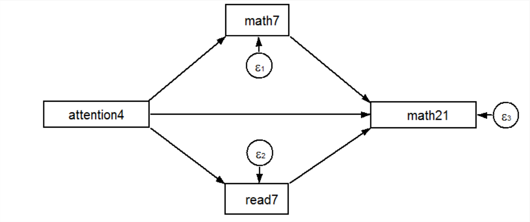
- Example from Alan Acock (2013)
- The researcher wants to analyze whether children with better span skills at four have advantages in their academic development for math at a later age
- Variables
  - attention4: span of attention at 4
  - math7: performance on math at 7
  - read7: performance on reading at 7
  - math21: performance on math at 21




```
. sem (math7 <- attention4) (read7 <- attention4) (math21 <- attention4 math7 read7)
```

Example 1: Mediation models

- Example from Alan Acock (2013)
- The researcher wants to analyze whether children with better span skills at four have advantages in their academic development for math at a later age
- Variables
  - attention4: span of attention at 4
  - math7: performance on math at 7
  - read7: performance on reading at 7
  - math21: performance on math at 21




```
. sem (math7 <- attention4)  (read7 <- attention4) (math21 <- attention4 math7 read7)
```

```
. sem (math7 <- attention4) (read7 <- attention4) (math21 <- attention4 math7 read7), ///
>     nolog noheader nodescribe
(92 observations with missing values excluded)
```

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
math7 <-						
attention4	.1129353	.0493128	2.29	0.022	.016284	.2095865
_cons	8.719295	.899051	9.70	0.000	6.957187	10.4814
read7 <-						
attention4	.3024263	.1441203	2.10	0.036	.0199556	.5848969
_cons	26.41214	2.627544	10.05	0.000	21.26225	31.56204
math21 <-						
math7	.2938502	.0475641	6.18	0.000	.2006263	.3870741
read7	.0825399	.0162747	5.07	0.000	.0506421	.1144378
attention4	.0813543	.0421041	1.93	0.053	-.0011683	.1638769
_cons	3.987873	.9165827	4.35	0.000	2.191404	5.784342
var(e.math7)	7.841701	.6032078			6.744244	9.117743
var(e.read7)	66.97947	5.152267			57.6056	77.87871
var(e.math21)	5.58987	.42999			4.80756	6.499482

```
LR test of model vs. saturated: chi2(1)   =     23.72, Prob > chi2 = 0.0000
```

```
. estat teffects,compact nodirect
```

Indirect effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
math21 <-						
attention4	.0581483	.0197686	2.94	0.003	.0194026	.096894

Total effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
math7 <-						
attention4	.1129353	.0493128	2.29	0.022	.016284	.2095865
read7 <-						
attention4	.3024263	.1441203	2.10	0.036	.0199556	.5848969
math21 <-						
math7	.2938502	.0475641	6.18	0.000	.2006263	.3870741
read7	.0825399	.0162747	5.07	0.000	.0506421	.1144378
attention4	.1395026	.045661	3.06	0.002	.0500086	.2289966

Example 2.1: Measurement model - one factor

- The researcher is interested in a latent variable (e.g. Consumer satisfaction, verbal abilities, alienation)
- The model specifies the relation between latent variables and measured indicator variables
- Modification indices are normally used to refine the model
- The model can also be incorporated as part of a larger model

Example 2.1: Measurement model - one factor

- The researcher is interested in a latent variable (e.g. Consumer satisfaction, verbal abilities, alienation)
- The model specifies the relation between latent variables and measured indicator variables
- Modification indices are normally used to refine the model
- The model can also be incorporated as part of a larger model

Example 2.1: Measurement model - one factor

- The researcher is interested in a latent variable (e.g. Consumer satisfaction, verbal abilities, alienation)
- The model specifies the relation between latent variables and measured indicator variables
- Modification indices are normally used to refine the model
- The model can also be incorporated as part of a larger model

Example 2.1: Measurement model - one factor

- The researcher is interested in a latent variable (e.g. Consumer satisfaction, verbal abilities, alienation)
- The model specifies the relation between latent variables and measured indicator variables
- Modification indices are normally used to refine the model
- The model can also be incorporated as part of a larger model

