Adventures in modeling

Testing differences between models

The power of simulation

I. Simulation of structures allows us to know truth.
II. If we can recover truth, then perhaps our models work
III. If we can not recover truth, then our models do not work
IV. Simulated structure
V. Simulated error

Simulated structure

I. Simulated structure

   A. $R = F F' + U^2$

   B. $R = F \phi F' + U^2$

II. Simulated error

   A. $E = N(0,1)$ of dimensionality of $n * r$

   B. $R = EF \phi F'E' + U^2$

Simulating structure in R

A. sim.structural(fx,fy,f, Phi,N)
B. sim.hierarchical(gload,fload,N)
C. sim.congeneric(loads,N)
D. sim.item(nvar,nsub, ...)
E. sim.circ(nvar,nsub, ...) (special case of sim.item

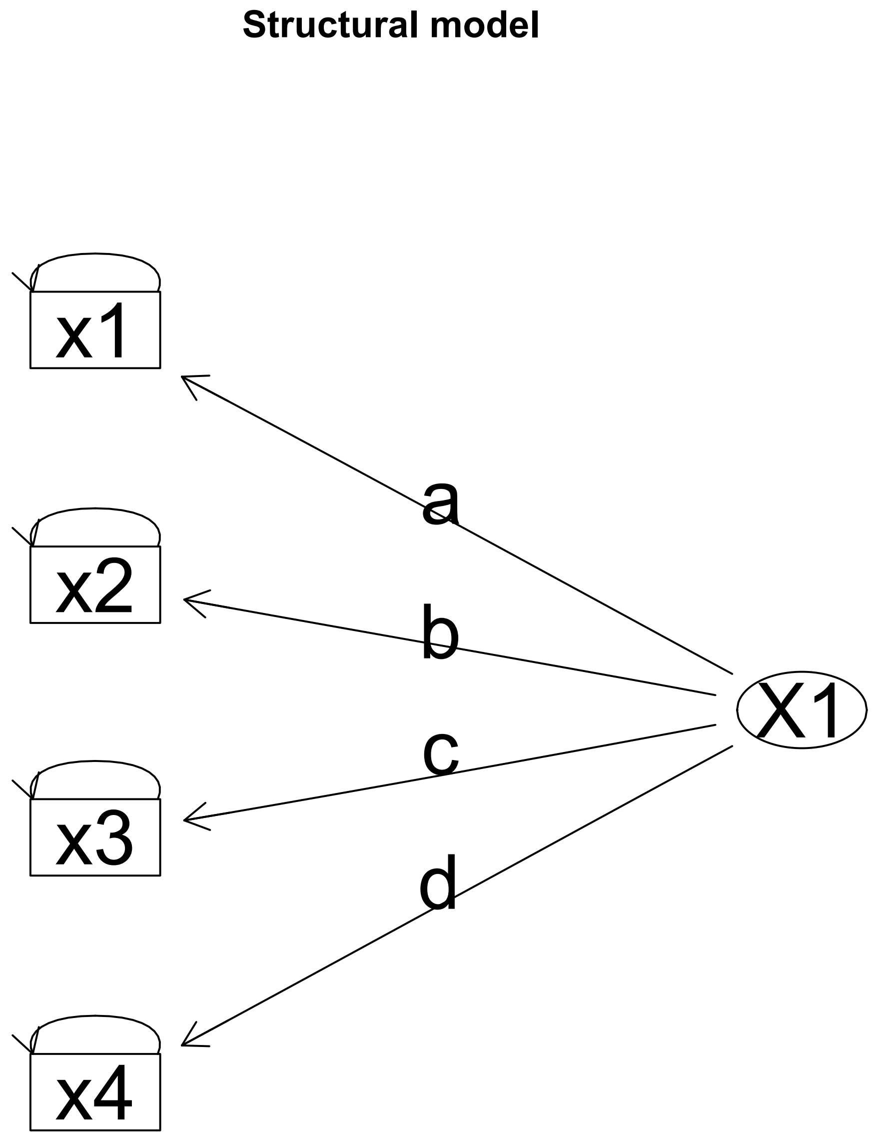
Congeneric model

```
>R<- sim.congeneric(loads=c(.9,.8,.7,.6), N=200,short=FALSE)
>R

 $model (Population correlation matrix)
     V1   V2   V3   V4
V1 1.00 0.72 0.63 0.54
V2 0.72 1.00 0.56 0.48
V3 0.63 0.56 1.00 0.42
V4 0.54 0.48 0.42 1.00

$r  (Sample correlation matrix  for sample size =  200 )
     V1   V2   V3   V4
V1 1.00 0.77 0.58 0.63
V2 0.77 1.00 0.57 0.50
V3 0.58 0.57 1.00 0.41
V4 0.63 0.50 0.41 1.00
```

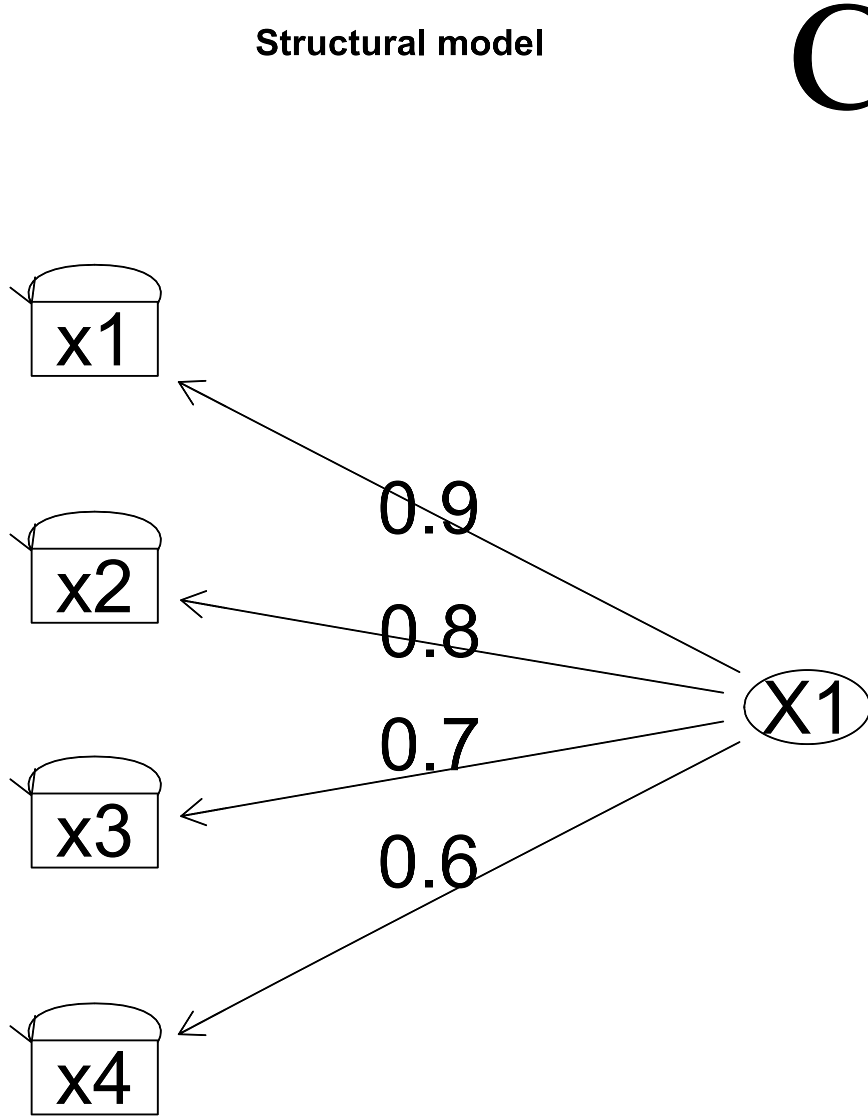
Exploratory factor analysis

```
> c1 <- factanal(covmat=R$model,factors=1,n.obs=100)
> c1
Call:
factanal(factors = 1, covmat = R$model, n.obs = 100)
Uniquenesses:
  V1   V2   V3   V4
0.19 0.36 0.51 0.64
Loadings:
   Factor1
V1 0.9
V2 0.8
V3 0.7
V4 0.6
               Factor1
SS loadings      2.300
Proportion Var   0.575
Test of the hypothesis that 1 factor is sufficient.
The chi square statistic is 0 on 2 degrees of freedom.
The p-value is 1
```

Congeneric model



Congeneric model



sem commands

```
> con.mod <- structure.graph(c1)
> con.mod
      [,1]                   [,2]  [,3]
 [1,] "Factor1->V1"          "V1"  NA
 [2,] "Factor1->V2"          "V2"  NA
 [3,] "Factor1->V3"          "V3"  NA
 [4,] "Factor1->V4"          "V4"  NA
 [5,] "V1<->V1"              "v1e" NA
 [6,] "V2<->V2"              "v2e" NA
 [7,] "V3<->V3"              "v3e" NA
 [8,] "V4<->V4"              "v4e" NA
 [9,] "Factor1<->Factor1" NA    "1"
```

congeneric for sem (population correlations)

```
> sem.con <- sem(con.mod,R$model,100)
> summary(sem.con)

 Model Chisquare =  5.201e-11   Df =  2 Pr(>Chisq) = 1
 Chisquare (null model) =  163.79   Df =  6
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  1
 RMSEA index =  0   90% CI: (NA, NA)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1.0380
 Bentler CFI =  1
 SRMR =  2.1696e-07
 BIC =  -9.2103

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-3.78e-06 -3.59e-07  5.80e-07  2.84e-07  1.17e-06  2.37e-06
```

sem paths

```
Parameter Estimates
    Estimate Std Error z value Pr(>|z|)
V1  0.90     0.085184  10.5654 0.0000e+00 V1 <--- Factor1
V2  0.80     0.089029   8.9858 0.0000e+00 V2 <--- Factor1
V3  0.70     0.092728   7.5489 4.3965e-14 V3 <--- Factor1
V4  0.60     0.096544   6.2148 5.1397e-10 V4 <--- Factor1
v1e 0.19     0.069000   2.7536 5.8941e-03 V1 <--> V1
v2e 0.36     0.072952   4.9348 8.0253e-07 V2 <--> V2
v3e 0.51     0.084619   6.0270 1.6698e-09 V3 <--> V3
v4e 0.64     0.098855   6.4742 9.5338e-11 V4 <--> V4

 Iterations =  12
```

sem on sample data

```
> sem.con.samp <- sem(con.mod,R$r,200)
> summary(sem.con.samp)

 Model Chisquare =  5.573   Df =  2 Pr(>Chisq) = 0.061637
 Chisquare (null model) =  370.55   Df =  6
 Goodness-of-fit index =  0.98614
 Adjusted goodness-of-fit index =  0.93071
 RMSEA index =  0.094749   90% CI: (NA, 0.19237)
 Bentler-Bonnett NFI =  0.98496
 Tucker-Lewis NNFI =  0.9706
 Bentler CFI =  0.9902
 SRMR =  0.021766
 BIC =  -5.0237

 Normalized Residuals
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-0.51700 -0.13200 -0.00318 -0.00207  0.04360  0.62700
```

