Estimating Latent Models

Part II: Exploratory and confirmatory factor models

Data = Model + Error

I. A fundamental process in science is the modeling of data. Statistics are estimates of how well the model fits the data and reflect the size of the error

II. Residuals = Data - Model

III. Theoretical models attempt to represent the data

IV. Empirical fits can be evaluated in terms of size of residuals

Fitting latent models

I. Latent variable models are path (matrix) models of observed covariances

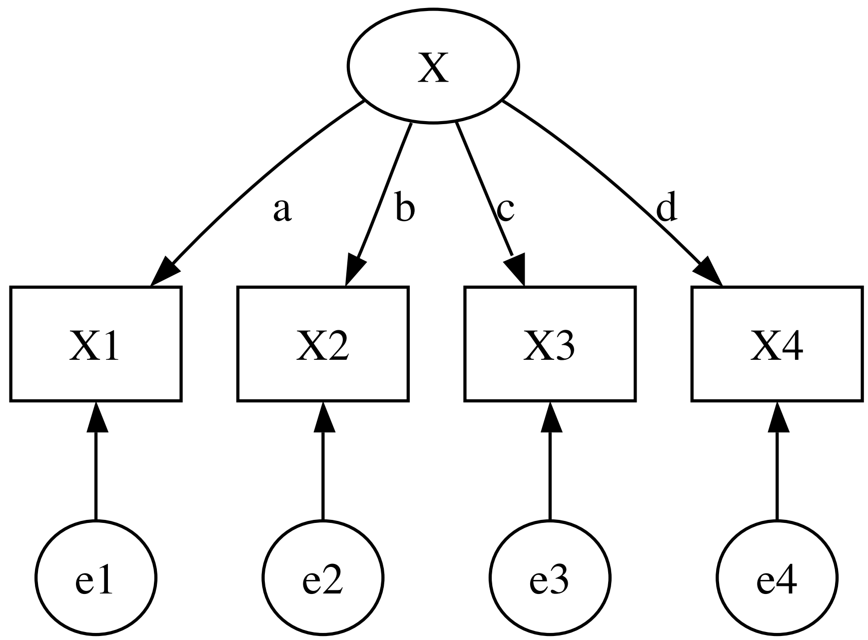
II. Parameters are estimated by finding model parameters that minimize some criterion of misfit.

III. Need to consider multiple multiples of data and compare adequacy of fit versus complexity of the model

Know the fundamentals

I. An understanding of the fundamental concepts allows one to derive the applications

The basic congeneric test model



Example data

	V1	V2	V3	V4
V1	1.00	0.54	0.52	0.41
V2	0.54	1.00	0.41	0.32
V3	0.52	0.41	1.00	0.32
V4	0.41	0.32	0.32	1.00

The 1 factor model

```
path label initial estimate
[1,] "theta -> V1" "a" NA
[2,] "theta -> V2" "b" NA
[3,] "theta -> V3" "c" NA
[4,] "theta -> V4" "d" NA
[5,] "V1 <-> V1" "u" NA
[6,] "V2 <-> V2" "v" NA
[7,] "V3 <-> V3" "w" NA
[8,] "V4 <-> V4" "x" NA
[9,] "theta <-> theta" NA "1"
```

Model statistics

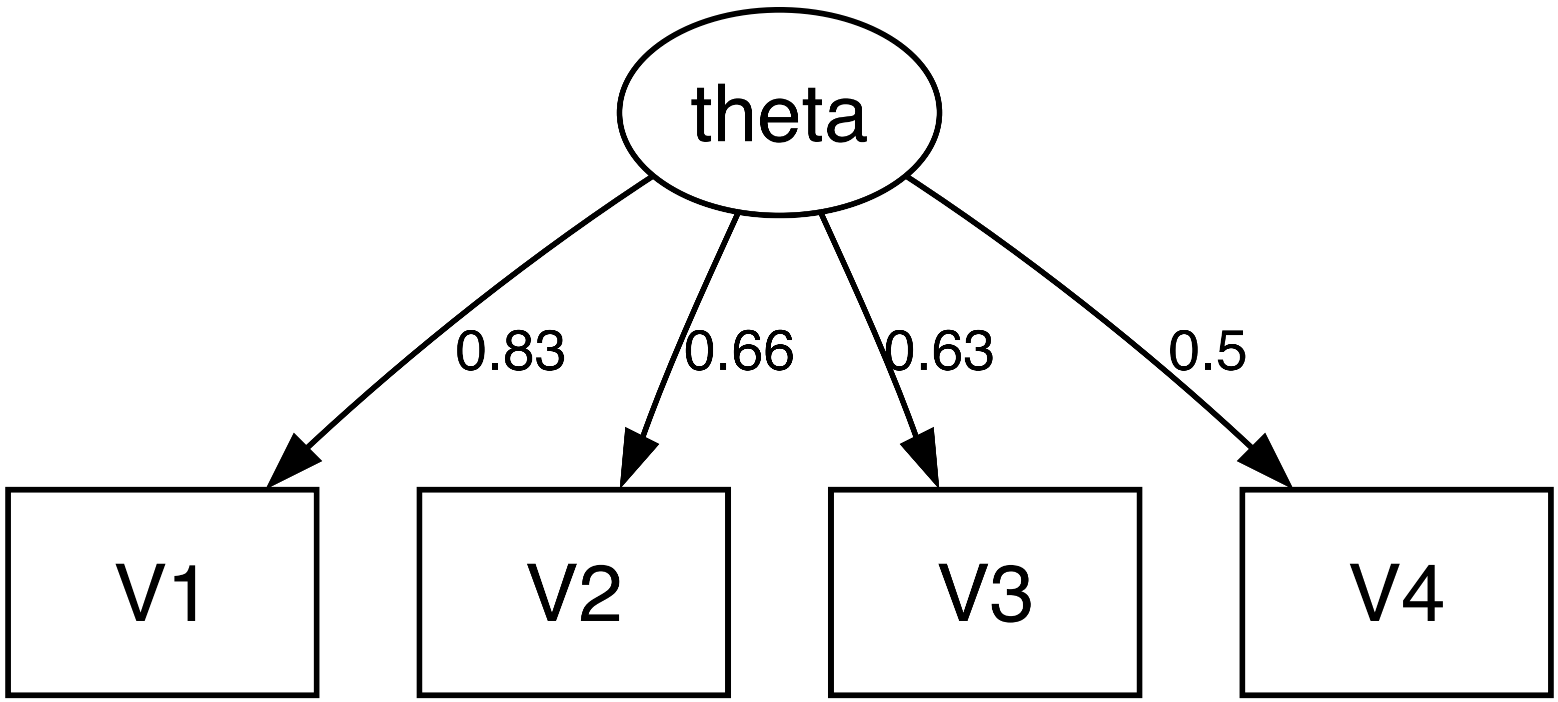
Model Chisquare = 0.46 Df = 2 Pr(>Chisq) = 0.795
Chisquare (null model) = 910 Df = 6
Goodness-of-fit index = 1
Adjusted goodness-of-fit index = 0.999
RMSEA index = 0 90% CI: (NA, 0.0398)
Bentler-Bonnett NFI = 1
Tucker-Lewis NNFI = 1.01
Bentler CFI = 1
BIC = -13.4

