Psychology 454: Latent Variable Modeling

Using lavaan to model structure

Department of Psychology
Northwestern University
Evanston, Illinois USA

February, 2011

Outline

CFA confirms structure; SEM tests for "causality"

- We have examined the use of Confirmatory Factor Models for testing:
  - quality of measures (e.g., congeneric measurement models)
  - equivalence of measures over time or form (e.g., factor invariance models).
- Structural models combine measurement models with regression models
  - Regressions may be of observed variables but more typically, of latent variables.
  - In this case, the modeling is essentially just correcting for attenuation.
  - But can also add in additional fitting parameters to estimate correlations between residuals.
- Whenever discussing "causality" we need to remember the interchangeability of directional arrows.
  - Causality is not tested in any particular model. Causality reflects relative plausibility of alternative models.

Examples from lavaan version .4-6

.4-6 was released 2/14/11

```
install.packages("lavaan", repos="http://www.da.ugent.be", type="source")
library(lavaan)
data(``PoliticalDemocracy')'
```

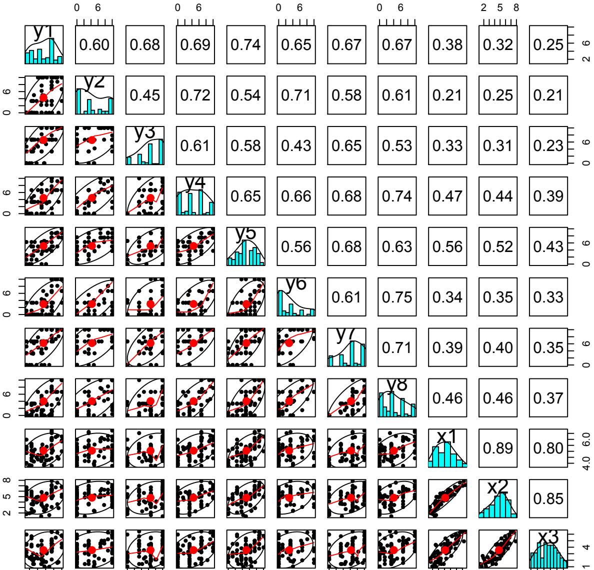
```
Industrialization And Political Democracy Dataset
Description
The "famous" Industrialization and Political Democracy dataset. This dataset is used throughout
Bollen's 1989 book (see pages 12, 17, 36 in chapter 2, pages 228 and following in chapter 7,
pages 321 and following in chapter 8).
The dataset contains various measures of political democracy and industrialization in
developing countries.

Usage
A data frame of 75 observations of 11 variables.

y1  Expert ratings of the freedom of the press in 1960
y2  The freedom of political opposition in 1960
y3  The fairness of elections in 1960
y4  The effectiveness of the elected legislature in 1960
y5  Expert ratings of the freedom of the press in 1965
y6  The freedom of political opposition in 1965
y7  The fairness of elections in 1965
y8   The effectiveness of the elected legislature in 1965
x1  The gross national product (GNP) per capita in 1960
x2  The inanimate energy consumption per capita in 1960
x3  The percentage of the labor force in industry in 1960
```

Basic descriptive statistics

```
> describe(PoliticalDemocracy)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
y1	1	75	5.46	2.62	5.40	5.46	3.11	1.25	10.00	8.75	-0.09	-1.10	0.30
y2	2	75	4.26	3.95	3.33	4.09	4.94	0.00	10.00	10.00	0.32	-1.44	0.46
y3	3	75	6.56	3.28	6.67	6.92	4.94	0.00	10.00	10.00	-0.59	-0.62	0.38
y4	4	75	4.45	3.35	3.33	4.33	4.94	0.00	10.00	10.00	0.12	-1.16	0.39
y5	5	75	5.14	2.61	5.00	5.23	3.71	0.00	10.00	10.00	-0.23	-0.68	0.30
y6	6	75	2.98	3.37	2.23	2.51	3.31	0.00	10.00	10.00	0.89	-0.34	0.39
y7	7	75	6.20	3.29	6.67	6.47	4.94	0.00	10.00	10.00	-0.55	-0.63	0.38
y8	8	75	4.04	3.25	3.33	3.82	4.40	0.00	10.00	10.00	0.45	-0.88	0.37
x1	9	75	5.05	0.73	5.08	5.03	0.82	3.78	6.74	2.95	0.25	-0.66	0.08
x2	10	75	4.79	1.51	4.96	4.85	1.53	1.39	7.87	6.49	-0.35	-0.46	0.17
x3	11	75	3.56	1.41	3.57	3.53	1.51	1.00	6.42	5.42	0.08	-0.86	0.16

A splom plot of the democracy data



The LISREL model of the data – note the correlated residuals

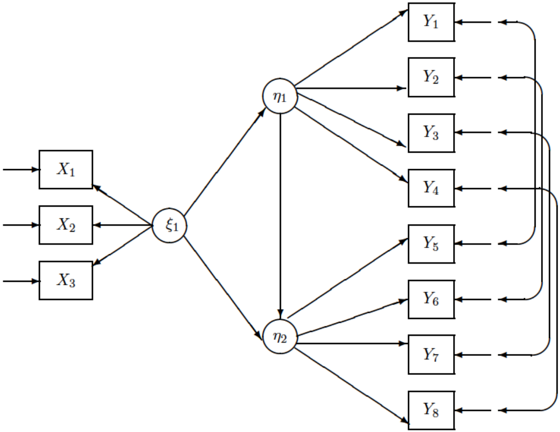


Figure 2: Panel Model of Democracy and Industrialization

http://www.ssicentral.com/lisrel/techdocs/obsres.pdf

Fitting as a simple measurement model – no structure