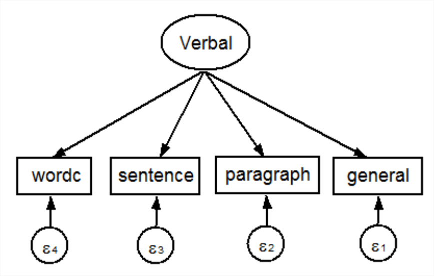
Example 2: Measurement model - one factor

- Example with Holzinger and Swineford (1939) Data

  The researcher wants to analyze the verbal ability based on indices associated to tests on word classification, sentence completion and paragraph comprehension

  Variables:

Verbal:	latent variable for verbal ability
wordc:	scores on word classification test
sentence:	scores on sentence completion test
paragraph:	scores on paragraph comprehension test
general:	scores on general information test



```
. sem (Verbal -> wordc sentence paragraph general)
```

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Measurement						
wordc <-						
Verbal	1	(constrained)				
_cons	26.12625	.3265833	80.00	0.000	25.48615	26.76634
sentence <-						
Verbal	1.080072	.0698177	15.47	0.000	.9432322	1.216912
_cons	17.36213	.2970317	58.45	0.000	16.77995	17.9443
paragraph <-						
Verbal	.6603575	.0471812	14.00	0.000	.5678841	.7528309
_cons	9.182724	.200961	45.69	0.000	8.788848	9.576601
general <-						
Verbal	2.351367	.1645309	14.29	0.000	2.028892	2.673842
_cons	40.59136	.7124269	56.98	0.000	39.19503	41.98769
var(e.wordc)	13.8878	1.336773			11.50008	16.77127
var(e.sentence)	5.306741	.813477			3.92958	7.166542
var(e.paragr~h)	4.21255	.4386365			3.43489	5.166273
var(e.general)	52.05898	5.597825			42.16646	64.27234
var(Verbal)	18.21586	2.46415			13.97345	23.7463

```
. sem (Verbal -> wordc sentence paragraph general), standardize
```

Standardized	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Measurement						
wordc <-						
Verbal	.7532647	.0289445	26.02	0.000	.6965345	.8099949
_cons	4.611049	.1965727	23.46	0.000	4.225774	4.996324
sentence <-						
Verbal	.8945235	.0185182	48.31	0.000	.8582284	.9308185
_cons	3.369123	.1489219	22.62	0.000	3.077242	3.661005
paragraph <-						
Verbal	.8083679	.0243277	33.23	0.000	.7606865	.8560492
_cons	2.633762	.1218401	21.62	0.000	2.39496	2.872565
general <-						
Verbal	.811936	.0245424	33.08	0.000	.7638338	.8600382
_cons	3.284053	.1457311	22.54	0.000	2.998425	3.56968
var(e.wordc)	.4325923	.0436058			.3550395	.5270852
var(e.sentence)	.1998278	.03313			.1443886	.2765533
var(e.paragr~h)	.3465414	.0393314			.2774255	.4328764
var(e.general)	.34076	.0398537			.2709543	.4285496
var(Verbal)	1	.			.	.

```
. sem (Verbal -> wordc sentence paragraph general), standardize nomeans
```

Standardized	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Measurement						
wordc <-						
Verbal	.7532647	.0289445	26.02	0.000	.6965345	.8099949
sentence <-						
Verbal	.8945235	.0185182	48.31	0.000	.8582284	.9308185
paragraph <-						
Verbal	.8083679	.0243277	33.23	0.000	.7606865	.8560492
general <-						
Verbal	.811936	.0245424	33.08	0.000	.7638338	.8600382
var(e.wordc)	.4325923	.0436058			.3550395	.5270852
var(e.sentence)	.1998278	.03313			.1443886	.2765533
var(e.paragr~h)	.3465414	.0393314			.2774255	.4328764
var(e.general)	.34076	.0398537			.2709543	.4285496
var(Verbal)	1	.			.	.