Path coefficients

Estimate Std Error z value Pr(>|z|)
a 0.829 0.0320 25.90 0 V1 <--- theta
b 0.657 0.0325 20.23 0 V2 <--- theta
c 0.632 0.0325 19.43 0 V3 <--- theta
d 0.503 0.0340 14.80 0 V4 <--- theta
u 0.316 0.0346 9.12 0 V1 <--> V1
v 0.580 0.0334 17.35 0 V2 <--> V2
w 0.604 0.0337 17.94 0 V3 <--> V3
x 0.776 0.0382 20.31 0 V4 <--> V4

```
> round(residuals(sem.congeneric), 2)
   V1    V2    V3   V4
V1  0  0.00  0.00 0.00
V2  0  0.00 -0.01 0.00
V3  0 -0.01  0.00 0.01
V4  0  0.00  0.01 0.00
```

Path coefficients



A more constrained model: Tau Equivalent

```
     path               label initial estimate
[1,] "theta -> V1"      "a"   NA
[2,] "theta -> V2"      "a"   NA
[3,] "theta -> V3"      "a"   NA
[4,] "theta -> V4"      "a"   NA
[5,] "V1 <-> V1"        "u"   NA
[6,] "V2 <-> V2"        "v"   NA
[7,] "V3 <-> V3"        "w"   NA
[8,] "V4 <-> V4"        "x"   NA
[9,] "theta <-> theta"  NA    "1"
```

Goodness of fit

Model Chisquare = 56.1 Df = 5 Pr(>Chisq) = 7.64e-11
Chisquare (null model) = 910 Df = 6
Goodness-of-fit index = 0.974
Adjusted goodness-of-fit index = 0.947
RMSEA index = 0.101 90% CI: (0.0783, 0.126)
Bentler-Bonnett NFI = 0.938
Tucker-Lewis NNFI = 0.932
Bentler CFI = 0.943
BIC = 21.6

Parameter Estimates

	Estimate	Std Error	z value	Pr(>|z|)	
a	0.668	0.0202	33.2	0	V1 <--- theta
u	0.448	0.0270	16.6	0	V1 <--> V1
v	0.565	0.0315	18.0	0	V2 <--> V2
w	0.576	0.0319	18.1	0	V3 <--> V3
x	0.730	0.0386	18.9	0	V4 <--> V4

Residuals for tau model

```
round(residuals(sem.tau),2)
      V1    V2    V3    V4
V1  0.11  0.10  0.08 -0.03
V2  0.10  0.00 -0.04 -0.12
V3  0.08 -0.04 -0.02 -0.12
V4 -0.03 -0.12 -0.12 -0.15
```