Sample path data

```
Parameter Estimates
    Estimate Std Error z value Pr(>|z|)
V1  0.93752  0.082254  11.3979 0.0000e+00 V1 <--- Factor1
V2  0.81842  0.087310   9.3737 0.0000e+00 V2 <--- Factor1
V3  0.63168  0.095472   6.6164 3.6812e-11 V3 <--- Factor1
V4  0.65745  0.092431   7.1129 1.1360e-12 V4 <--- Factor1
v1e 0.12106  0.064563   1.8750 6.0793e-02 V1 <--> V1
v2e 0.33018  0.067984   4.8568 1.1929e-06 V2 <--> V2
v3e 0.60099  0.094555   6.3559 2.0717e-10 V3 <--> V3
v4e 0.56775  0.087146   6.5150 7.2712e-11 V4 <--> V4

 Iterations =  13
```

```
> sem.con.samp <- sem(con.mod,R,400)
> R <- sim.congeneric(loads=c(.9,.8,.7,.6), N= 400)
> round(R,2)
     V1   V2   V3   V4
V1 1.00 0.72 0.61 0.51
V2 0.72 1.00 0.54 0.47
V3 0.61 0.54 1.00 0.44
V4 0.51 0.47 0.44 1.00
> sem.con.samp <- sem(con.mod,R,400)
> summary(sem.con.samp)
```

Sample with N = 400

```
 Model Chisquare =  2.5386   Df =  2 Pr(>Chisq) = 0.28103
 Chisquare (null model) =  635.03   Df =  6
 Goodness-of-fit index =  0.99683
 Adjusted goodness-of-fit index =  0.98413
 RMSEA index =  0.025979   90% CI: (NA, 0.10633)
 Bentler-Bonnett NFI =  0.996
 Tucker-Lewis NNFI =  0.99743
 Bentler CFI =  0.99914
 SRMR =  0.013338
 BIC =  -9.4444
```

sem: sample 400

```
 Parameter Estimates
    Estimate Std Error z value Pr(>|z|)
V1  0.89009  0.042846  20.7740 0.0000e+00 V1 <--- Factor1
V2  0.80363  0.044342  18.1233 0.0000e+00 V2 <--- Factor1
V3  0.68581  0.046682  14.6910 0.0000e+00 V3 <--- Factor1
V4  0.58946  0.048623  12.1231 0.0000e+00 V4 <--- Factor1
v1e 0.20774  0.035165   5.9075 3.4734e-09 V1 <--> V1
v2e 0.35418  0.036381   9.7354 0.0000e+00 V2 <--> V2
v3e 0.52967  0.043589  12.1515 0.0000e+00 V3 <--> V3
v4e 0.65254  0.050426  12.9403 0.0000e+00 V4 <--> V4

 Iterations =  13
```

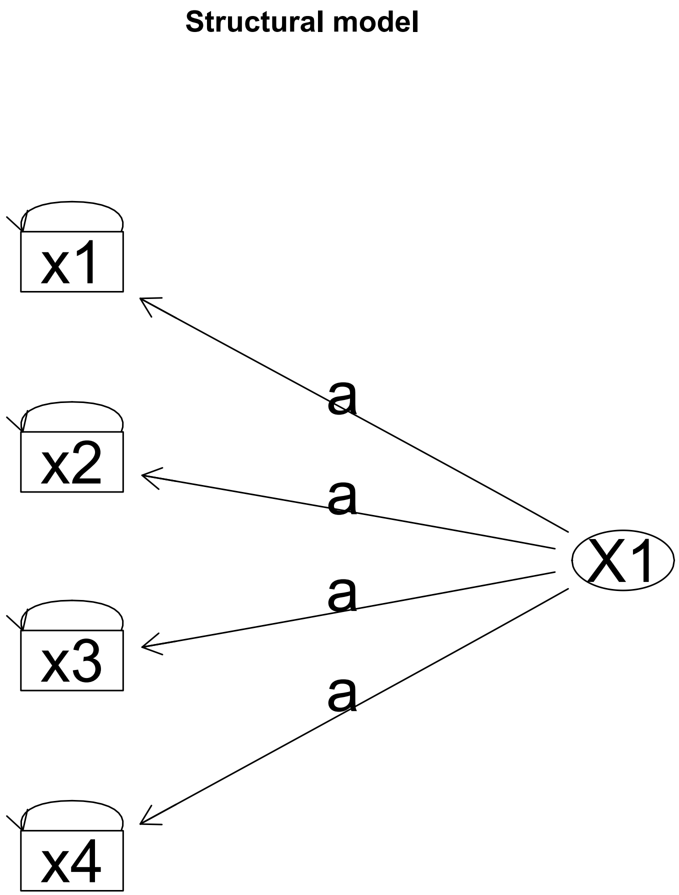
Testing alternative models

I. Congeneric model allows paths to be different

II. Tau equivalent requires paths to be the same.

III. Test a tau equivalent model

Tau equivalence



Changing models

I. edit a prior model using edit function

II. tau.mod <- edit(con.mod)

Tau equivalent model

```
> tau.mod <- edit(con.mod)
> tau.mod
      col1                    col2  col3
 [1,] "Factor1->V1"           "a"   NA
 [2,] "Factor1->V2"           "a"   NA
 [3,] "Factor1->V3"           "a"   NA
 [4,] "Factor1->V4"           "a"   NA
 [5,] "V1<->V1"               "v1e" NA
 [6,] "V2<->V2"               "v2e" NA
 [7,] "V3<->V3"               "v3e" NA
 [8,] "V4<->V4"               "v4e" NA
 [9,] "Factor1<->Factor1" NA    "1"
```

Tau equivalent fit

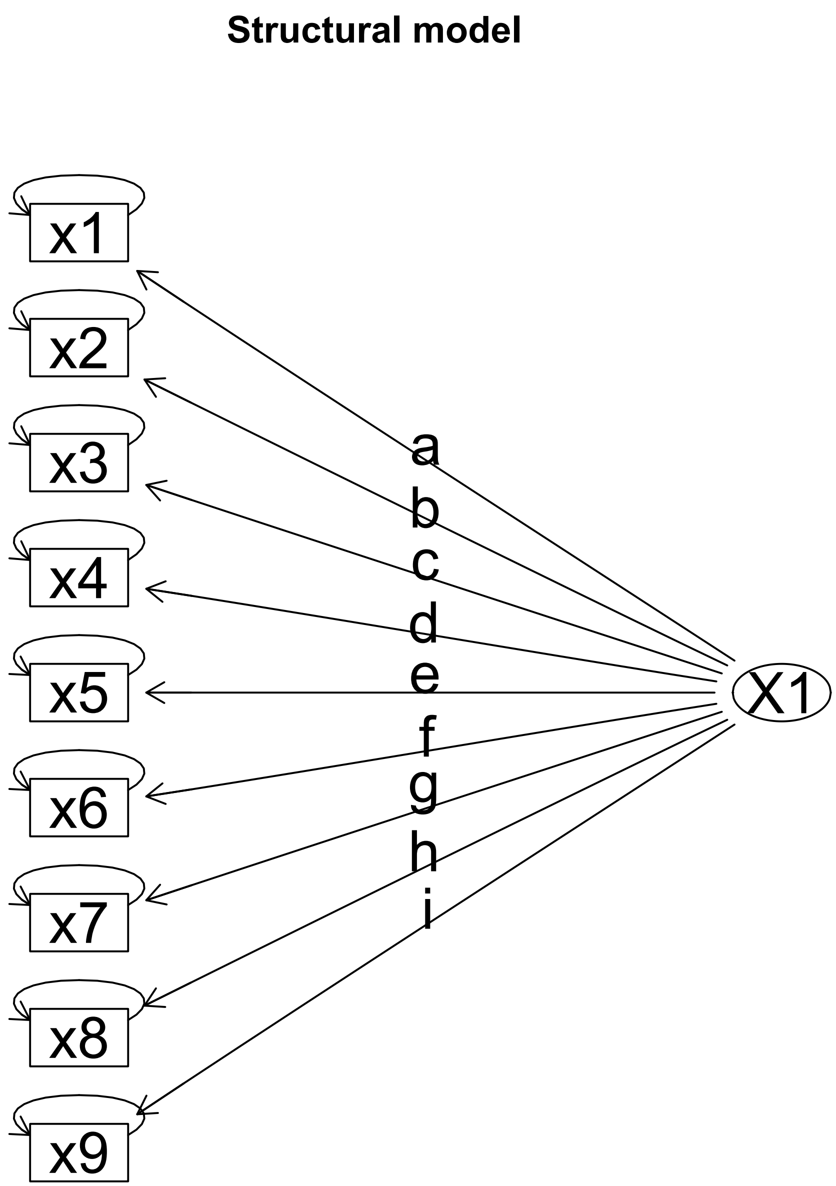
```
>  sem.tau <- sem(tau.mod,R$model,100)
> summary(sem.tau)

 Model Chisquare =  8.8453   Df =  5 Pr(>Chisq) = 0.11540
 Chisquare (null model) =  163.79   Df =  6
 Goodness-of-fit index =  0.95793
 Adjusted goodness-of-fit index =  0.91585
 RMSEA index =  0.088138   90% CI: (NA, 0.18136)
 Bentler-Bonnett NFI =  0.946
 Tucker-Lewis NNFI =  0.97076
 Bentler CFI =  0.97563
 SRMR =  0.12051
 BIC =  -14.181
```