```
f3  <- fa(PoliticalDemocracy,3)
```

```
Factor Analysis using method =  minres
Call: fa(r = PoliticalDemocracy, nfactors = 3)
Standardized loadings based upon correlation matrix
     MR1   MR2   MR3   h2    u2
y1 -0.05  0.20  0.75 0.78 0.219
y2 -0.09  0.76  0.10 0.66 0.345
y3 -0.01 -0.08  0.84 0.61 0.388
y4  0.12  0.48  0.37 0.72 0.278
y5  0.24  0.09  0.65 0.71 0.290
y6  0.04  0.91 -0.06 0.78 0.217
y7  0.07  0.31  0.53 0.66 0.336
y8  0.14  0.62  0.21 0.73 0.269
x1  0.90 -0.09  0.15 0.87 0.133
x2  0.96  0.01  0.00 0.94 0.065
x3  0.90  0.09 -0.13 0.78 0.223

                 MR1  MR2  MR3
SS loadings     2.80 2.71 2.73
Proportion Var 0.25 0.25 0.25
Cumulative Var 0.25 0.50 0.75

 With factor correlations of
     MR1  MR2  MR3
MR1 1.00 0.38 0.41
MR2 0.38 1.00 0.72
MR3 0.41 0.72 1.00
```

This is not the desired model. For it groups by content rather than time.

With goodness of fit statistics

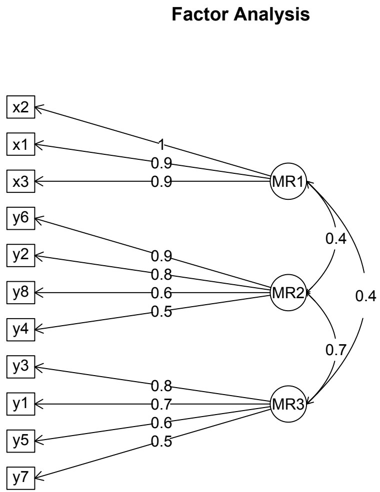
Although it fits pretty well.

```
Test of the hypothesis that 3 factors are sufficient.

The degrees of freedom for the null model are  55  and the objective function was  9.74
            with Chi Square of  677.07
The degrees of freedom for the model are 25  and the objective function was  0.43

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.03
The number of observations was  75  with Chi Square =  29.01  with prob <  0.26

Tucker Lewis Index of factoring reliability =  0.985
RMSEA index =  0.061  and the 90 % confidence intervals are  0.059 0.069
BIC =  -78.93
Fit based upon off diagonal values = 1
Measures of factor score adequacy
                                                MR1  MR2  MR3
Correlation of scores with factors             0.98 0.95 0.95
Multiple R square of scores with factors       0.96 0.90 0.90
Minimum correlation of possible factor scores  0.92 0.81 0.79
```

But this is not the model we wanted to fit!

The exploratory factor solution is pretty, but not what we want

Factors are forming based upon content rather than temporal relation.



Two ways of estimating χ^2

- There are two ways to calculate χ^2, both are available in *lavaan* as options
  - The sem, LISREL, EQS, or AMOS way is to use a "Wishart" distribution based upon covariance matrix calculated with N-1 observations.
  - The MPlus way is to use N. (The normal approach to maximum likelihood estimation.)
    - "'N' is used as a multiplier to compute both standard errors and test statistics. One advantage of the 'normal' approach is that it extends naturally to the missing-data case. The 'wishart' approach does not apply to missing data, and software programs tend to 'switch' to the normal approach if any data is missing." (Rosseel, personal communication, 2/14/11).
- As of 2/14/11, (version .4-6) `cfa` and `sem` can do either approach. Specify the likehood="wishart" or "normal".

CFA model of the data - is not very good