Comparisons of fit

	chisq	df
congeneric	0.46	2
tau	56.13	5
parallel	91.23	8
fixed	98.57	9

2 uncorrelated latent traits

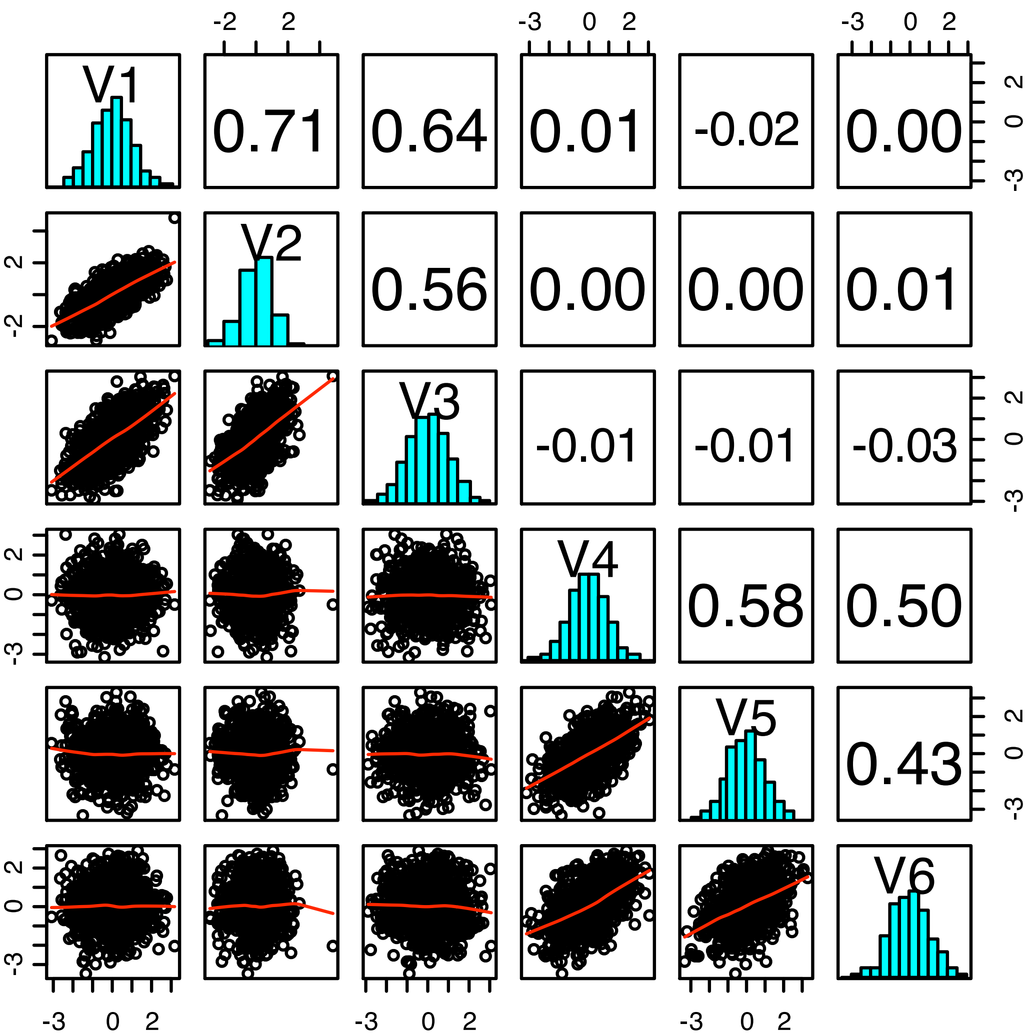
F

	theta1	theta2
V1	0.9	0.0
V2	0.8	0.0
V3	0.7	0.0
V4	0.0	0.8
V5	0.0	0.7
V6	0.0	0.6

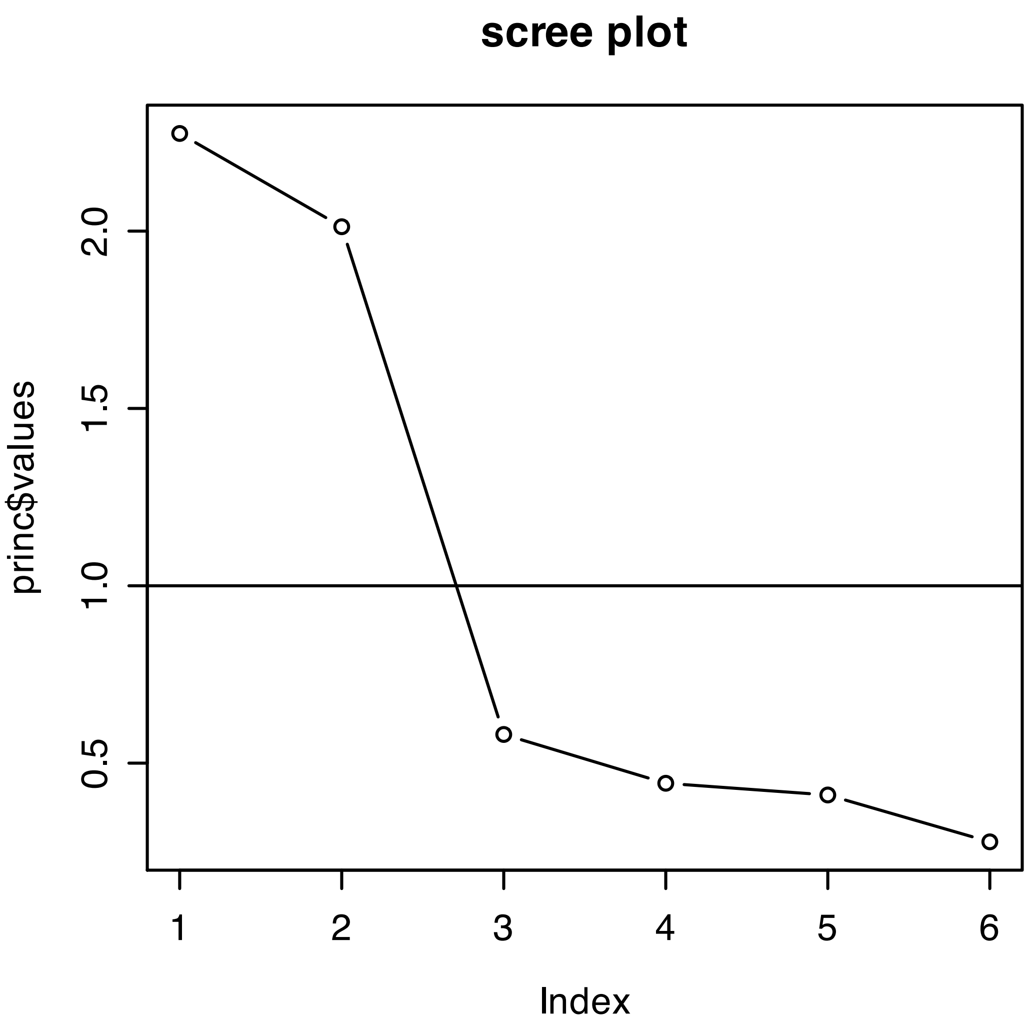
$R = FF' + U^2$

	V1	V2	V3	V4	V5	V6
V1	1.00	0.72	0.63	0.00	0.00	0.00
V2	0.72	1.00	0.56	0.00	0.00	0.00
V3	0.63	0.56	1.00	0.00	0.00	0.00
V4	0.00	0.00	0.00	1.00	0.56	0.48
V5	0.00	0.00	0.00	0.56	1.00	0.42
V6	0.00	0.00	0.00	0.48	0.42	1.00

Sample correlations



Scree plot



Exploratory factor analysis

	Factor1	Factor2
V1	0.894	
V2	0.791	
V3	0.713	
V4		0.819
V5		0.709
V6		0.611

	Factor1	Factor2
SS loadings	1.934	1.548
Proportion Var	0.322	0.258
Cumulative Var	0.322	0.580

The chi square statistic is 3.97 on 4 df

sem: confirmatory

```
model.two
     path                label initial estimate
 [1,] "theta1 ->  V1"     "a"   NA
 [2,] "theta1 -> V2"      "b"   NA
 [3,] "theta1 -> V3"      "c"   NA
 [4,] "theta2 -> V4"      "d"   NA
 [5,] "theta2 -> V5"      "e"   NA
 [6,] "theta2 -> V6"      "f"   NA
 [7,] "theta1 <-> theta2" "g"   NA
 [8,] "V1 <-> V1"         "u"   NA
 [9,] "V2 <-> V2"         "v"   NA
[10,] "V3 <-> V3"         "w"   NA
[11,] "V4 <-> V4"         "x"   NA
[12,] "V5 <-> V5"         "y"   NA
[13,] "V6 <-> V6"         "z"   NA
[14,] "theta1 <-> theta1" NA    "1"
[15,] "theta2 <-> theta2" NA    "1"
```

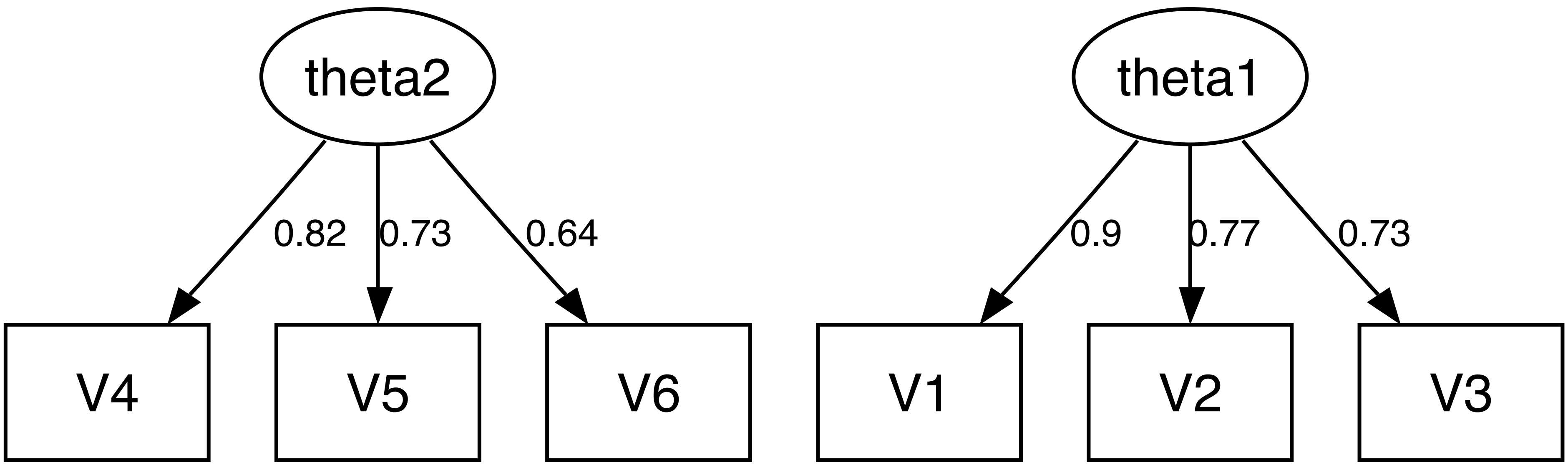
Goodness of fit

Model Chisquare = 4.9 Df = 8 Pr(>Chisq) = 0.768
Chisquare (null model) = 2004 Df = 15
Goodness-of-fit index = 0.998
Adjusted goodness-of-fit index = 0.996
RMSEA index = 0 90% CI: (NA, 0.0255)
Bentler-Bonnett NFI = 0.998
Tucker-Lewis NNFI = 1.00
Bentler CFI = 1
BIC = -50.4

Parameters

```
  Estimate Std Error z value Pr(>|z|)
a  0.89759 0.0283    31.6630 0.00e+00 V1 <--- theta1
b  0.77024 0.0282    27.3110 0.00e+00 V2 <--- theta1
c  0.72734 0.0301    24.1851 0.00e+00 V3 <--- theta1
d  0.81649 0.0340    24.0207 0.00e+00 V4 <--- theta2
e  0.73146 0.0345    21.1873 0.00e+00 V5 <--- theta2
f  0.64406 0.0348    18.5048 0.00e+00 V6 <--- theta2
g -0.00361 0.0383    -0.0941 9.25e-01 theta2 <--> theta1
u  0.20420 0.0267     7.6376 2.22e-14 V1 <--> V1
v  0.35373 0.0244    14.5264 0.00e+00 V2 <--> V2
w  0.51037 0.0282    18.1124 0.00e+00 V3 <--> V3
x  0.32728 0.0391     8.3602 0.00e+00 V4 <--> V4
y  0.52691 0.0375    14.0570 0.00e+00 V5 <--> V5
z  0.69579 0.0385    18.0894 0.00e+00 V6 <--> V6
```