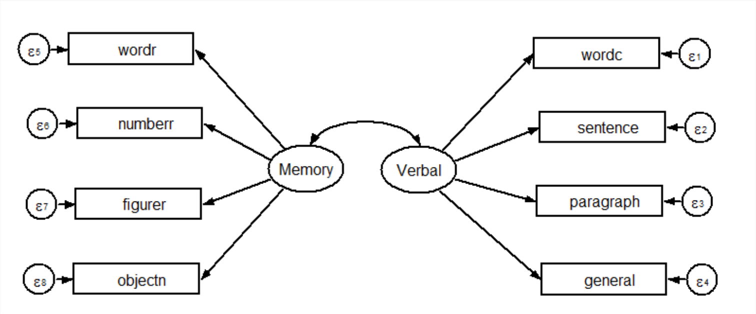
Example 2.2: Measurement model - two factors

- Holzinger and Swineford (1939) Data
- Variables
  - **Verbal**: latent variable for verbal ability
  - **wordc**: scores on word classification test
  - **sentence**: scores on sentence completion test
  - **paragraph**: scores on paragraph comprehension test
  - **general**: scores on general information test
  - **Memory**: latent variable for Memory condition
  - **wordr**: scores on word recognition test
  - **number**: scores on number recognition test
  - **figurer**: scores on figurer recognition test
  - **object**: scores on object-number test

Example 2.2: Measurement model - two factors

- Holzinger and Swineford (1939) Data
- Variables
  - **Verbal**: latent variable for verbal ability
  - **wordc**: scores on word classification test
  - **sentence**: scores on sentence completion test
  - **paragraph**: scores on paragraph comprehension test
  - **general**: scores on general information test
  - **Memory**: latent variable for Memory condition
  - **wordr**: scores on word recognition test
  - **number**: scores on number recognition test
  - **figurer**: scores on figurer recognition test
  - **object**: scores on object-number test

Example 2.2: Measurement model - two factors



```
. sem (Verbal -> wordc sentence paragraph general)            ///
>     (Memory -> wordr numberr figurer objectn), standardize  ///
>     cov( Memory*Verbal) nomeans noheader nocnsreport nolog
Endogenous variables
Measurement:  wordc sentence paragraph general wordr numberr figurer objectn
Exogenous variables
Latent:       Verbal Memory
```

Standardized	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Measurement						
wordc <-						
Verbal	.7558212	.0287657	26.28	0.000	.6994414	.812201
sentence <-						
Verbal	.890738	.0186103	47.86	0.000	.8542625	.9272136
paragraph <-						
Verbal	.8118064	.0241118	33.67	0.000	.7645481	.8590646
general <-						
Verbal	.8114395	.02443	33.21	0.000	.7635576	.8593214

wordr <-						
Memory	.697023	.0530879	13.13	0.000	.5929727	.8010734
numberr <-						
Memory	.5658826	.0555916	10.18	0.000	.456925	.6748401
figurer <-						
Memory	.5741969	.0559689	10.26	0.000	.4645	.6838939
objectn <-						
Memory	.4994731	.0575943	8.67	0.000	.3865904	.6123558
var(e.wordc)	.4287343	.0434835			.3514445	.5230216
var(e.sentence)	.2065857	.0331538			.1508327	.282947
var(e.paragraph)	.3409704	.0391482			.2722618	.4270185
var(e.general)	.3415659	.0396469			.2720646	.4288219
var(e.wordr)	.5141589	.074007			.3877726	.681738
var(e.numberr)	.6797769	.0629166			.5670007	.8149843
var(e.figurer)	.6702979	.0642743			.5554524	.8088889
var(e.objectn)	.7505266	.0575336			.6458253	.8722022
var(Verbal)	1	.			.	.
var(Memory)	1	.			.	.
cov(Verbal,Memory)	.2579165	.0705776	3.65	0.000	.1195869	.3962461

LR test of model vs. saturated: chi2(19) = 41.13, Prob > chi2 = 0.0023

```
. estat mindices

Modification indices
```

	MI	df	P>MI	EPC	Standard EPC
Measurement					
sentence <-					
Memory	5.272	1	0.02	-.064809	-.1007062
paragraph <-					
Memory	6.152	1	0.01	.0518729	.1191386
numberr <-					
Verbal	8.098	1	0.00	-.3086542	-.1712958
figurer <-					
Verbal	5.759	1	0.02	.2570111	.144435
cov(e.wordc,e.paragraph)	4.056	1	0.04	-1.228282	-.1626198
cov(e.wordc,e.figurer)	9.964	1	0.00	4.905813	.2119483
cov(e.sentence,e.numberr)	4.560	1	0.03	-2.528942	-.1697062

```
EPC = expected parameter change
```