```
mod.cfa <- '
# latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + y6 + y7 + y8'
 cfa.fit <- cfa(mod.cfa,data=PoliticalDemocracy,
                      std.ov=TRUE,std.lv=TRUE)
 summary(cfa.fit,fit.measures=TRUE)
```

```
model converged normally after 38 iterations using ML

  Minimum Function Chi-square                   72.462
  Degrees of freedom                                41
  P-value                                       0.0018

Chi-square test baseline model:

  Minimum Function Chi-square                  730.654
  Degrees of freedom                                55
  P-value                                       0.0000
```

```
Full model versus baseline model:

  Comparative Fit Index (CFI)                    0.953
  Tucker-Lewis Index (TLI)                       0.938

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)               -835.991
  Loglikelihood unrestricted model (H1)       -799.760

  Akaike (AIC)                                1721.982
  Bayesian (BIC)                              1779.919

Root Mean Square Error of Approximation:

  RMSEA                                          0.101
  90 Percent Confidence Interval          0.061  0.139
  P-value RMSEA <= 0.05                          0.021

Standardized Root Mean Square Residual:

  SRMR                                           0.055
```

With parameters of

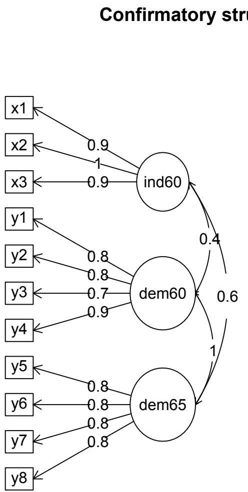
```
                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  ind60 =~
    x1                0.913    0.088   10.339    0.000
    x2                0.967    0.085   11.428    0.000
    x3                0.866    0.091    9.477    0.000
  dem60 =~
    y1                0.750    0.086    8.676    0.000
    y2                0.675    0.090    7.462    0.000
    y3                0.626    0.093    6.747    0.000
    y4                0.763    0.086    8.902    0.000
  dem65 =~
    y5                0.130    0.113    1.150    0.250
    y6                0.126    0.110    1.149    0.251
    y7                0.132    0.115    1.150    0.250
    y8                0.137    0.119    1.150    0.250

Regressions:
  dem60 ~
    ind60             0.501    0.143    3.511    0.000
  dem65 ~
    ind60             0.895    0.827    1.082    0.279
    dem60             5.017    4.485    1.119    0.263
```

```
Variances:
    x1                0.152    0.036    4.180    0.000
    x2                0.052    0.031    1.689    0.091
    x3                0.236    0.046    5.174    0.000
    y1                0.282    0.057    4.910    0.000
    y2                0.417    0.076    5.479    0.000
    y3                0.496    0.088    5.662    0.000
    y4                0.257    0.054    4.731    0.000
    y5                0.350    0.065    5.351    0.000
    y6                0.382    0.070    5.456    0.000
    y7                0.325    0.062    5.252    0.000
    y8                0.279    0.056    5.019    0.000
    ind60             1.000
    dem60             1.000
    dem65             1.000
```

Find the fits with "Wishart" distribution

```
mod.cfa <- '
# latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + y6 + y7 + y8'
 cfa.fit <- cfa(mod.cfa,data=PoliticalDemocracy,
                      std.ov=TRUE,std.lv=TRUE,
                      likelihood='wishart')
 summary(cfa.fit,fit.measures=TRUE)

Lavaan (0.4-6) converged normally after 36 iterations

  Number of observations                            75

  Estimator                                         ML
  Minimum Function Chi-square                   71.495
  Degrees of freedom                                41
  P-value                                        0.002

Chi-square test baseline model:

  Minimum Function Chi-square                  720.912
  Degrees of freedom                                55
  P-value                                        0.000

Full model versus baseline model:

  Comparative Fit Index (CFI)                    0.954
  Tucker-Lewis Index (TLI)                       0.939

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)               -841.528
  Loglikelihood unrestricted model (H1)       -805.297

  Number of free parameters                         25
  Akaike (AIC)                                1733.056
  Bayesian (BIC)                              1790.993
  Sample-size adjusted Bayesian (BIC)         1712.200

Root Mean Square Error of Approximation:

  RMSEA                                          0.100
  90 Percent Confidence Interval          0.059  0.137
  P-value RMSEA <= 0.05                          0.025

Standardized Root Mean Square Residual:

  SRMR                                           0.055
```

CFA of the model



Structural models versus Confirmatory Factor models

- The CFA models are just glorified factor models – pure measurement.
  - Typically with fewer parameters because the cross loadings are ignored.
  - Cross loadings can be specified if needed (because of poor fit).
- Structural models combine the measurement models of cfa with regression.
  - The use of regression implies causal paths.
  - Always need to examine direction of these paths. Is the model plausible.

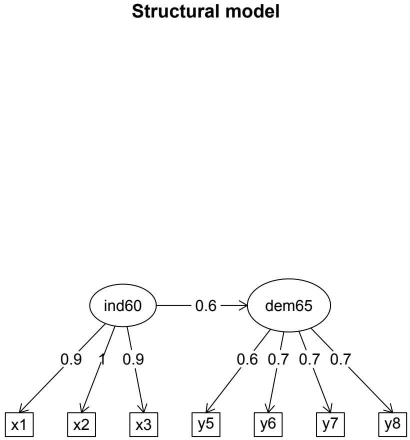
The most simple model: Industrialization -> Democracy

```
mod.sem0 <- '
# latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem65 =~ y5 + y6 + y7 + y8
     # regressions
    dem65 ~ ind60 '

sem.fit <- cfa(mod.sem0,data=PoliticalDemocracy,
                       std.ov=TRUE,std.lv=TRUE)
 summary(sem.fit,fit.measures=TRUE)
```

```
Lavaan (0.4-6) converged normally after 28 iterations

  Number of observations                            75
  Estimator                                         ML
  Minimum Function Chi-square                   20.862
  Degrees of freedom                                13
  P-value                                        0.076
Chi-square test baseline model:
  Minimum Function Chi-square                  426.897
  Degrees of freedom                                21
  P-value                                        0.000

Full model versus baseline model:
  Comparative Fit Index (CFI)                    0.981
  Tucker-Lewis Index (TLI)                       0.969
  RMSEA                                          0.090
  90 Percent Confidence Interval          0.000  0.158
  P-value RMSEA <= 0.05                          0.173
Standardized Root Mean Square Residual:
  SRMR                                           0.054
```

The most simple model: Industrialization leads to Democracy

lavaan.diagram(sem.fit,lr=FALSE,e.size=.2)



Consider a simple model, one causal factor, two dependent time points

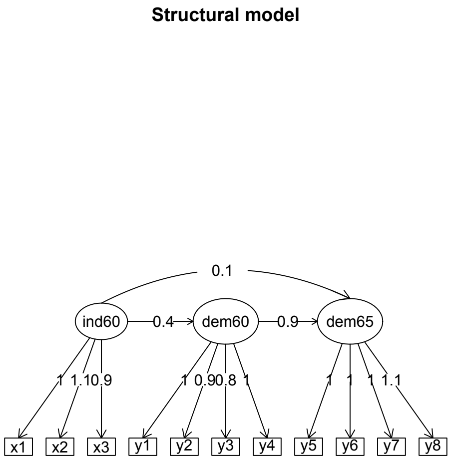
```
mod.sem1 <- '
# latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + y6 + y7 + y8
     # regressions
    dem60 ~ ind60
    dem65 ~ ind60 + dem60'
sem.fit <- cfa(mod.sem1,data=PoliticalDemocracy,std.ov=TRUE)
summary(sem.fit,fit.measures=TRUE)
lavaan.diagram(sem.fit,cut=0,lr=FALSE)
```

```
lavaan (0.4-6) converged normally after 34 iterations
  Number of observations                            75
  Estimator                                         ML
  Minimum Function Chi-square                   72.462
  Degrees of freedom                                41
  P-value                                        0.002
Chi-square test baseline model:
  Minimum Function Chi-square                  730.654
  Degrees of freedom                                55
  P-value                                        0.000
Full model versus baseline model:
  Comparative Fit Index (CFI)                    0.953
  Tucker-Lewis Index (TLI)                       0.938
Root Mean Square Error of Approximation:
  RMSEA                                          0.101
  90 Percent Confidence Interval          0.061  0.139
  P-value RMSEA <= 0.05                          0.021
Standardized Root Mean Square Residual:
  SRMR                                           0.055
```

Examine the residuals

```
> resid(sem.fit,type="normalized")

$cov
   x1     x2     x3     y1     y2     y3     y4     y5     y6     y7     y8
x1  0.000
x2 -0.004  0.000
x3 -0.018  0.012  0.000
y1  0.274 -0.389 -0.698  0.000
y2 -0.840 -0.685 -0.731 -0.285  0.000
y3  0.301  0.045 -0.425  0.593 -0.672  0.000
y4  0.896  0.537  0.438 -0.237  0.462  0.017  0.000
y5  1.170  0.685  0.339  0.524 -0.409  0.167 -0.173  0.000
y6 -0.445 -0.558 -0.388  0.018  0.870 -0.898  0.003 -0.483  0.000
y7 -0.232 -0.323 -0.361 -0.017 -0.209  0.618 -0.060  0.145 -0.239  0.000
y8  0.249 -0.012 -0.314 -0.242 -0.178 -0.417  0.174 -0.370  0.625  0.130  0.000