Unstandardized paths



Factor invariance

I. Are the factor loadings in the first scale the same as the factor loadings in the second scale?

II. (Usually asked if the two measures are thought to be measures of the same construct and are highly correlated)

III. Test the model of equality of ordered loadings (e.g., a=d, b=e, c=f)

Model factor invariance

```
      path                   label initial estimate
 [1,] "theta1 ->  V1"      "a"   NA
 [2,] "theta1 -> V2"       "b"   NA
 [3,] "theta1 -> V3"       "c"   NA
 [4,] "theta2 -> V4"       "a"   NA
 [5,] "theta2 -> V5"       "b"   NA
 [6,] "theta2 -> V6"       "c"   NA
 [7,] "V1 <-> V1"          "u"   NA
 [8,] "V2 <-> V2"          "v"   NA
 [9,] "V3 <-> V3"          "w"   NA
[10,] "V4 <-> V4"          "x"   NA
[11,] "V5 <-> V5"          "y"   NA
[12,] "V6 <-> V6"          "z"   NA
[13,] "theta1 <-> theta1" NA    "1"
[14,] "theta2 <-> theta2" NA    "1"
```

Parameters

	Estimate	Std Error	z value	Pr(>|z|)	
a	0.862	0.0214	40.26	0	V1 <--- theta1
b	0.750	0.0215	34.91	0	V2 <--- theta1
c	0.690	0.0225	30.67	0	V3 <--- theta1
u	0.211	0.0249	8.48	0	V1 <--> V1
v	0.350	0.0235	14.86	0	V2 <--> V2
w	0.513	0.0277	18.53	0	V3 <--> V3
x	0.312	0.0315	9.89	0	V4 <--> V4
y	0.536	0.0330	16.24	0	V5 <--> V5
z	0.692	0.0371	18.66	0	V6 <--> V6

Goodness of fit

Model Chisquare = 10.7 Df = 12 Pr(>Chisq) = 0.557
Chisquare (null model) = 2004 Df = 15
Goodness-of-fit index = 0.996
Adjusted goodness-of-fit index = 0.994
RMSEA index = 0 90% CI: (NA, 0.0293)
Bentler-Bonnett NFI = 0.995
Tucker-Lewis NNFI = 1
Bentler CFI = 1
BIC = -72.2

Are the factors invariant?

I. Population values are different

   A. .9,.8,.7

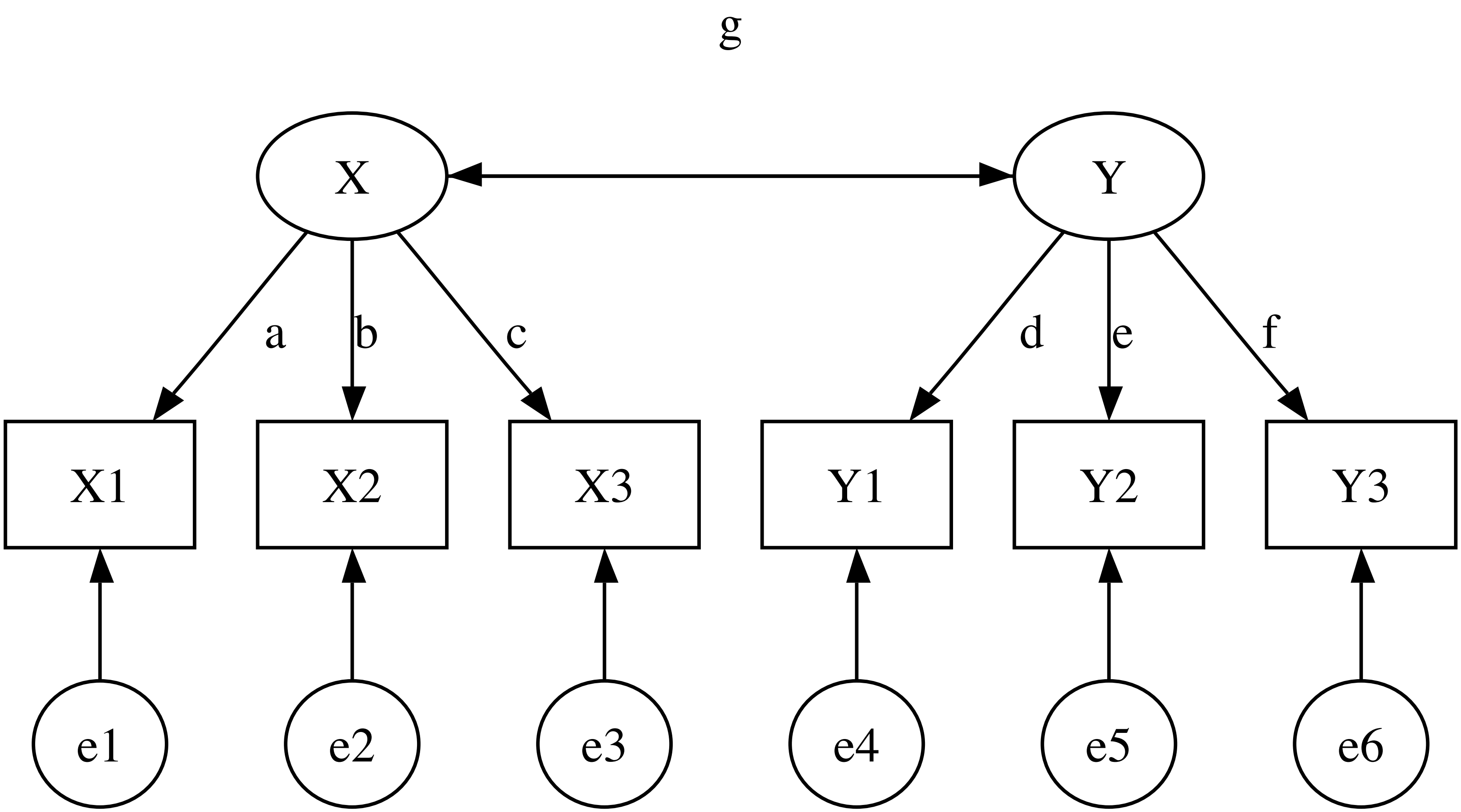
   B. .8,.7,.6

II. But sample values allow data to be fit with equal estimates

Standardized estimates

		Std. Estimate	
1	a	0.88236	V1 <--- theta1
2	b	0.78527	V2 <--- theta1
3	c	0.69405	V3 <--- theta1
4	a	0.83927	V4 <--- theta2
5	b	0.71551	V5 <--- theta2
6	c	0.63875	V6 <--- theta2

Two correlated latent traits



Pattern

	theta1	theta2	e1	e2	e3	e4	e5	e6
V1	0.9	0.0	0.44	0.0	0.00	0.0	0.00	0.0
V2	0.8	0.0	0.00	0.6	0.00	0.0	0.00	0.0
V3	0.7	0.0	0.00	0.0	0.71	0.0	0.00	0.0
V4	0.0	0.8	0.00	0.0	0.00	0.6	0.00	0.0
V5	0.0	0.7	0.00	0.0	0.00	0.0	0.71	0.0
V6	0.0	0.6	0.00	0.0	0.00	0.0	0.00	0.8