- Based on **estat mindices**, let's add *figurer* and *numberr* to the equation for *verbal*, and also cov(e.wordc*e.paragraph): (but we should first check the theoretical framework)

```
sem (Verbal -> wordc sentence paragraph general figurer numberr )      ///
    (Memory -> wordr numberr figurer objectn), standardize nocnsreport ///
    nolog nomeans cov( Memory*Verbal) noheader cov(e.wordc*e.paragraph)
```

- Compare the **chi2_ms** before and after adding the suggestions from estat mindices:

```
. estat gof    /* Before adding suggestions from estat mindices */
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(19)	41.132	model vs. saturated
p > chi2	0.002	

```
. estat gof /* After adding suggestions from estat mindices */
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(16)	25.107	model vs. saturated
p > chi2	0.068	

- Based on **estat mindices**, let's add *figurer* and *numberr* to the equation for *verbal*, and also cov(e.wordc*e.paragraph): (but we should first check the theoretical framework)

```
sem (Verbal -> wordc sentence paragraph general figurer numberr )      ///
    (Memory -> wordr numberr figurer objectn), standardize nocnsreport ///
   nolog nomeans cov( Memory*Verbal) noheader cov(e.wordc*e.paragraph)
```

- Compare the **chi2_ms** before and after adding the suggestions from estat mindices:

```
. estat gof    /* Before adding suggestions from estat mindices */
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(19)	41.132	model vs. saturated
p > chi2	0.002	

```
. estat gof /* After adding suggestions from estat mindices */
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(16)	25.107	model vs. saturated
p > chi2	0.068	

- Based on **estat mindices**, let's add *figurer* and *numberr* to the equation for *verbal*, and also cov(e.wordc*e.paragraph): (but we should first check the theoretical framework)

```
sem (Verbal -> wordc sentence paragraph general figurer numberr )      ///
    (Memory -> wordr numberr figurer objectn), standardize nocnsreport ///
    nolog nomeans cov( Memory*Verbal) noheader cov(e.wordc*e.paragraph)
```

- Compare the **chi2_ms** before and after adding the suggestions from estat mindices:

```
. estat gof    /* Before adding suggestions from estat mindices */
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(19)	41.132	model vs. saturated
p > chi2	0.002	

```
. estat gof /* After adding suggestions from estat mindices */
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(16)	25.107	model vs. saturated
p > chi2	0.068	

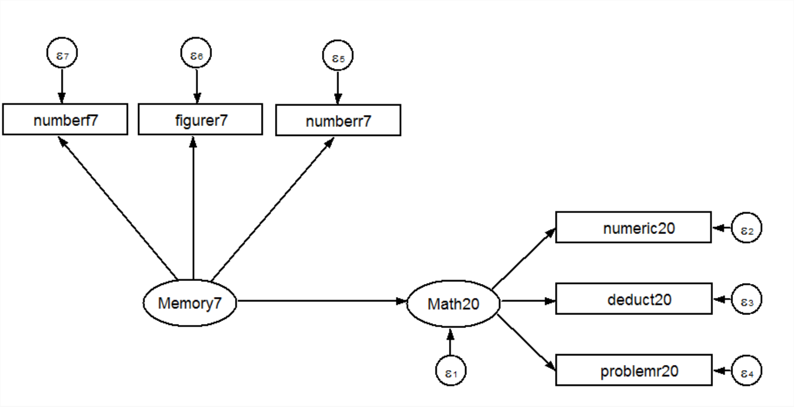
Example 3: Structural equation model

- The model Integrates two components
  - A structural component that specifies the relationship among the latent variables
  - A measurement component that specifies the relationship between the latent variables and the corresponding indicators
- Modification indices are normally used to refine the model
- The model can incorporate causal relationships among the latent variables