Parameters for N = 100

```
 Parameter Estimates
    Estimate Std Error z value Pr(>|z|)
a   0.77793  0.065658  11.8481 0.0000e+00 V1 <--- Factor1
v1e 0.27519  0.057904   4.7525 2.0089e-06 V1 <--> V1
v2e 0.35502  0.067115   5.2898 1.2248e-07 V2 <--> V2
v3e 0.48247  0.083979   5.7452 9.1831e-09 V3 <--> V3
v4e 0.62200  0.103498   6.0098 1.8580e-09 V4 <--> V4
 Iterations =  11

> std.coef(sem.tau)
    Std. Estimate
1 a 0.82910       V1 <--- Factor1
2 a 0.79389       V2 <--- Factor1
3 a 0.74593       V3 <--- Factor1
4 a 0.70224       V4 <--- Factor1
```

```
> f <- c(.9,.8,.7,.8,.7,.6,.7,.6,.5)
> R <- sim.congeneric(loads=f)
> f1 <- factor.pa(R)
> f1
  V PA1
1 1 0.9
2 2 0.8
3 3 0.7
4 4 0.8
5 5 0.7
6 6 0.6
7 7 0.7
8 8 0.6
9 9 0.5
               PA1
SS loadings    4.53
Proportion Var 0.50
Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the model is 27  and the fit was
0
```

Extend model to 9 items



9 variables congeneric model

Automatic model not quite right

```
> mod.con9 <- structure.graph(f1)
> mod.con9
      [,1]          [,2]  [,3]
 [1,] "PA1->V1"     "V1"  NA
 [2,] "PA1->V2"     "V2"  NA
 [3,] "PA1->V3"     "V3"  NA
 [4,] "PA1->V4"     "V4"  NA
 [5,] "PA1->V5"     "V5"  NA
 [6,] "PA1->V6"     "V6"  NA
 [7,] "PA1->V7"     "V7"  NA
 [8,] "PA1->V8"     "V8"  NA
 [9,] "PA1->V9"     "V9"  NA
[10,] "V1<->V1"     "v1e" NA
...
[[17,] "V8<->V8"     "v8e" NA
[18,] "V9<->V9"     "v9e" NA
[19,] "PA1<->PA1" NA    "1"
[20,] NA            NA    NA
> mod.con9 <- mod.con9[1:19,]
```

bug produces extra line
remove it

sem congeneric 9

```
> sem.con9 <- sem(mod.con9,R,100)
> summary(sem.con9)

 Model Chisquare =  4.2611e-10   Df =  27 Pr(>Chisq) = 1
 Chisquare (null model) =  428.39   Df =  36
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  1
 RMSEA index =  0   90% CI: (NA, NA)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1.0917
 Bentler CFI =  1
 SRMR =  5.5811e-07
 BIC =  -124.34

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-2.87e-06 -2.25e-06 -1.02e-06  1.03e-06  7.93e-07  1.53e-05
```

path values

Parameter Estimates

```
     Estimate Std Error z value Pr(>|z|)
V1  0.90   0.079610  11.3051 0.0000e+00 V1 <--- PA1
V2  0.80   0.085248   9.3844 0.0000e+00 V2 <--- PA1
V3  0.70   0.090098   7.7694 7.9936e-15 V3 <--- PA1
V4  0.80   0.085248   9.3844 0.0000e+00 V4 <--- PA1
V5  0.70   0.090098   7.7693 7.9936e-15 V5 <--- PA1
V6  0.60   0.094137   6.3737 1.8450e-10 V6 <--- PA1
V7  0.70   0.090098   7.7693 7.9936e-15 V7 <--- PA1
V8  0.60   0.094137   6.3737 1.8449e-10 V8 <--- PA1
V9  0.50   0.097434   5.1317 2.8716e-07 V9 <--- PA1
v1e 0.19   0.042114   4.5115 6.4358e-06 V1 <--> V1
v2e 0.36   0.060241   5.9760 2.2870e-09 V2 <--> V2
v3e 0.51   0.078831   6.4695 9.8304e-11 V3 <--> V3
v4e 0.36   0.060241   5.9760 2.2870e-09 V4 <--> V4
v5e 0.51   0.078831   6.4695 9.8304e-11 V5 <--> V5
v6e 0.64   0.095443   6.7056 2.0062e-11 V6 <--> V6
v7e 0.51   0.078831   6.4695 9.8304e-11 V7 <--> V7
v8e 0.64   0.095443   6.7056 2.0062e-11 V8 <--> V8
v9e 0.75   0.109647   6.8401 7.9132e-12 V9 <--> V9
```

A more complicated model

I. 9 variables (e.g., 9 measures of Neuroticism)

II. 3 group factors (e.g., anxiety, depression, aggression)

III. 1 higher order factor (e.g., Neuroticism)

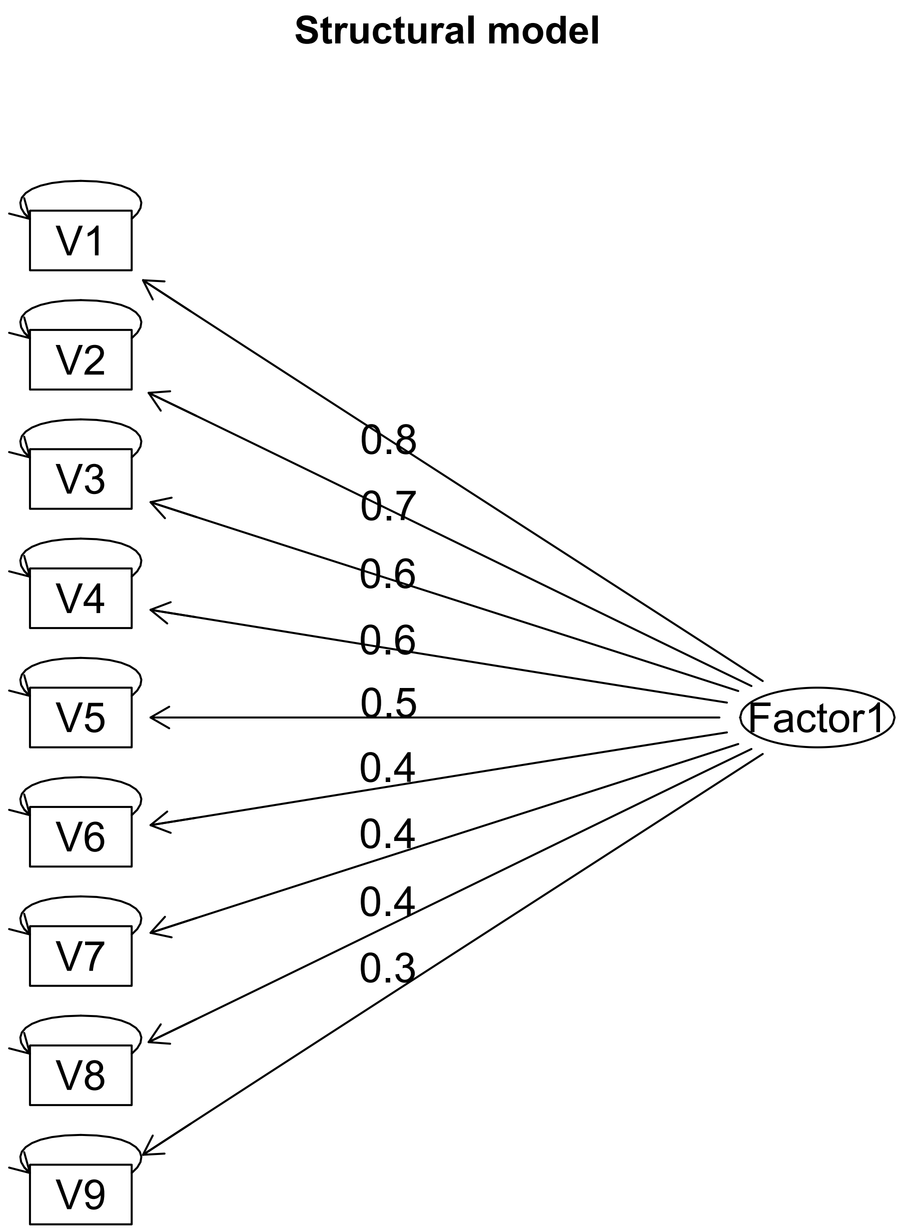
IV. Simulate using sim.hierarchical() (could use sim.structural()

Nine variables

```
> nine <- sim.hierarchical()
> round(nine,2)
     V1   V2   V3   V4   V5   V6   V7   V8   V9
V1 1.00 0.56 0.48 0.40 0.35 0.29 0.30 0.25 0.20
V2 0.56 1.00 0.42 0.35 0.30 0.25 0.26 0.22 0.18
V3 0.48 0.42 1.00 0.30 0.26 0.22 0.23 0.19 0.15
V4 0.40 0.35 0.30 1.00 0.42 0.35 0.24 0.20 0.16
V5 0.35 0.30 0.26 0.42 1.00 0.30 0.20 0.17 0.13
V6 0.29 0.25 0.22 0.35 0.30 1.00 0.17 0.14 0.11
V7 0.30 0.26 0.23 0.24 0.20 0.17 1.00 0.30 0.24
V8 0.25 0.22 0.19 0.20 0.17 0.14 0.30 1.00 0.20
V9 0.20 0.18 0.15 0.16 0.13 0.11 0.24 0.20 1.00
>
```

ML factor analysis

```
> f1 <- factanal(covmat=nine,factors=1,n.obs=200)
Loadings:
   Factor1
V1 0.757
V2 0.678
V3 0.585
V4 0.574
V5 0.503
V6 0.425
V7 0.423
V8 0.359
V9 0.291
               Factor1
SS loadings      2.531
Proportion Var   0.281
Test of the hypothesis that 1 factor is sufficient.
The chi square statistic is 31.51 on 27 degrees of
freedom.
The p-value is 0.251
```



factor model

SEM of nine variables

```
> mod1 <- mod1[1:19,]
> sem.1 <- sem(mod1,nine,200)
Error in sem.default(ram = ram, S = S, N = N, param.names = pars, 
var.names = vars,  : 
  S must be a square triangular or symmetric matrix
     (test in sem is whether R == t(R) and some elements are not 
exactly equal)

> nine <- round(nine,4)
> sem.1 <- sem(mod1,nine,200)
```