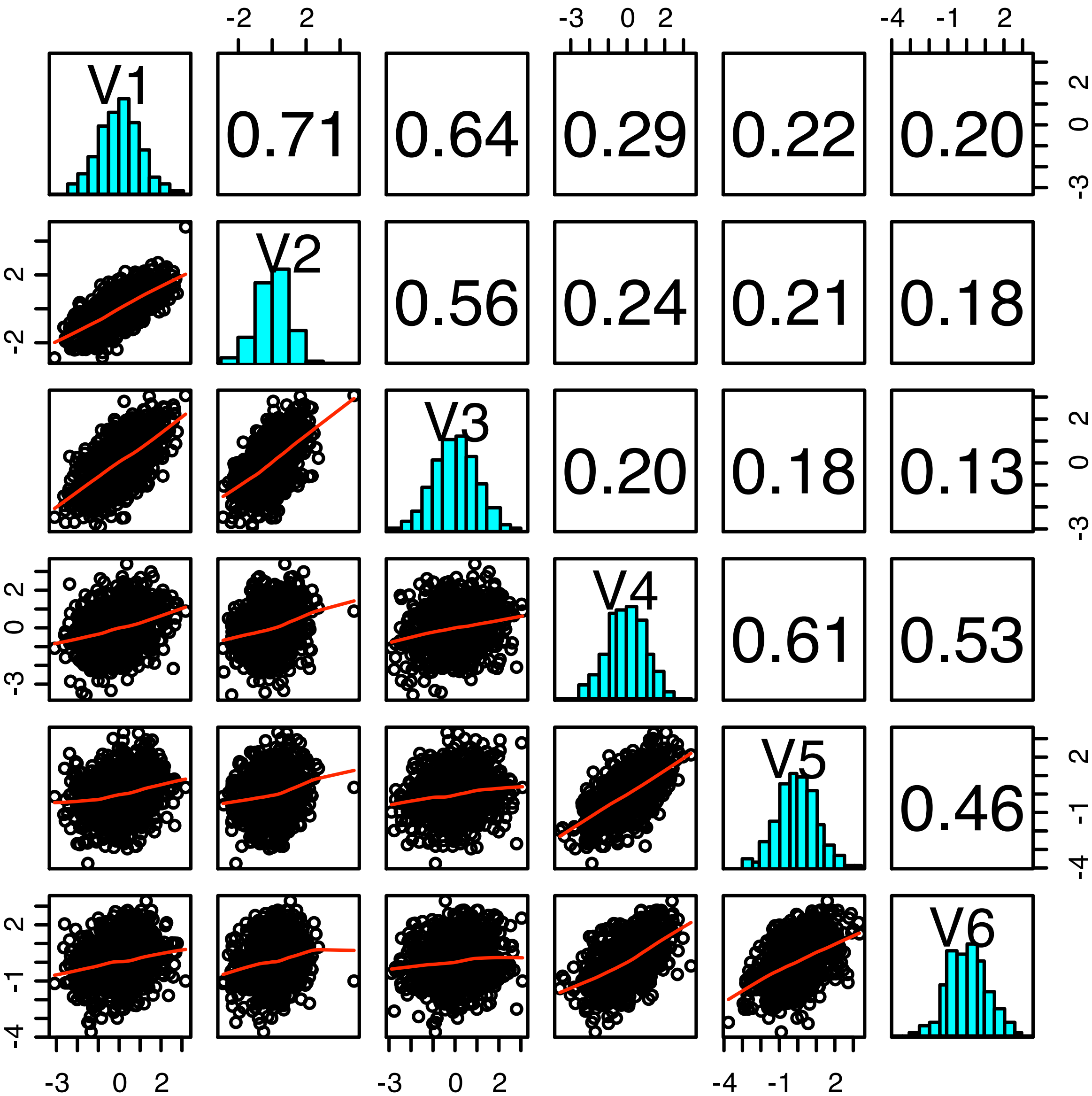
Structure

	theta1	theta2	e1	e2	e3	e4	e5	e6
V1	0.90	0.36	0.44	0.0	0.00	0.0	0.00	0.0
V2	0.80	0.32	0.00	0.6	0.00	0.0	0.00	0.0
V3	0.70	0.28	0.00	0.0	0.71	0.0	0.00	0.0
V4	0.32	0.80	0.00	0.0	0.00	0.6	0.00	0.0
V5	0.28	0.70	0.00	0.0	0.00	0.0	0.71	0.0
V6	0.24	0.60	0.00	0.0	0.00	0.0	0.00	0.8

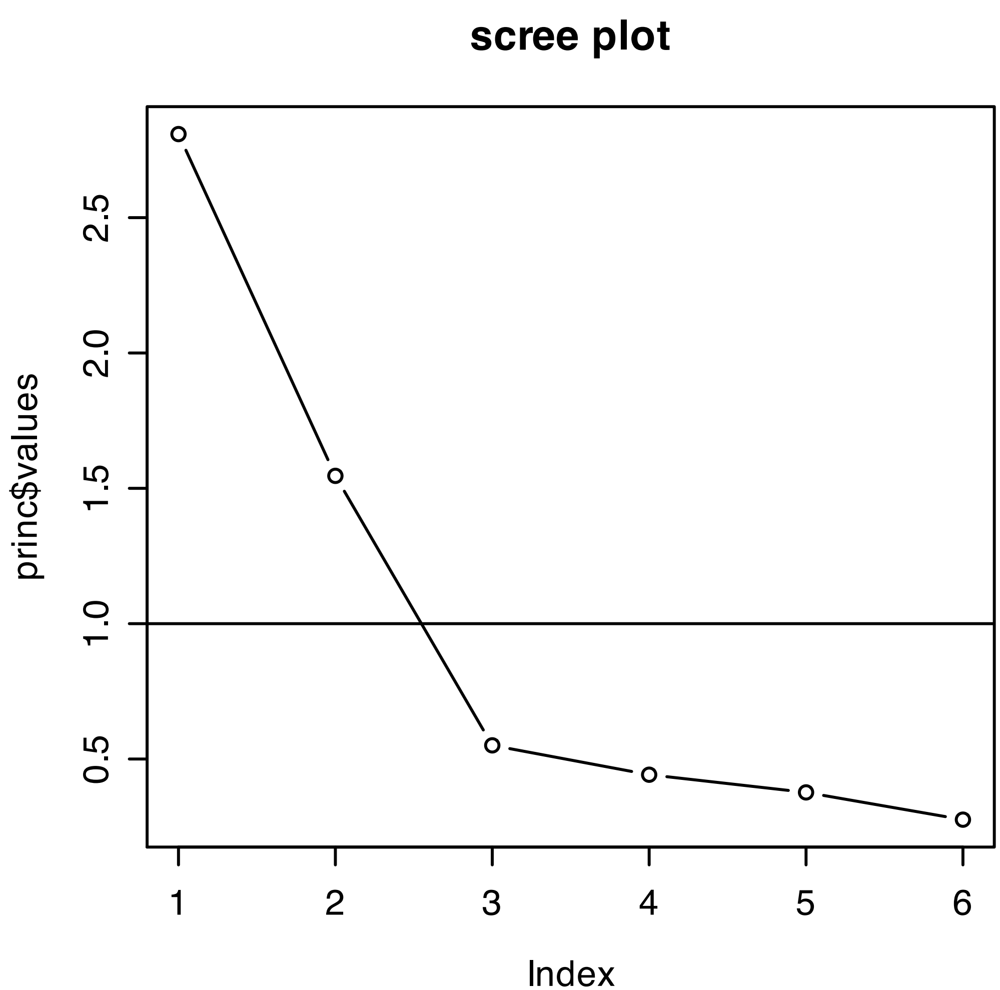
Population correlation

	V1	V2	V3	V4	V5	V6
V1	1.00	0.72	0.63	0.29	0.25	0.22
V2	0.72	1.00	0.56	0.26	0.22	0.19
V3	0.63	0.56	1.00	0.22	0.20	0.17
V4	0.29	0.26	0.22	1.00	0.56	0.48
V5	0.25	0.22	0.20	0.56	1.00	0.42
V6	0.22	0.19	0.17	0.48	0.42	1.00