$mean
x1 x2 x3 y1 y2 y3 y4 y5 y6 y7 y8
 0  0  0  0  0  0  0  0  0  0  0
```

Show the structure



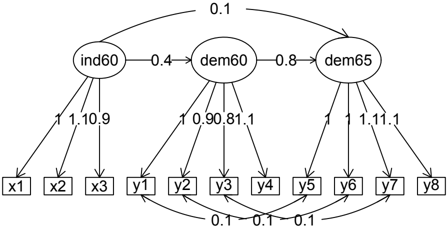
Correlate some of the residuals

```
mod.sem2 <- '
# latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + y6 + y7 + y8
     # regressions
    dem60 ~ ind60
    dem65 ~ ind60 + dem60
     y1 ~~ y5
    y2 ~~ y6
    y3 ~~ y7
    y4 ~~ y8'
fit2 <- sem(mod.sem2, data=PoliticalDemocracy,std.ov=TRUE)
summary(fit2, fit.measures=TRUE)
lavaan.diagram(fit2,cut=0,lr=FALSE)
```

```
  Minimum Function Chi-square                     50.835
  Degrees of freedom                                  37
  P-value                                          0.064
Chi-square test baseline model:
  Minimum Function Chi-square                    730.654
  Degrees of freedom                                  55
  P-value                                          0.000
Full model versus baseline model:
  Comparative Fit Index (CFI)                      0.980
  Tucker-Lewis Index (TLI)                         0.970
Root Mean Square Error of Approximation:
  RMSEA                                            0.071
  90 Percent Confidence Interval          0.000    0.115
  P-value RMSEA <= 0.05                            0.234
Standardized Root Mean Square Residual:
  SRMR                                             0.050
```

Correlated residuals

Structural model



Examine the residuals

resid(fit2)

```
> resid(fit2)
$cov
     x1     x2     x3     y1     y2     y3     y4     y5     y6     y7     y8
x1  0.000
x2 -0.001  0.000
x3 -0.003  0.002  0.000
y1  0.042 -0.037 -0.073  0.001
y2 -0.097 -0.080 -0.085 -0.021  0.001
y3  0.041  0.011 -0.045  0.106 -0.074  0.003
y4  0.104  0.058  0.045 -0.032  0.049 -0.003 -0.001
y5  0.165  0.100  0.053  0.004 -0.013  0.067  0.005  0.002
y6 -0.060 -0.074 -0.053  0.033 -0.007 -0.090  0.000 -0.049 -0.006
y7 -0.030 -0.041 -0.045  0.038 -0.006 -0.005 -0.001  0.042 -0.039 -0.004
y8  0.018 -0.016 -0.051 -0.009 -0.017 -0.039  0.000 -0.045  0.064  0.000  0.001

$mean
x1 x2 x3 y1 y2 y3 y4 y5 y6 y7 y8
 0  0  0  0  0  0  0  0  0  0  0
```

Consider the modification indices

```
modificationIndices(fit2)
```

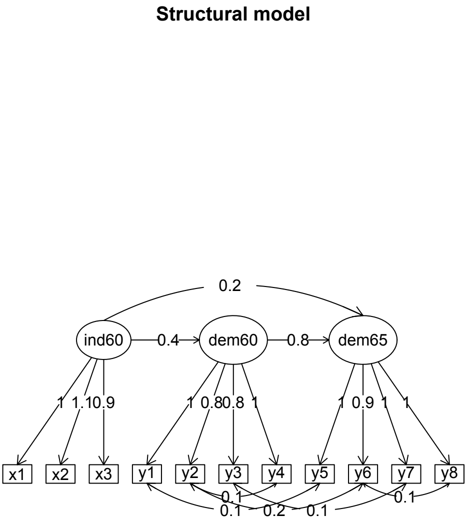
```
      lhs op   rhs    mi    epc sepc.lv sepc.all
1   ind60 =~    x1    NA     NA      NA       NA
2   ind60 =~    x2 0.000  0.000   0.000    0.000
3   ind60 =~    x3 0.000  0.000   0.000    0.000
4   ind60 =~    y1 1.951 -0.128  -0.117   -0.118
...
7   ind60 =~    y4 3.760  0.179   0.163    0.164
8   ind60 =~    y5 6.346  0.264   0.241    0.243
...
26  dem65 =~    y1 0.108  0.169   0.130    0.131
27  dem65 =~    y2 0.617 -0.410  -0.316   -0.319
28  dem65 =~    y3 0.423 -0.357  -0.275   -0.278
29  dem65 =~    y4 0.744  0.456   0.351    0.354
30  dem65 =~    y5    NA     NA      NA       NA
...
73     y2 ~~    y3 1.333 -0.061  -0.061   -0.062
74     y2 ~~    y4 5.668  0.113   0.113    0.115
75     y2 ~~    y5 0.560  0.035   0.035    0.035
...
95     y6 ~~    y7 0.495 -0.032  -0.032   -0.032
96     y6 ~~    y8 8.813  0.132   0.132    0.133
97     y7 ~~    y7 0.000  0.000   0.000    0.000
...
109 dem65  ~ ind60 0.000  0.000   0.000    0.000
110 ind60  ~ dem60    NA     NA      NA       NA
111 ind60  ~ dem65    NA     NA      NA       NA
```

Even more correlated residuals