Fit statistics show a good fit

```
>  summary(sem.1)

 Model Chisquare =  32.242   Df =  27 Pr(>Chisq) =
0.22323
 Chisquare (null model) =  340.21   Df =  36
 Goodness-of-fit index =  0.96133
 Adjusted goodness-of-fit index =  0.93555
 RMSEA index =  0.031236   90% CI: (NA, 0.066409)
 Bentler-Bonnett NFI =  0.90523
 Tucker-Lewis NNFI =  0.97702
 Bentler CFI =  0.98277
 SRMR =  0.047258
 BIC =  -110.81
```

Parameters

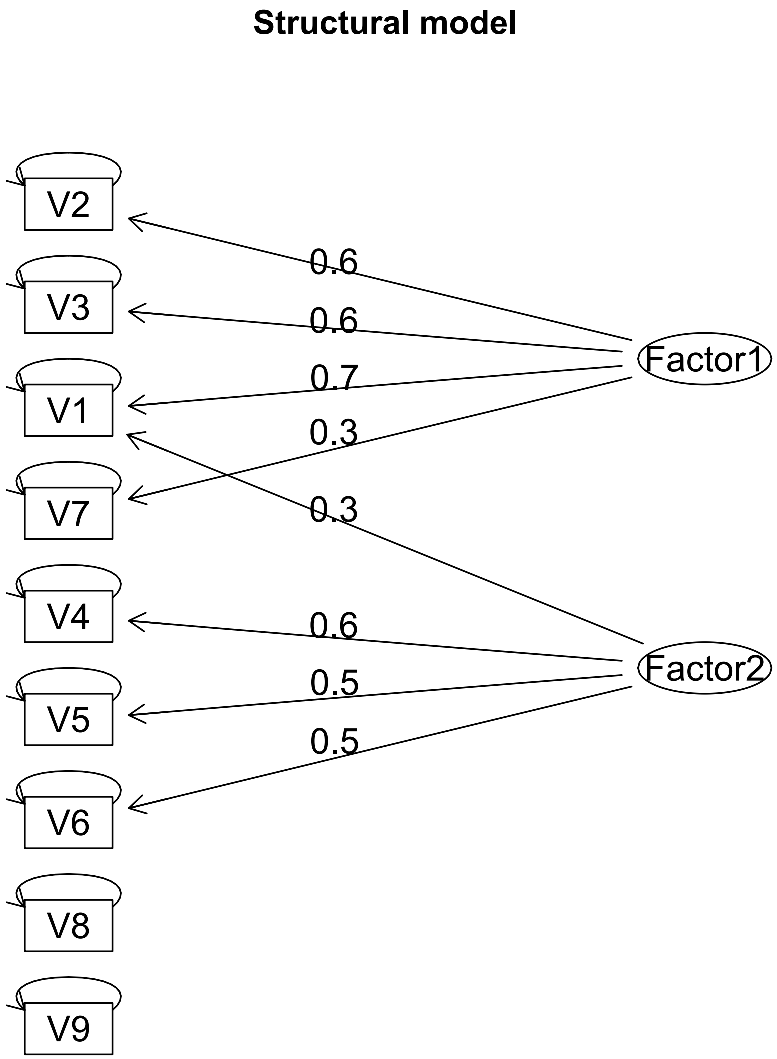
```
 Parameter Estimates
    Estimate Std Error z value Pr(>|z|)
V1   0.75705  0.067644  11.1917 0.0000e+00 V1 <--- Factor1
V2   0.67797  0.069404   9.7686 0.0000e+00 V2 <--- Factor1
V3   0.58514  0.071706   8.1603 4.4409e-16 V3 <--- Factor1
V4   0.57428  0.072784   7.8902 3.1086e-15 V4 <--- Factor1
V5   0.50305  0.074297   6.7707 1.2813e-11 V5 <--- Factor1
V6   0.42489  0.075611   5.6194 1.9163e-08 V6 <--- Factor1
V7   0.42303  0.075416   5.6092 2.0326e-08 V7 <--- Factor1
V8   0.35911  0.076439   4.6980 2.6269e-06 V8 <--- Factor1
V9   0.29107  0.077285   3.7662 1.6574e-04 V9 <--- Factor1
v1e 0.42687  0.064051   6.6645 2.6553e-11 V1 <--> V1
v2e 0.54035  0.068381   7.9021 2.6645e-15 V2 <--> V2
v3e 0.65761  0.075405   8.7210 0.0000e+00 V3 <--> V3
v4e 0.67020  0.077258   8.6748 0.0000e+00 V4 <--> V4
v5e 0.74694  0.082185   9.0885 0.0000e+00 V5 <--> V5
v6e 0.81947  0.087059   9.4128 0.0000e+00 V6 <--> V6
v7e 0.82105  0.087030   9.4341 0.0000e+00 V7 <--> V7
v8e 0.87104  0.090651   9.6087 0.0000e+00 V8 <--> V8
v9e 0.91528  0.093896   9.7478 0.0000e+00 V9 <--> V9
```

Residuals suggest some problems

```
> round(residuals(sem.1),2)
      V1    V2    V3    V4    V5    V6    V7    V8    V9
V1  0.00  0.05  0.04 -0.03 -0.04 -0.03 -0.02 -0.02 -0.02
V2  0.05  0.00  0.02 -0.04 -0.04 -0.04 -0.02 -0.02 -0.02
V3  0.04  0.02  0.00 -0.03 -0.04 -0.03 -0.02 -0.02 -0.02
V4 -0.03 -0.04 -0.03  0.00  0.13  0.11 -0.01 -0.01 -0.01
V5 -0.04 -0.04 -0.04  0.13  0.00  0.09 -0.01 -0.01 -0.01
V6 -0.03 -0.04 -0.03  0.11  0.09  0.00 -0.01 -0.01 -0.01
V7 -0.02 -0.02 -0.02 -0.01 -0.01 -0.01  0.00  0.15  0.12
V8 -0.02 -0.02 -0.02 -0.01 -0.01 -0.01  0.15  0.00  0.10
V9 -0.02 -0.02 -0.02 -0.01 -0.01 -0.01  0.12  0.10  0.00
```

Try two factors (increase sample size)

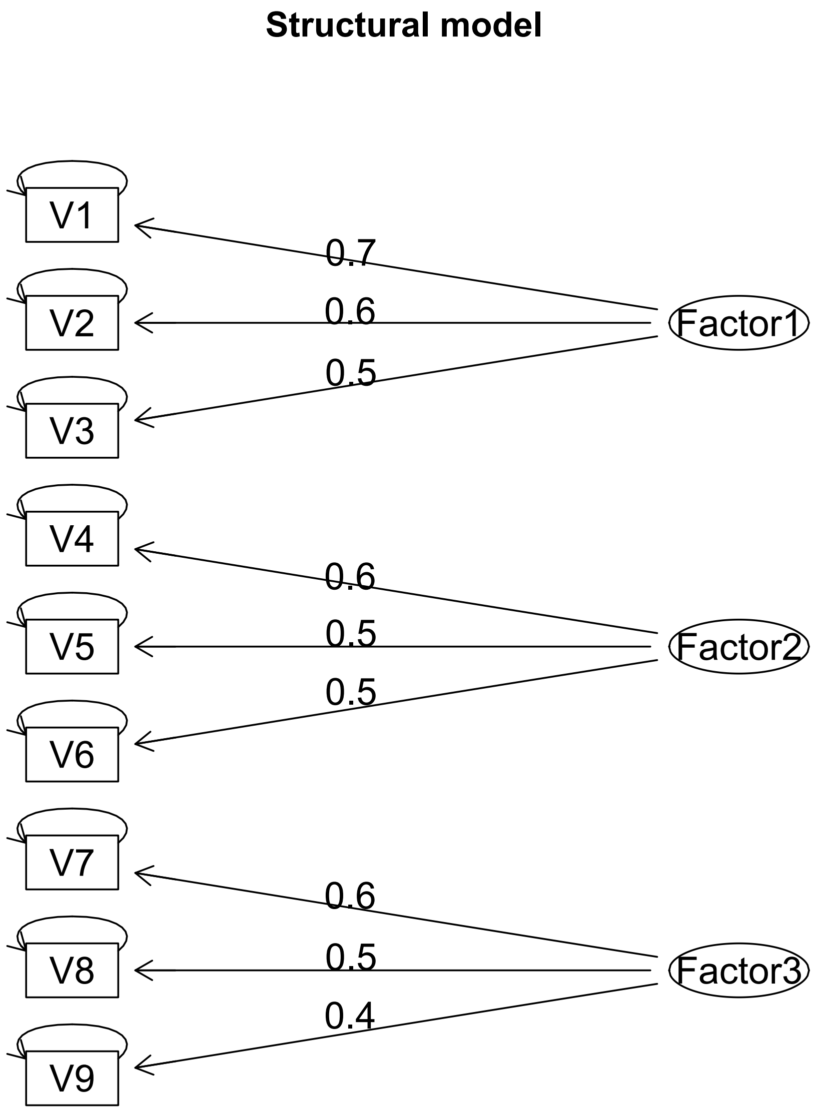
```
> f2 <- factanal(covmat=nine,factors=2,n.obs=400)
> f2
Loadings:
   Factor1 Factor2
V1 0.730   0.302
V2 0.650   0.259
V3 0.557   0.222
V4 0.287   0.633
V5 0.244   0.548
V6 0.203   0.457
V7 0.341   0.235
V8 0.290   0.198
V9 0.235   0.160
               Factor1 Factor2
SS loadings      1.704   1.237
Proportion Var   0.189   0.137
Cumulative Var   0.189   0.327

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 26.89 on 19 degrees of freedom.
The p-value is 0.107
>
```