Sample correlation



Scree test suggests 2



Unrotated factors

Loadings:

	Factor1	Factor2
V1	0.845	-0.291
V2	0.749	-0.256
V3	0.667	-0.254
V4	0.552	0.633
V5	0.463	0.568
V6	0.404	0.487

	Factor1	Factor2
SS loadings	2.401	1.175
Proportion Var	0.400	0.196
Cumulative Var	0.400	0.596

Varimax rotated factors

	Factor1	Factor2
V1	0.888	
V2	0.786	
V3	0.711	
V4	0.226	0.809
V5	0.174	0.711
V6	0.156	0.613

	Factor1	Factor2
SS loadings	2.018	1.558
Proportion Var	0.336	0.260
Cumulative Var	0.336	0.596

Oblimin pattern

	Factor1	Factor2
V1	0.89	0.01
V2	0.79	0.01
V3	0.72	-0.02
V4	0.01	0.84
V5	-0.01	0.74
V6	-0.01	0.64

Test 2 orthogonal

```
     path                    label initial estimate
 [1,] "theta1 -> V1"      "a"   NA
 [2,] "theta1 -> V2"      "b"   NA
 [3,] "theta1 -> V3"      "c"   NA
 [4,] "theta2 -> V4"      "d"   NA
 [5,] "theta2 -> V5"      "e"   NA
 [6,] "theta2 -> V6"      "f"   NA
 [7,] "V1 <-> V1"         "u"   NA
 [8,] "V2 <-> V2"         "v"   NA
 [9,] "V3 <-> V3"         "w"   NA
[10,] "V4 <-> V4"         "x"   NA
[11,] "V5 <-> V5"         "y"   NA
[12,] "V6 <-> V6"         "z"   NA
[13,] "theta1 <-> theta1" NA    "1"
[14,] "theta2 <-> theta2" NA    "1"
```

Statistics

Model Chisquare = 101 Df = 9 Pr(>Chisq) = 0
Chisquare (null model) = 2206 Df = 15
Goodness-of-fit index = 0.969
Adjusted goodness-of-fit index = 0.927
RMSEA index = 0.101 90% CI: (0.0838, 0.119)
Bentler-Bonnett NFI = 0.954
Tucker-Lewis NNFI = 0.93
Bentler CFI = 0.958
BIC = 38.5

Parameters

	Estimate	Std Error	z value	Pr(>|z|)	
a	0.898	0.0283	31.66	0.00e+00	V1 <--- theta1
b	0.770	0.0282	27.31	0.00e+00	V2 <--- theta1
c	0.727	0.0301	24.18	0.00e+00	V3 <--- theta1
d	0.883	0.0341	25.88	0.00e+00	V4 <--- theta2
e	0.782	0.0344	22.74	0.00e+00	V5 <--- theta2
f	0.682	0.0346	19.72	0.00e+00	V6 <--- theta2
u	0.204	0.0267	7.64	2.22e-14	V1 <--> V1
v	0.354	0.0244	14.53	0.00e+00	V2 <--> V2
w	0.510	0.0282	18.11	0.00e+00	V3 <--> V3
x	0.328	0.0401	8.18	2.22e-16	V4 <--> V4
y	0.527	0.0376	14.02	0.00e+00	V5 <--> V5
z	0.696	0.0383	18.19	0.00e+00	V6 <--> V6

Residuals suggest factors are correlated

	V1	V2	V3	V4	V5	V6
V1	0.00	0.00	0.00	0.30	0.23	0.22
V2	0.00	0.00	0.00	0.25	0.22	0.19
V3	0.00	0.00	0.00	0.22	0.20	0.15
V4	0.30	0.25	0.22	0.00	0.00	0.00
V5	0.23	0.22	0.20	0.00	0.00	0.00
V6	0.22	0.19	0.15	0.00	0.00	0.00

model 2 correlated

```
      path                  label initial estimate
 [1,] "theta1 -> V1"        "a"   NA
 [2,] "theta1 -> V2"        "b"   NA
 [3,] "theta1 -> V3"        "c"   NA
 [4,] "theta2 -> V4"        "d"   NA
 [5,] "theta2 -> V5"        "e"   NA
 [6,] "theta2 -> V6"        "f"   NA
 [7,] "theta1 <-> theta2"   "g"   NA
 [8,] "V1 <-> V1"           "u"   NA
 [9,] "V2 <-> V2"           "v"   NA
[10,] "V3 <-> V3"           "w"   NA
[11,] "V4 <-> V4"           "x"   NA
[12,] "V5 <-> V5"           "y"   NA
[13,] "V6 <-> V6"           "z"   NA
[14,] "theta1 <-> theta1"   NA    "1"
[15,] "theta2 <-> theta2"   NA    "1"
```