```
mod.sem3<- '
# latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + y6 + y7 + y8
     # regressions
    dem60 ~ ind60
    dem65 ~ ind60 + dem60
    y1 ~~ y5
    y2 ~~ y4 + y6
    y3 ~~ y7
    y4 ~~ y8
    y6 ~~ y8'
fit3 <- sem(mod.sem3, data=PoliticalDemocracy,std.ov=TRUE)
summary(fit3, fit.measures=TRUE)
lavaan.diagram(fit3,cut=0,lr=FALSE)
```

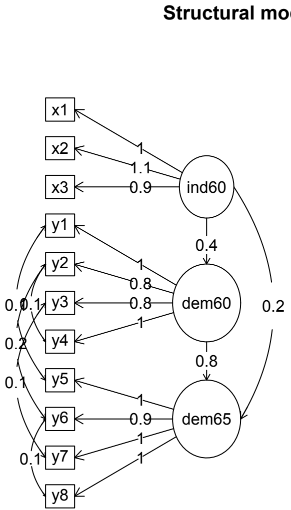
```
Lavaan (0.4-6) converged normally after 46 iterations
  Number of observations                              75
  Estimator                                           ML
  Minimum Function Chi-square                     38.125
  Degrees of freedom                                  35
  P-value                                          0.329
Chi-square test baseline model:
  Minimum Function Chi-square                    730.654
  Degrees of freedom                                  55
  P-value                                          0.000
Root Mean Square Error of Approximation:
  RMSEA                                            0.035
  90 Percent Confidence Interval          0.000    0.092
  P-value RMSEA <= 0.05                            0.611
Standardized Root Mean Square Residual:
  SRMR                                             0.044
```

The very modified model



The very modified model – another way

lavaan.diagram(fit3,cut=0)



A model with constraints at time 2 as well as correlated residuals

```
## The industrialization and Political Democracy Example
## Bollen (1989), page 332
model <- '
  # latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + equal("dem60=~y2")*y6
                 + equal("dem60=~y3")*y7
                 + equal("dem60=~y4")*y8
  # regressions
    dem60 ~ ind60
    dem65 ~ ind60 + dem60

  # residual correlations
    y1 ~~ y5
    y2 ~~ y4 + y6
    y3 ~~ y7
    y4 ~~ y8
    y6 ~~ y8
'
fit <- sem(model, data=PoliticalDemocracy)
summary(fit, fit.measures=TRUE)

  Minimum Function Chi-square                   40.179
  Degrees of freedom                                38
  P-value                                        0.374
Chi-square test baseline model:
  Minimum Function Chi-square                  730.654
  Degrees of freedom                                55
  P-value                                        0.000
```

```
Full model versus baseline model:
  Comparative Fit Index (CFI)                    0.997
  Tucker-Lewis Index (TLI)                       0.995
Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)              -1548.818
  Loglikelihood unrestricted model (H1)      -1528.728

  Number of free parameters                         28
  Akaike (AIC)                                3153.636
  Bayesian (BIC)                              3218.526
  Sample-size adjusted Bayesian (BIC)         3130.277

Root Mean Square Error of Approximation:

  RMSEA                                          0.028
  90 Percent Confidence Interval          0.000  0.087
  P-value RMSEA <= 0.05                          0.665

Standardized Root Mean Square Residual:

  SRMR                                           0.056
```

Constrained model with parameter values