Messy solution

```
> f3 <- factanal(covmat=nine,factors=3,n.obs=400)
> f3
Loadings:
   Factor1 Factor2 Factor3
V1 0.697   0.298   0.254
V2 0.610   0.261   0.222
V3 0.523   0.224   0.191
V4 0.242   0.631   0.182
V5 0.207   0.541   0.156
V6 0.173   0.451   0.130
V7 0.167   0.148   0.557
V8 0.139   0.123   0.464
V9 0.112           0.371

               Factor1 Factor2 Factor3
SS loadings      1.324   1.149   0.887
Proportion Var   0.147   0.128   0.099
Cumulative Var   0.147   0.275   0.373

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 0 on 12 degrees of freedom.
The p-value is 1
```

Try 3



Clearer solution

But sem is poor

```
> sem3 <- sem(mod3,nine,400)
> summary(sem3)

 Model Chisquare =  198.13   Df =  27 Pr(>Chisq) =
0
 Chisquare (null model) =  682.12   Df =  36
 Goodness-of-fit index =  0.89582
 Adjusted goodness-of-fit index =  0.82637
 RMSEA index =  0.12603   90% CI: (0.10990,
0.14278)
 Bentler-Bonnett NFI =  0.70954
 Tucker-Lewis NNFI =  0.64687
 Bentler CFI =  0.73515
 SRMR =  0.18682
 BIC =  36.356
```

Parameters are good

```
 Parameter Estimates
     Estimate Std Error z value Pr(>|z|)
F1V1 0.80000  0.055518  14.4099 0.0000e+00 V1 <--- Factor1
F1V2 0.70000  0.054290  12.8937 0.0000e+00 V2 <--- Factor1
F1V3 0.60000  0.053202  11.2778 0.0000e+00 V3 <--- Factor1
F2V4 0.70000  0.071104   9.8447 0.0000e+00 V4 <--- Factor2
F2V5 0.60000  0.066180   9.0662 0.0000e+00 V5 <--- Factor2
F2V6 0.50000  0.061705   8.1031 4.4409e-16 V6 <--- Factor2
F3V7 0.60000  0.094407   6.3554 2.0785e-10 V7 <--- Factor3
F3V8 0.50000  0.083403   5.9950 2.0355e-09 V8 <--- Factor3
F3V9 0.40000  0.073178   5.4662 4.5973e-08 V9 <--- Factor3
v1e  0.36000  0.064642   5.5691 2.5612e-08 V1 <--> V1
v2e  0.51000  0.058077   8.7815 0.0000e+00 V2 <--> V2
v3e  0.64000  0.056309  11.3659 0.0000e+00 V3 <--> V3
v4e  0.51000  0.086646   5.8860 3.9554e-09 V4 <--> V4
v5e  0.64000  0.073502   8.7072 0.0000e+00 V5 <--> V5
v6e  0.75000  0.066602  11.2609 0.0000e+00 V6 <--> V6
v7e  0.64000  0.109240   5.8587 4.6652e-09 V7 <--> V7
v8e  0.75000  0.087096   8.6112 0.0000e+00 V8 <--> V8
v9e  0.84000  0.074098  11.3363 0.0000e+00 V9 <--> V9
```

Residuals show problem

```
> round(residuals(sem3),2)
     V1   V2   V3   V4   V5   V6   V7   V8   V9
V1 0.00 0.00 0.00 0.40 0.35 0.29 0.30 0.25 0.20
V2 0.00 0.00 0.00 0.35 0.30 0.25 0.26 0.22 0.18
V3 0.00 0.00 0.00 0.30 0.26 0.22 0.23 0.19 0.15
V4 0.40 0.35 0.30 0.00 0.00 0.00 0.24 0.20 0.16
V5 0.35 0.30 0.26 0.00 0.00 0.00 0.20 0.17 0.13
V6 0.29 0.25 0.22 0.00 0.00 0.00 0.17 0.14 0.11
V7 0.30 0.26 0.23 0.24 0.20 0.17 0.00 0.00 0.00
V8 0.25 0.22 0.19 0.20 0.17 0.14 0.00 0.00 0.00
V9 0.20 0.18 0.15 0.16 0.13 0.11 0.00 0.00 0.00
```

```
> f3p <- Promax(f3)
> f3p
  V Factor1 Factor2 Factor3
V1 1  0.7588
V2 2  0.6639
V3 3  0.5691
V4 4          0.6888
V5 5          0.5904
V6 6          0.4920
V7 7                  0.597
V8 8                  0.498
V9 9                  0.398
               Factor1 Factor2 Factor3
SS loadings       1.34    1.07    0.77
Proportion Var    0.15    0.12    0.09
Cumulative Var    0.15    0.27    0.35
 With factor correlations of
        Factor1 Factor2 Factor3
Factor1    1.00    0.68    0.60
Factor2    0.68    1.00    0.55
Factor3    0.60    0.55    1.00
>
```

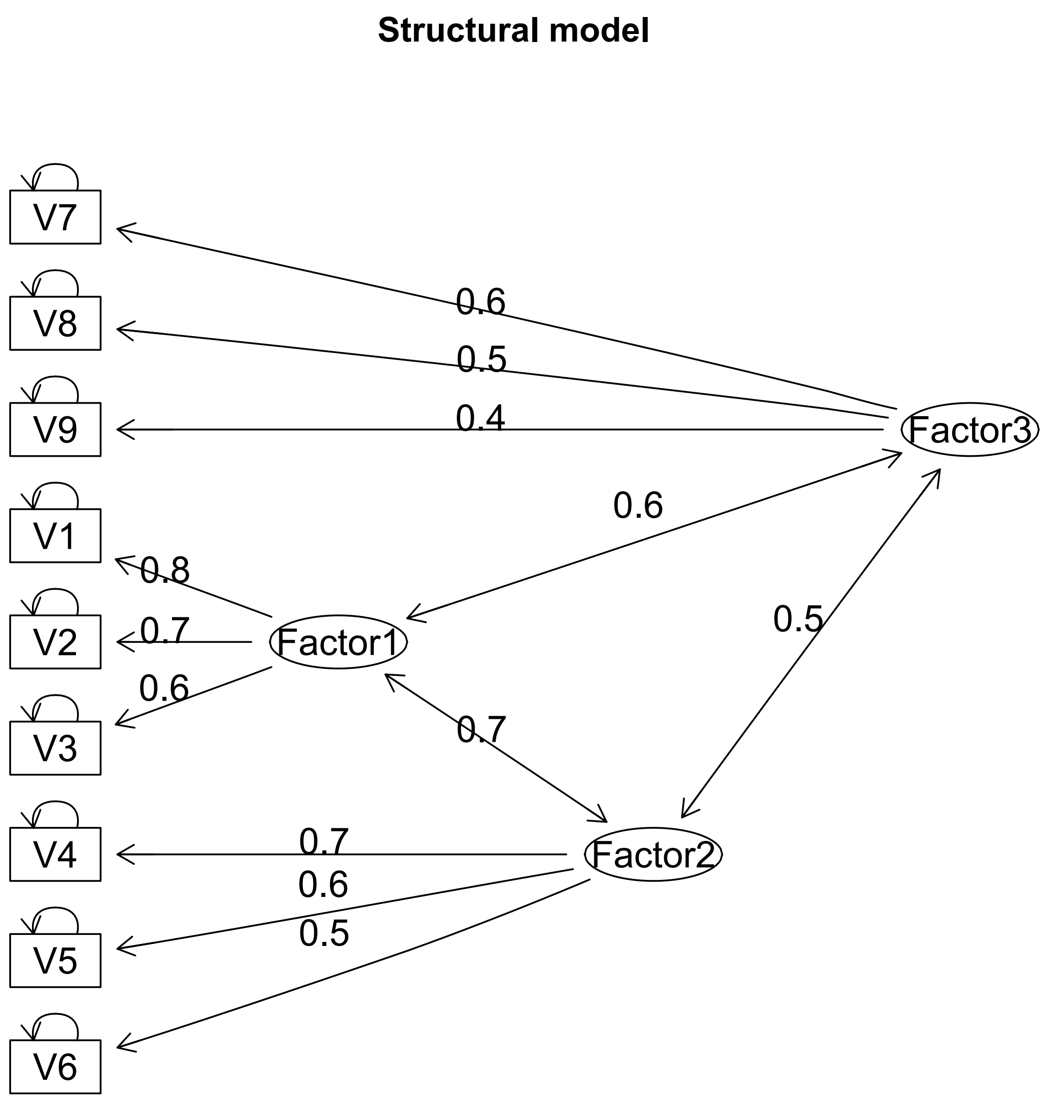
Allow oblique factors

Promax

Oblimin transformation

```
 f3o <- oblimin(loadings(f3))
Oblique rotation method Oblimin Quartimin converged.
     Factor1   Factor2   Factor3
V1  8.00e-01 -1.21e-06 -2.38e-07
V2  7.00e-01  5.10e-07  7.61e-07
V3  6.00e-01  1.39e-07  4.68e-07
V4  7.31e-07  7.00e-01  2.12e-07
V5 -6.83e-07  6.00e-01 -1.98e-07
V6 -8.24e-07  5.00e-01 -2.39e-07
V7  1.01e-05  1.04e-05  6.00e-01
V8  5.39e-06  7.27e-06  5.00e-01
V9  4.83e-06  6.06e-06  4.00e-01
Rotating matrix:
       [,1]   [,2]   [,3]
[1,]  1.463 -0.507 -0.359
[2,] -0.470  1.364 -0.221
[3,] -0.315 -0.210  1.244
Phi:
     [,1] [,2] [,3]
[1,] 1.00 0.72 0.63
[2,] 0.72 1.00 0.56
[3,] 0.63 0.56 1.00
```

Correlated factors



```
 modf3p <- structure.graph(f3p)
 modf3p
     [,1]                      [,2]    [,3]
[1,] "Factor1->V1"             "F1V1"  NA
[2,] "Factor1->V2"             "F1V2"  NA
[3,] "Factor1->V3"             "F1V3"  NA
[4,] "Factor2->V4"             "F2V4"  NA
...
[8,] "Factor3->V8"             "F3V8"  NA
[9,] "Factor3->V9"             "F3V9"  NA
10,] "V1<->V1"                 "v1e"   NA
11,] "V2<->V2"                 "v2e"   NA
..
[17,] "V8<->V8"                 "v8e"   NA
18,] "V9<->V9"                 "v9e"   NA
19,] "Factor2<->Factor1" "rF2F1" NA
20,] "Factor3<->Factor1" "rF3F1" NA
21,] "Factor3<->Factor2" "rF3F2" NA
22,] "Factor1<->Factor1" NA      "1"
23,] "Factor2<->Factor2" NA      "1"
```