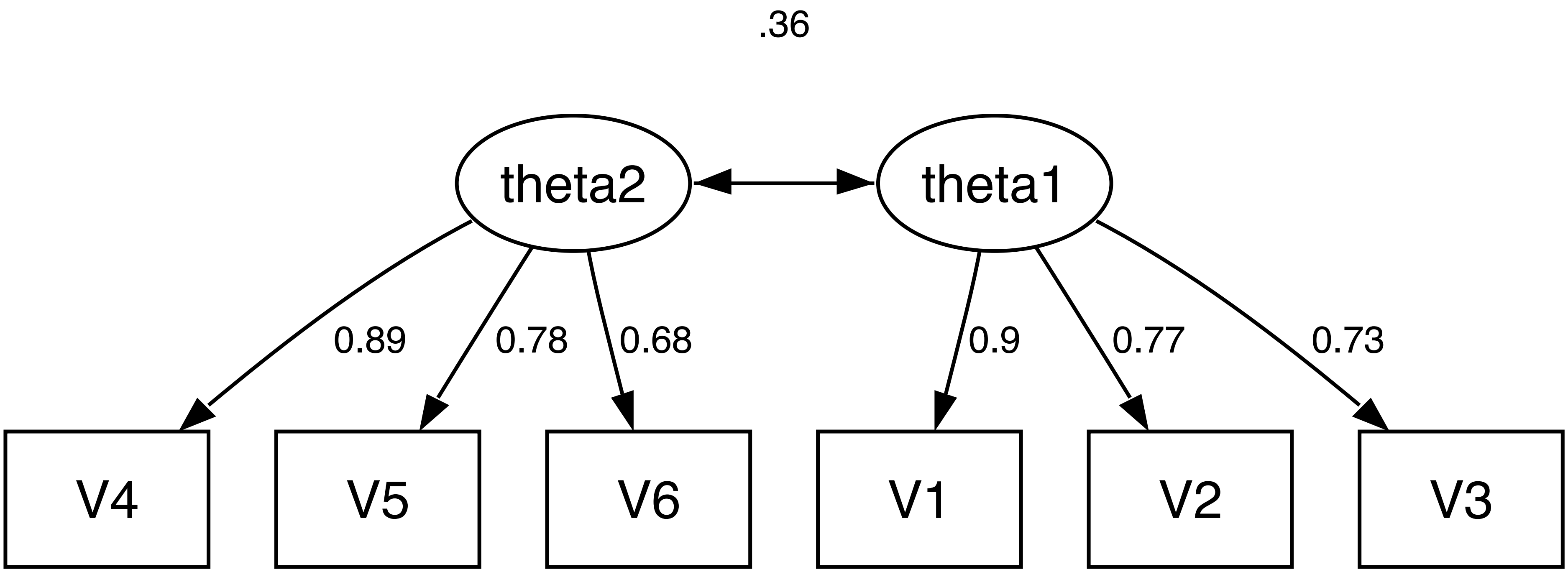
Statistics for 2 correlated

Model Chisquare = 5.39   Df = 8 Pr(>Chisq) = 0.715
Chisquare (null model) = 2206   Df = 15
Goodness-of-fit index = 0.998
Adjusted goodness-of-fit index = 0.995
RMSEA index = 0   90% CI: (NA, 0.0278)
Bentler-Bonnett NFI = 0.998
Tucker-Lewis NNFI = 1.00
Bentler CFI = 1
BIC = -49.9

Residuals

```
> round(residuals(sem.two),2)
      V1   V2    V3    V4    V5    V6
V1  0.00 0.00  0.00  0.02 -0.02  0.00
V2  0.00 0.00  0.00  0.00  0.00  0.01
V3  0.00 0.00  0.00 -0.01 -0.01 -0.03
V4  0.02 0.00 -0.01  0.00  0.00  0.00
V5 -0.02 0.00 -0.01  0.00  0.00  0.01
V6  0.00 0.01 -0.03  0.00  0.01  0.00
```