```
 Information                                 Expected Covariances:
  Standard Errors                             Standard  y1 ~~
                                                          y5              0.583    0.356    1.637    0.102
                    Estimate  Std.err  Z-value  P(>|z|)  y2 ~~
Latent variables:                                         y4              1.440    0.689    2.092    0.036
  ind60 =~                                                y6              2.183    0.737    2.960    0.003
    x1              1.000                                y3 ~~
    x2              2.180    0.138   15.751    0.000      y7              0.712    0.611    1.165    0.244
    x3              1.818    0.152   11.971    0.000   y4 ~~
  dem60 =~                                                y8              0.363    0.444    0.817    0.414
    y1              1.000                                y6 ~~
    y2              1.191    0.139    8.551    0.000      y8              1.372    0.577    2.378    0.017
    y3              1.175    0.120    9.755    0.000
    y4              1.251    0.117   10.712    0.000Variances:
  dem65 =~                                                x1              0.081    0.019    4.182    0.000
    y5              1.000                                 x2              0.120    0.070    1.729    0.084
    y6              1.191                                 x3              0.467    0.090    5.177    0.000
    y7              1.175                                 y1              1.855    0.433    4.279    0.000
    y8              1.251                                 y2              7.581    1.366    5.549    0.000
                                                          y3              4.956    0.956    5.182    0.000
Regressions:                                              y4              3.224    0.723    4.458    0.000
  dem60 ~                                                 y5              2.313    0.479    4.831    0.000
    ind60           1.471    0.392    3.750    0.000      y6              4.968    0.921    5.393    0.000
  dem65 ~                                                 y7              3.560    0.710    5.018    0.000
    ind60           0.600    0.226    2.661    0.008      y8              3.308    0.704    4.701    0.000
    dem60           0.865    0.075   11.554    0.000      ind60           0.449    0.087    5.175    0.000
                                                          dem60           3.875    0.866    4.477    0.000
                                                          dem65           0.164    0.227    0.725    0.469
```

Standardized solution

```
> standardizedSolution(fit)


     lhs op   rhs   est est.std est.std.all
1  ind60 =~    x1 1.000   0.670       0.920
2  ind60 =~    x2 2.180   1.460       0.973
3  ind60 =~    x3 1.818   1.218       0.872
4  dem60 =~    y1 1.000   2.201       0.850
5  dem60 =~    y2 1.191   2.621       0.690
6  dem60 =~    y3 1.175   2.586       0.758
7  dem60 =~    y4 1.251   2.754       0.838
8  dem65 =~    y5 1.000   2.154       0.817
9  dem65 =~    y6 1.191   2.565       0.755
10 dem65 =~    y7 1.175   2.530       0.802
11 dem65 =~    y8 1.251   2.694       0.829
12 dem60  ~ ind60 1.471   0.448       0.448
13 dem65  ~ ind60 0.600   0.187       0.187
14 dem65  ~ dem60 0.865   0.884       0.884
15    y1 ~~    y5 0.583   0.583       0.085
16    y2 ~~    y4 1.440   1.440       0.115
17    y2 ~~    y6 2.183   2.183       0.169
18    y3 ~~    y7 0.712   0.712       0.066
19    y4 ~~    y8 0.363   0.363       0.034
20    y6 ~~    y8 1.372   1.372       0.124
21    x1 ~~    x1 0.081   0.081       0.154
22    x2 ~~    x2 0.120   0.120       0.053
23    x3 ~~    x3 0.467   0.467       0.239
24    y1 ~~    y1 1.855   1.855       0.277
25    y2 ~~    y2 7.581   7.581       0.525
26    y3 ~~    y3 4.956   4.956       0.426
27    y4 ~~    y4 3.224   3.224       0.298
28    y5 ~~    y5 2.313   2.313       0.333
29    y6 ~~    y6 4.968   4.968       0.430
30    y7 ~~    y7 3.560   3.560       0.357
31    y8 ~~    y8 3.308   3.308       0.313
32 ind60 ~~ ind60 0.449   1.000       1.000
33 dem60 ~~ dem60 3.875   0.800       0.800
34 dem65 ~~ dem65 0.164   0.035       0.035
```

Lisrel example

```
model <- '
  # latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + equal("dem60=~y2")*y6
                 + equal("dem60=~y3")*y7
                 + equal("dem60=~y4")*y8

  # regressions
    dem60 ~ ind60
    dem65 ~ ind60 + dem60
  # residual correlations
    y1 ~~ y5
    y2 ~~ y6
    y3 ~~ y7
    y4 ~~ y8'
fit <- sem(model, data=PoliticalDemocracy,
   likelihood='wishart')
summary(fit, fit.measures=TRUE)