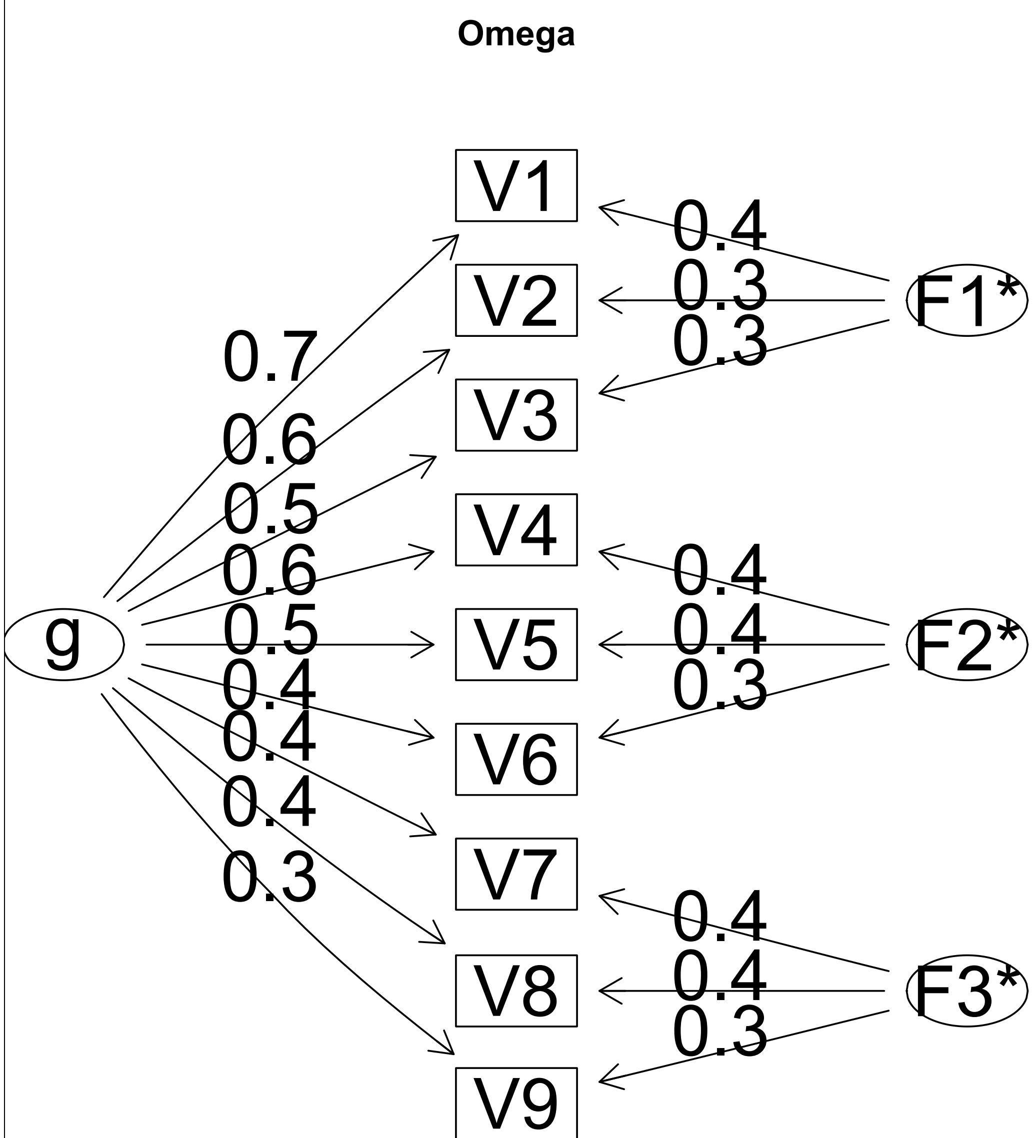
sem model

Very good fit!

```
> sem.3p <- sem(modf3p,nine, 200)
> summary(sem.3p)

 Model Chisquare =  8.0118e-09   Df =  24 Pr(>Chisq) = 1
 Chisquare (null model) =  340.21   Df =  36
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  1
 RMSEA index =  0   90% CI: (NA, NA)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1.1183
 Bentler CFI =  1
 SRMR =  1.6130e-06
 BIC =  -127.16

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-1.83e-05  7.07e-06  1.80e-05  1.78e-05  2.77e-05  4.49e-05
```

Parameters

```
 Parameter Estimates
      Estimate Std Error z value Pr(>|z|)
F1V1  0.80     0.068775  11.6321 0.0000e+00 V1 <--- Factor1
F1V2  0.70     0.070168   9.9761 0.0000e+00 V2 <--- Factor1
F1V3  0.60     0.072095   8.3224 0.0000e+00 V3 <--- Factor1
F2V4  0.70     0.078136   8.9587 0.0000e+00 V4 <--- Factor2
F2V5  0.60     0.077947   7.6975 1.3767e-14 V5 <--- Factor2
F2V6  0.50     0.079099   6.3212 2.5954e-10 V6 <--- Factor2
F3V7  0.60     0.092802   6.4654 1.0106e-10 V7 <--- Factor3
F3V8  0.50     0.089432   5.5908 2.2607e-08 V8 <--- Factor3
F3V9  0.40     0.088929   4.4980 6.8602e-06 V9 <--- Factor3
v1e   0.36     0.068296   5.2712 1.3554e-07 V1 <--> V1
v2e   0.51     0.069492   7.3389 2.1538e-13 V2 <--> V2
v3e   0.64     0.075297   8.4997 0.0000e+00 V3 <--> V3
...
v8e   0.75     0.096220   7.7947 6.4393e-15 V8 <--> V8
v9e   0.84     0.095905   8.7587 0.0000e+00 V9 <--> V9
rF2F1 0.72     0.070286  10.2439 0.0000e+00 Factor1 <-->
Factor2
rF3F1 0.63     0.092891   6.7821 1.1841e-11 Factor1 <-->
Factor3
rF3F2 0.56     0.107225   5.2227 1.7637e-07 Factor2 <-->
```



Bifactor model from exploratory factor analysis

sem for bifactor

```
> om.bf <- omega(nine)
> mod.bf <- omega.graph(om.bf)
> mod.bf
      [,1]          [,2]    [,3]
 [1,] "g->V1"       "V1"    NA
 [2,] "g->V2"       "V2"    NA
 [3,] "g->V3"       "V3"    NA
...
 [9,] "g->V9"       "V9"    NA
[10,] "F1*->V1"     "F1*V1" NA
[11,] "F1*->V2"     "F1*V2" NA
...
[17,] "F3*->V8"     "F3*V8" NA
[18,] "F3*->V9"     "F3*V9" NA
[19,] "V1<->V1"     "e1"    NA
[20,] "V2<->V2"     "e2"    NA
...
[27,] "V9<->V9"     "e9"    NA
[28,] "F1*<->F1*" NA        "1"
[29,] "F2*<->F2*" NA        "1"
[30,] "F3*<->F3*" NA        "1"
[31,] "g <->g"    NA        "1"
```

Very good fit

```
> sem.bf <- sem(mod.bf,nine,400)
> summary(sem.bf)

 Model Chisquare =  6.8971e-08   Df =  18 Pr(>Chisq) = 1
 Chisquare (null model) =  682.12   Df =  36
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  1
 RMSEA index =  0   90% CI: (NA, NA)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1.0557
 Bentler CFI =  1
 SRMR =  1.9516e-06
 BIC =  -107.85

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-1.23e-04 -3.16e-06  1.20e-05  1.46e-05  3.21e-05  1.18e-04
```