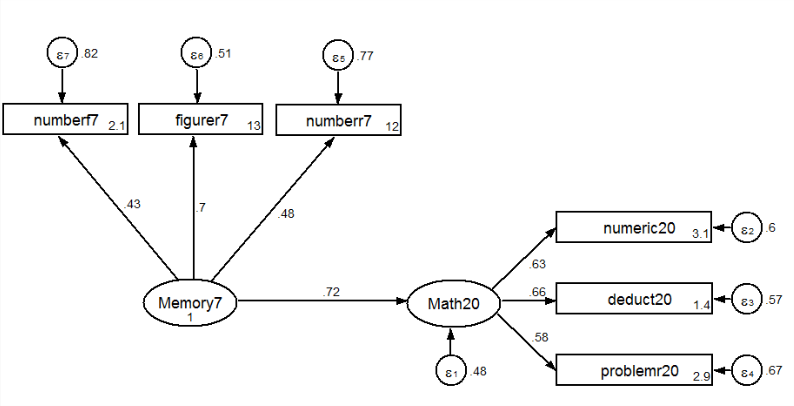
Example 3: Structural equation model

- The model Integrates two components
  - A structural component that specifies the relationship among the latent variables
  - A measurement component that specifies the relationship between the latent variables and the corresponding indicators
- Modification indices are normally used to refine the model
- The model can incorporate causal relationships among the latent variables

Example 3: Structural equation model (Fictitious data)



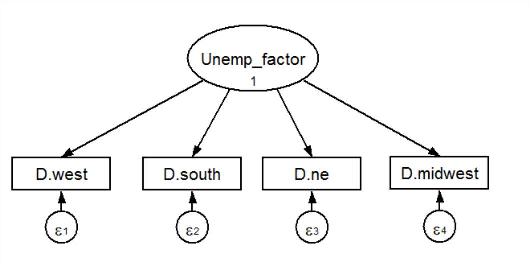
Example 3: Structural equation model (Fictitious data)



Example 4: Dynamic Factor

- **Unobserved (factor) effect on unemployment in four different regions**

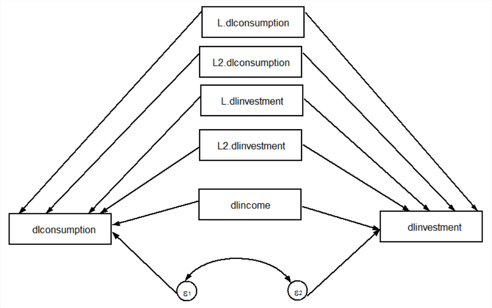
```
use http://www.stata-press.com/data/r13/urate,clear
dfactor (D.(west south ne midwest) = ,  ) (Unemp_factor = )
sem (D.(west south ne midwest) <- Unemp_factor),var(Unemp_factor@1)
```



Example 5: VAR model

- **VAR model** for consumption and investment with income as an exogenous variable

```
use http://www.stata-press.com/data/r13/lutkepohl,clear
var dlconsumption dlinvestment,lags(1/2) exog(dlincome)
sem (dlconsumption <- l.dlconsumption l2.dlconsumption          ///
                      l.dlinvestment l2.dlinvestment dlincome)  ///
    (dlinvestment   <- l.dlconsumption l2.dlconsumption         ///
                      l.dlinvestment l2.dlinvestment dlincome), ///
    cov(e.dlinvestment*e.dlconsumption)
```



Summary

Structural Equation Models (SEM):

- Applications, Concepts and components
- Examples
  - Mediation Model
  - Measurement Model
  - SEM Model
  - Other Models
- What is next?
  - **Generalization (gsem)**
    - Extensions for nonlinear models
    - Multilevel models and more

Summary

Structural Equation Models (SEM):

- Applications, Concepts and components
- Examples
  - Mediation Model
  - Measurement Model
  - SEM Model
  - Other Models
- What is next?
  - **Generalization (gsem)**
    - Extensions for nonlinear models
    - Multilevel models and more

References

Acock, Alan 2013. Discovering structural equation modeling using Stata. Stata Press

Holzinger, K. J. and F. Swineford, 1939. A study in factor analysis: The stability of a bi-factor solution. Supplementary Education Monographs, 48. Chicago, IL: University of Chicago.

Haavelmo, T. 1943. The statistical implications of a system of simultaneous equations. Econometrica 11 (1), 1—12.

Koopmans, T. C. 1950. When is an Equation System Complete for Statistical Purposes? (in Statistical Inference in Dynamic Economic Models, T. C. Koopmans (ed.), Cowles Commission Monograph 10, Wiley, New York, 1950, pp. 393—409) pp. 527—537

Koopmans, T.C. 1953 Identification Problems in Economic Model Construction (in Studies in Econometric Method, Cowles Commission Monograph 14), New Haven, Yale University Press

Pearl J. 2000 Causality: Models, Reasoning, and Inference. Cambridge University Press.

Wright, Sewall S. 1921. Correlation and causation. Journal of Agricultural Research 20: 557—85.

Questions - Comments