Lavaan (0.4-6) converged normally after 88 iterations
  Number of observations                            75
  Estimator                                         ML
  Minimum Function Chi-square                   53.156
  Degrees of freedom                                40
  P-value                                        0.080
Chi-square test baseline model:
  Minimum Function Chi-square                  720.912
  Degrees of freedom                                55
  P-value                                        0.000
Full model versus baseline model:
  Comparative Fit Index (CFI)                    0.980
```

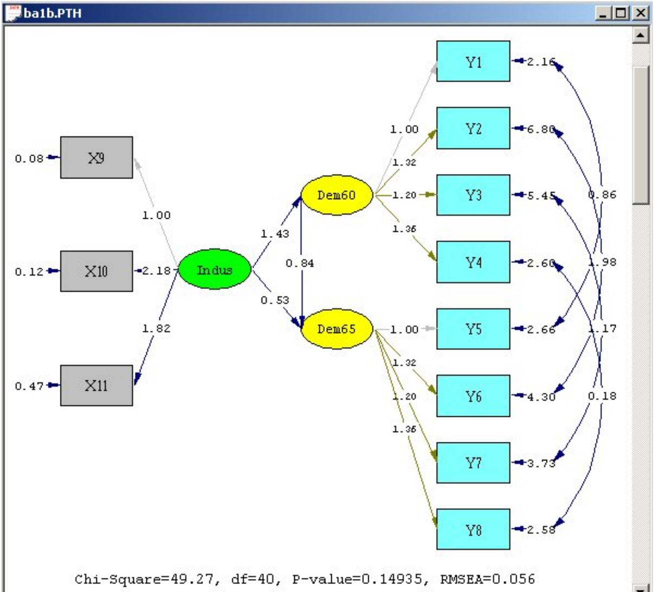
```
Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)              -1561.202
  Loglikelihood unrestricted model (H1)      -1534.265

  Number of free parameters                         26
  Akaike (AIC)                                3174.405
  Bayesian (BIC)                              3234.659
  Sample-size adjusted Bayesian (BIC)         3152.714

Root Mean Square Error of Approximation:

  RMSEA                                          0.066
  90 Percent Confidence Interval          0.000  0.110
  P-value RMSEA <= 0.05                          0.277

Standardized Root Mean Square Residual:

  SRMR                                           0.061
```

With parameters

```
                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  ind60 =~
    x1                1.000
    x2                2.180    0.140   15.620    0.000
    x3                1.819    0.153   11.888    0.000
  dem60 =~
    y1                1.000
    y2                1.318    0.149    8.869    0.000
    y3                1.203    0.133    9.034    0.000
    y4                1.363    0.128   10.686    0.000
  dem65 =~
    y5                1.000
    y6                1.318
    y7                1.203
    y8                1.363

Regressions:
  dem60 ~
    ind60             1.427    0.381    3.747    0.000
  dem65 ~
    ind60             0.525    0.215    2.445    0.014
    dem60             0.845    0.071   11.850    0.000
```

```
ovariances:
  y1 ~~
    y5                0.862    0.370    2.327    0.020
  y2 ~~
    y6                1.978    0.786    2.516    0.012
  y3 ~~
    y7                1.171    0.638    1.836    0.066
  y4 ~~
    y8                0.179    0.478    0.374    0.708

Variances:
    x1                0.083    0.020    4.147    0.000
    x2                0.122    0.071    1.702    0.089
    x3                0.473    0.092    5.138    0.000
    y1                2.157    0.454    4.747    0.000
    y2                6.799    1.265    5.374    0.000
    y3                5.452    1.021    5.340    0.000
    y4                2.601    0.669    3.890    0.000
    y5                2.664    0.515    5.172    0.000
    y6                4.301    0.841    5.117    0.000
    y7                3.729    0.724    5.151    0.000
    y8                2.576    0.620    4.158    0.000
    ind60             0.455    0.088    5.138    0.000
    dem60             3.659    0.845    4.329    0.000
    dem65             0.367    0.199    1.851    0.064
```

LISREL output



Reverse causation: Democracy leads to industrialization

```
model <- '
  # latent variable definitions
     ind60 =~ x1 + x2 + x3
     dem60 =~ y1 + y2 + y3 + y4
     dem65 =~ y5 + equal("dem60=~y2")*y6
                 + equal("dem60=~y3")*y7
                 + equal("dem60=~y4")*y8

  # regressions
   ind60 ~ dem60
    dem65 ~ ind60 + dem60
  # residual correlations
    y1 ~~ y5
    y2 ~~ y6
    y3 ~~ y7
    y4 ~~ y8'
fit <- sem(model, data=PoliticalDemocracy,
   likelihood='wishart')
summary(fit, fit.measures=TRUE)
```

```
Lavaan (0.4-6) converged normally after 90 iterations
  Number of observations                            75
  Estimator                                         ML
  Minimum Function Chi-square                   53.156
  Degrees of freedom                                40
  P-value                                        0.080
Chi-square test baseline model:
  Minimum Function Chi-square                  720.912
  Degrees of freedom                                55
  P-value                                        0.000
Full model versus baseline model:
```