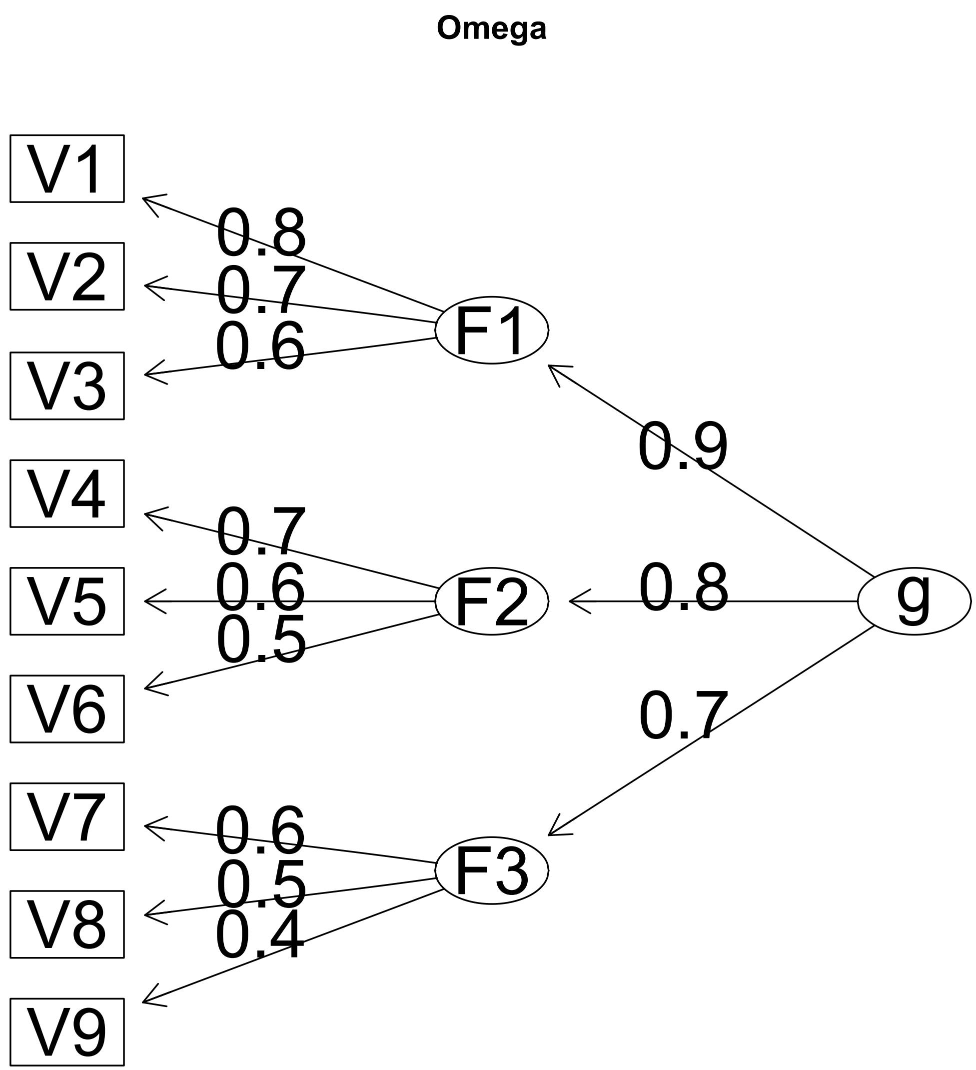
Parameters

```
Parameter Estimates
       Estimate Std Error z value Pr(>|z|)
V1     0.72000  0.066569  10.8159 0.0000e+00 V1 <--- g
V2     0.63000  0.067357   9.3531 0.0000e+00 V2 <--- g
V3     0.54000  0.068033   7.9373 1.9984e-15 V3 <--- g
V4     0.56000  0.061190   9.1518 0.0000e+00 V4 <--- g
V5     0.48000  0.061243   7.8376 4.6629e-15 V5 <--- g
V6     0.40000  0.061287   6.5267 6.7254e-11 V6 <--- g
V7     0.42000  0.057942   7.2487 4.2077e-13 V7 <--- g
V8     0.35000  0.058501   5.9828 2.1929e-09 V8 <--- g
V9     0.28000  0.058954   4.7494 2.0397e-06 V9 <--- g
F1*V1 0.34871  0.138809   2.5122 1.1999e-02 V1 <--- F1*
F1*V2 0.30512  0.133120   2.2920 2.1904e-02 V2 <--- F1*
F1*V3 0.26154  0.127954   2.0440 4.0953e-02 V3 <--- F1*
F2*V4 0.41999  0.114795   3.6586 2.5356e-04 V4 <--- F2*
F2*V5 0.36001  0.104622   3.4410 5.7950e-04 V5 <--- F2*
F2*V6 0.29999  0.095149   3.1529 1.6167e-03 V6 <--- F2*
F3*V7 0.42847  0.127678   3.3558 7.9125e-04 V7 <--- F3*
F3*V8 0.35708  0.111371   3.2062 1.3449e-03 V8 <--- F3*
F3*V9 0.28567  0.095973   2.9765 2.9154e-03 V9 <--- F3*
e1    0.36000  0.064640   5.5692 2.5584e-08 V1 <--> V1
e2    0.51000  0.058076   8.7817 0.0000e+00 V2 <--> V2
e3    0.64000  0.056312  11.3653 0.0000e+00 V3 <--> V3
```

Tiny residuals

```
> round(residuals(sem.bf),4)
   V1 V2 V3 V4 V5 V6 V7 V8 V9
V1  0  0  0  0  0  0  0  0  0
V2  0  0  0  0  0  0  0  0  0
V3  0  0  0  0  0  0  0  0  0
V4  0  0  0  0  0  0  0  0  0
V5  0  0  0  0  0  0  0  0  0
V6  0  0  0  0  0  0  0  0  0
V7  0  0  0  0  0  0  0  0  0
V8  0  0  0  0  0  0  0  0  0
V9  0  0  0  0  0  0  0  0  0
```

Hierarchical model



Hierarchical model

```
> mod.h
      [,1]        [,2]    [,3]
 [1,] "g->F1"     "gF1"   NA
 [2,] "g->F2"     "gF2"   NA
 [3,] "g->F3"     "gF3"   NA
 [4,] "F1->V1"    "F1V1"  NA
 [5,] "F1->V2"    "F1V2"  NA
 [6,] "F1->V3"    "F1V3"  NA
 [7,] "F2->V4"    "F2V4"  NA
 [8,] "F2->V5"    "F2V5"  NA
 [9,] "F2->V6"    "F2V6"  NA
[10,] "F3->V7"    "F3V7"  NA
[11,] "F3->V8"    "F3V8"  NA
[12,] "F3->V9"    "F3V9"  NA
[13,] "V1<->V1"   "e1"    NA
[14,] "V2<->V2"   "e2"    NA
...
[21,] "V9<->V9"   "e9"    NA
[22,] "F1<->F1"   "ef1"   NA
[23,] "F2<->F2"   "ef2"   NA
[24,] "F3<->F3"   "ef3"   NA
[25,] "g<->g"     NA      "1"
```

Model is not identified!

```
> sem.h <- sem(mod.h,nine,400)
Warning message:
In sem.default(ram = ram, S = S, N = N, param.names = pars,
var.names = vars,  :
  Negative parameter variances.
Model is probably underidentified.
```

But fit is great!

```
Model Chisquare =  4.1023e-08   Df =  21 Pr(>Chisq) = 1
 Chisquare (null model) =  682.12   Df =  36
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  1
 RMSEA index =  0   90% CI: (NA, NA)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1.0557
 Bentler CFI =  1
 SRMR =  1.5605e-06
 BIC =  -125.82
```

```
 Parameter Estimates
      Estimate Std Error z value Pr(>|z|)
gF1   0.95948        NaN     NaN         NaN F1 <--- g
gF2   0.80863        NaN     NaN         NaN F2 <--- g
gF3   0.66758        NaN     NaN         NaN F3 <--- g
F1V1  0.75040        NaN     NaN         NaN V1 <--- F1
F1V2  0.65660        NaN     NaN         NaN V2 <--- F1
F1V3  0.56280        NaN     NaN         NaN V3 <--- F1
F2V4  0.69253        NaN     NaN         NaN V4 <--- F2
F2V5  0.59360        NaN     NaN         NaN V5 <--- F2
F2V6  0.49466        NaN     NaN         NaN V6 <--- F2
F3V7  0.62914        NaN     NaN         NaN V7 <--- F3
F3V8  0.52428        NaN     NaN         NaN V8 <--- F3
F3V9  0.41943        NaN     NaN         NaN V9 <--- F3
e1    0.36000  0.048231   7.4640 8.3933e-14 V1 <--> V1
e2    0.51000  0.049076  10.3919 0.0000e+00 V2 <--> V2
...
e8    0.75001  0.067952  11.0373 0.0000e+00 V8 <--> V8
e9    0.84000  0.067730  12.4022 0.0000e+00 V9 <--> V9
ef1   0.21595        NaN     NaN         NaN F1 <--> F1
ef2   0.36780        NaN     NaN         NaN F2 <--> F2
ef3   0.46386        NaN     NaN         NaN F3 <--> F3
```

Warning messages bad identification

```
 Aliased parameters: gF1 gF2 gF3 F1V1 F1V2 F1V3 F2V4
F2V5 F2V6 F3V7 F3V8 F3V9 ef1 ef2 ef3
Warning message:
In sqrt(diag(object$cov)) : NaNs produced
>
```

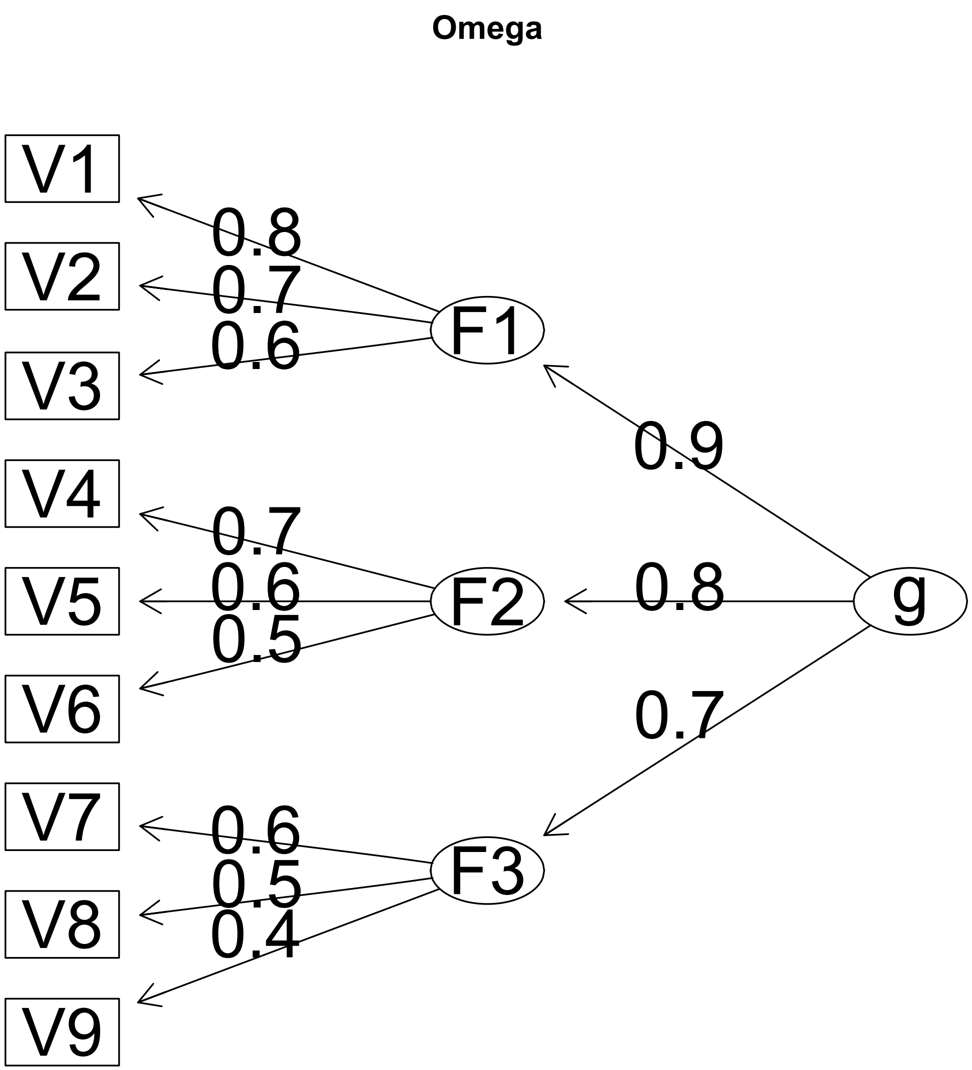
Problem of identication

I. More parameters than unknowns (negative degrees of freedom

II. Parameters undefined in terms of particular paths

III. Consider the hierarchical model -- need to fix some parameters for identification

Hierarchical model



```
> mod.h.f <- edit(mod.h)
> mod.h.f
      col1       col2   col3
 [1,] "g->F1"    "gF1"  NA
 [2,] "g->F2"    "gF2"  NA
 [3,] "g->F3"    "gF3"  NA
 [4,] "F1->V1"   NA     "1"
 [5,] "F1->V2"   "F1V2" NA
 [6,] "F1->V3"   "F1V3" NA
 [7,] "F2->V4"   NA     "1"
 [8,] "F2->V5"   "F2V5" NA
 [9,] "F2->V6"   "F2V6" NA
[10,] "F3->V7"   NA     "1"
[11,] "F3->V8"   "F3V8" NA
[12,] "F3->V9"   "F3V9" NA
[13,] "V1<->V1"  "e1"   NA
...
[21,] "V9<->V9"  "e9"   NA
[22,] "F1<->F1"  "ef1"  NA
[23,] "F2<->F2"  "ef2"  NA
[24,] "F3<->F3"  "ef3"  NA
[25,] "g<->g"    NA     "1"
```

Fix some covariances to 1

Fits are very good increased df

```
> sem.h.f <- sem(mod.h.f,nine,400)
> summary(sem.h.f)

 Model Chisquare =  2.8126e-08   Df =  24 Pr(>Chisq) = 1
 Chisquare (null model) =  682.12   Df =  36
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  1
 RMSEA index =  0   90% CI: (NA, NA)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1.0557
 Bentler CFI =  1
 SRMR =  1.3197e-06
 BIC =  -143.80
```

```
Parameter Estimates
      Estimate Std Error z value Pr(>|z|)
gF1   0.72001  0.065075  11.0642 0.0000e+00 F1 <--- g
gF2   0.56000  0.059879   9.3521 0.0000e+00 F2 <--- g
gF3   0.42000  0.056295   7.4606 8.6153e-14 F3 <--- g
F1V2  0.87500  0.073697  11.8729 0.0000e+00 V2 <--- F1
F1V3  0.75000  0.071288  10.5207 0.0000e+00 V3 <--- F1
F2V5  0.85714  0.099649   8.6017 0.0000e+00 V5 <--- F2
F2V6  0.71428  0.093817   7.6135 2.6645e-14 V6 <--- F2
F3V8  0.83334  0.140550   5.9291 3.0455e-09 V8 <--- F3
F3V9  0.66667  0.127252   5.2390 1.6148e-07 V9 <--- F3
e1    0.36000  0.048239   7.4628 8.4599e-14 V1 <--> V1
e2    0.51000  0.049077  10.3917 0.0000e+00 V2 <--> V2
e3    0.64000  0.053176  12.0355 0.0000e+00 V3 <--> V3
e4    0.51000  0.059713   8.5408 0.0000e+00 V4 <--> V4
e5    0.64000  0.058811  10.8823 0.0000e+00 V5 <--> V5
e6    0.75001  0.061226  12.2498 0.0000e+00 V6 <--> V6
e7    0.64000  0.072692   8.8044 0.0000e+00 V7 <--> V7
e8    0.75000  0.067954  11.0370 0.0000e+00 V8 <--> V8
e9    0.84000  0.067731  12.4021 0.0000e+00 V9 <--> V9
ef1   0.12159  0.073314   1.6585 9.7211e-02 F1 <--> F1
ef2   0.17640  0.058281   3.0268 2.4717e-03 F2 <--> F2
ef3   0.18360  0.057752   3.1791 1.4773e-03 F3 <--> F3
```

Standardized estimates

```
> std.coef(sem.h.f)
                 Std. Estimate
 1  gF1  0.9            F1 <--- g
 2  gF2  0.8            F2 <--- g
 3  gF3  0.7            F3 <--- g
4        0.8           V1 <--- F1
5   F1V2 0.7           V2 <--- F1
6   F1V3 0.6           V3 <--- F1
7        0.7           V4 <--- F2
8   F2V5 0.6           V5 <--- F2
9   F2V6 0.5           V6 <--- F2
10       0.6           V7 <--- F3
11  F3V8 0.5           V8 <--- F3
12  F3V9 0.4           V9 <--- F3
```

Apply model to sample data

```
>  R <- sim.hierarchical(n=200)
> R
 $model (Population correlation matrix)
     V1   V2   V3   V4   V5   V6   V7   V8   V9
V1 1.00 0.56 0.48 0.40 0.35 0.29 0.30 0.25 0.20
V2 0.56 1.00 0.42 0.35 0.30 0.25 0.26 0.22 0.18
V3 0.48 0.42 1.00 0.30 0.26 0.22 0.23 0.19 0.15
V4 0.40 0.35 0.30 1.00 0.42 0.35 0.24 0.20 0.16
V5 0.35 0.30 0.26 0.42 1.00 0.30 0.20 0.17 0.13
V6 0.29 0.25 0.22 0.35 0.30 1.00 0.17 0.14 0.11
V7 0.30 0.26 0.23 0.24 0.20 0.17 1.00 0.30 0.24
V8 0.25 0.22 0.19 0.20 0.17 0.14 0.30 1.00 0.20
V9 0.20 0.18 0.15 0.16 0.13 0.11 0.24 0.20 1.00
$r  (Sample correlation matrix  for sample size =  200 )
     V1   V2   V3   V4   V5    V6   V7    V8   V9
V1 1.00 0.66 0.49 0.35 0.33 0.240 0.35 0.391 0.38
V2 0.66 1.00 0.45 0.36 0.31 0.152 0.36 0.349 0.21
V3 0.49 0.45 1.00 0.38 0.26 0.232 0.23 0.290 0.22
V4 0.35 0.36 0.38 1.00 0.39 0.309 0.18 0.149 0.20
V5 0.33 0.31 0.26 0.39 1.00 0.271 0.20 0.158 0.10
```

sem statistics for n = 200

```
> sem.h.f.s <- sem(mod.h.f,R$r,200)
> summary(sem.h.f.s)

 Model Chisquare =  24.474   Df =  24 Pr(>Chisq) =
0.43479
 Chisquare (null model) =  409.03   Df =  36
 Goodness-of-fit index =  0.97365
 Adjusted goodness-of-fit index =  0.9506
 RMSEA index =  0.0099606   90% CI: (NA, 0.058646)
 Bentler-Bonnett NFI =  0.94017
 Tucker-Lewis NNFI =  0.9981
 Bentler CFI =  0.99873
 SRMR =  0.037009
 BIC =  -102.69
```

```
 Parameter Estimates
     Estimate   Std Error z value  Pr(>|z|)
gF1    0.8693327 0.092409   9.40744 0.0000e+00 F1 <--- g
gF2    0.4409825 0.075778   5.81936 5.9072e-09 F2 <--- g
gF3    0.4198795 0.080600   5.20943 1.8942e-07 F3 <--- g
F1V2   0.8970144 0.083345  10.76268 0.0000e+00 V2 <--- F1
F1V3   0.7063272 0.086836   8.13401 4.4409e-16 V3 <--- F1
F2V5   0.8660712 0.161514   5.36220 8.2212e-08 V5 <--- F2
F2V6   0.6577838 0.142485   4.61650 3.9026e-06 V6 <--- F2
F3V8   1.0740000 0.240679   4.46238 8.1053e-06 V8 <--- F3
F3V9   0.9045756 0.226100   4.00078 6.3135e-05 V9 <--- F3
e1     0.2771631 0.055516   4.99250 5.9603e-07 V1 <--> V1
e2     0.4183790 0.057656   7.25651 3.9724e-13 V2 <--> V2
e3     0.6393765 0.072127   8.86456 0.0000e+00 V3 <--> V3
e4     0.5393050 0.092560   5.82656 5.6582e-09 V4 <--> V4
e5     0.6544402 0.088630   7.38400 1.5365e-13 V5 <--> V5
e6     0.8006633 0.090994   8.79911 0.0000e+00 V6 <--> V6
e7     0.8143601 0.089335   9.11581 0.0000e+00 V7 <--> V7
e8     0.7858592 0.092967   8.45309 0.0000e+00 V8 <--> V8
e9     0.8480952 0.091921   9.22631 0.0000e+00 V9 <--> V9
ef1   -0.0329036 0.126991  -0.25910 7.9556e-01 F1 <--> F1
ef2    0.2662290 0.085303   3.12097 1.8026e-03 F2 <--> F2
ef3    0.0093447 0.045828   0.20391 8.3843e-01 F3 <--> F3
```

Residuals

```
> round(residuals(sem.h.f.s),2)
      V1    V2    V3    V4    V5    V6    V7    V8    V9
V1  0.00  0.01 -0.02 -0.03  0.00 -0.01 -0.01  0.00  0.05
V2  0.01  0.00 -0.01  0.01  0.01 -0.07  0.04  0.00 -0.08
V3 -0.02 -0.01  0.00  0.11  0.02  0.05 -0.02  0.01 -0.01
V4 -0.03  0.01  0.11  0.00 -0.01  0.01  0.00 -0.05  0.04
V5  0.00  0.01  0.02 -0.01  0.00  0.01  0.04 -0.01 -0.04
V6 -0.01 -0.07  0.05  0.01  0.01  0.00  0.03 -0.04  0.11
V7 -0.01  0.04 -0.02  0.00  0.04  0.03  0.00  0.00 -0.05
V8  0.00  0.00  0.01 -0.05 -0.01 -0.04  0.00  0.00  0.05
V9  0.05 -0.08 -0.01  0.04 -0.04  0.11 -0.05  0.05  0.00
```