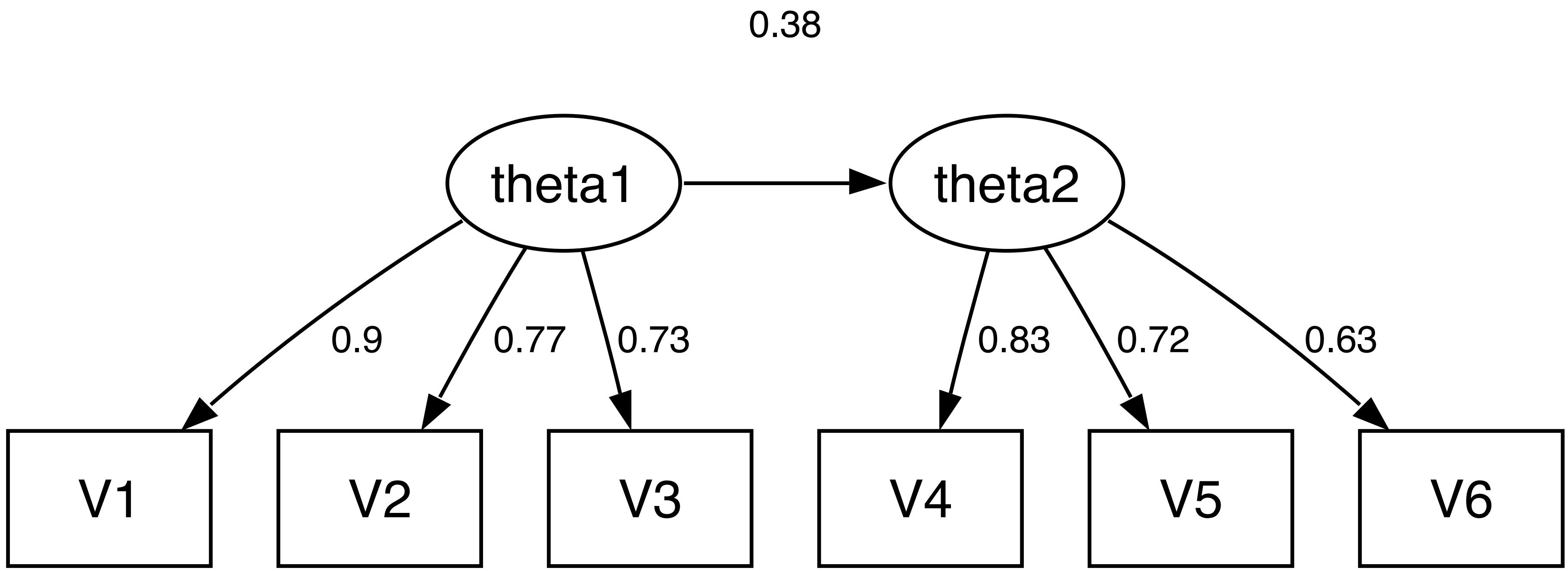
Raw coefficients



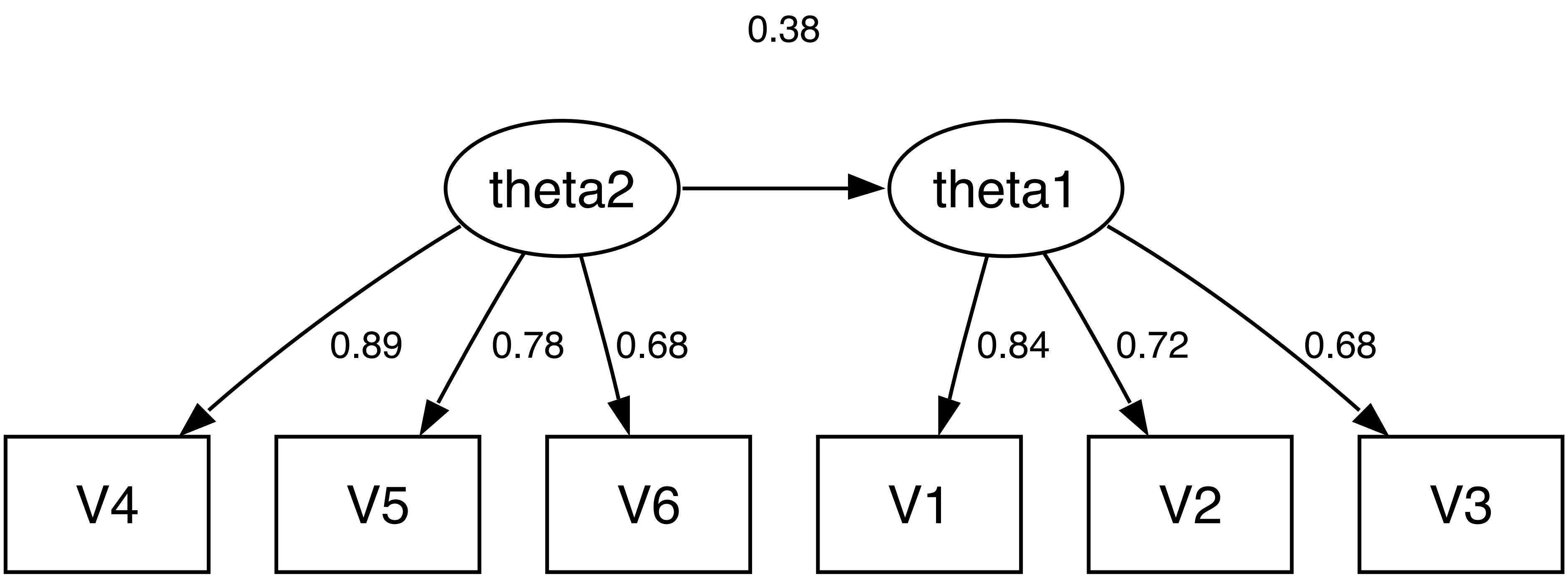
Standardized coefficients

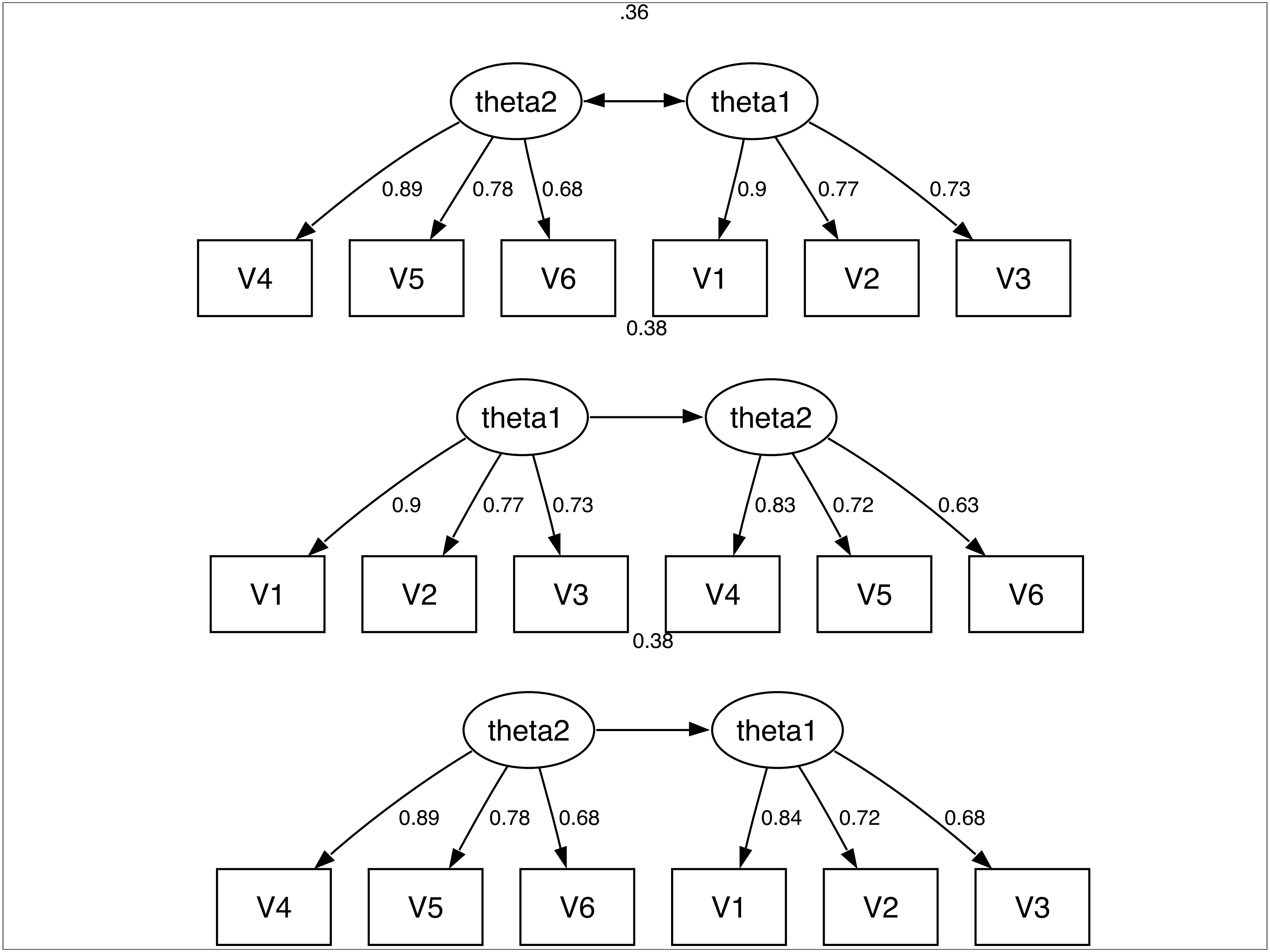
a	a	0.89472	V1 <--- theta1
b	b	0.79114	V2 <--- theta1
c	c	0.71163	V3 <--- theta1
d	d	0.84578	V4 <--- theta2
e	e	0.72766	V5 <--- theta2
f	f	0.63007	V6 <--- theta2

But what if we thought T_1 caused T_2



Or, if T_2 caused T_1




.36
theta2
theta1
0.89
0.78
0.68
0.9
0.77
0.73
V4
V5
V6
V1
V2
V3
0.38
theta1
theta2
0.9
0.77
0.73
0.83
0.72
0.63
V1
V2
V3
V4
V5
V6
0.38
theta2
theta1
0.89
0.78
0.68
0.84
0.72
0.68
V4
V5
V6
V1
V2
V3

		Std. Estimate	
a	a	0.89472	V1 <--- theta1
b	b	0.79114	V2 <--- theta1
c	c	0.71163	V3 <--- theta1
d	d	0.84578	V4 <--- theta2
e	e	0.72766	V5 <--- theta2
f	f	0.63007	V6 <--- theta2

		Std. Estimate	
a	a	0.89472	V1 <--- theta1
b	b	0.79114	V2 <--- theta1
c	c	0.71163	V3 <--- theta1
d	d	0.84578	V4 <--- theta2
e	e	0.72766	V5 <--- theta2
f	f	0.63007	V6 <--- theta2
g	g	0.35921	theta2 <--- theta1

		Std. Estimate	
a	a	0.89472	V1 <--- theta1
b	b	0.79114	V2 <--- theta1
c	c	0.71163	V3 <--- theta1
d	d	0.84578	V4 <--- theta2
e	e	0.72766	V5 <--- theta2
f	f	0.63007	V6 <--- theta2
g	g	0.35921	theta1 <--- theta2

	Estimate	Std Error	z value	Pr(>|z|)	
a	0.899	0.0280	32.10	0e+00	V1 <--- theta1
b	0.770	0.0280	27.46	0e+00	V2 <--- theta1
c	0.726	0.0300	24.20	0e+00	V3 <--- theta1
d	0.890	0.0332	26.84	0e+00	V4 <--- theta2
e	0.776	0.0338	22.98	0e+00	V5 <--- theta2
f	0.679	0.0344	19.74	0e+00	V6 <--- theta2
g	0.359	0.0337	10.65	0e+00	theta2 <--> theta1

	Estimate	Std Error	z value	Pr(>|z|)	
a	0.899	0.0280	32.10	0e+00	V1 <--- theta1
b	0.770	0.0280	27.46	0e+00	V2 <--- theta1
c	0.726	0.0300	24.20	0e+00	V3 <--- theta1
d	0.831	0.0320	25.97	0e+00	V4 <--- theta2
e	0.725	0.0321	22.58	0e+00	V5 <--- theta2
f	0.634	0.0325	19.50	0e+00	V6 <--- theta2
g	0.385	0.0415	9.28	0e+00	theta2 <--- theta1

	Estimate	Std Error	z value	Pr(>|z|)	
a	0.839	0.0272	30.81	0e+00	V1 <--- theta1
b	0.719	0.0267	26.89	0e+00	V2 <--- theta1
c	0.677	0.0285	23.75	0e+00	V3 <--- theta1
d	0.890	0.0332	26.84	0e+00	V4 <--- theta2
e	0.776	0.0338	22.98	0e+00	V5 <--- theta2
f	0.679	0.0344	19.74	0e+00	V6 <--- theta2
g	0.385	0.0415	9.28	0e+00	theta1 <--- theta2

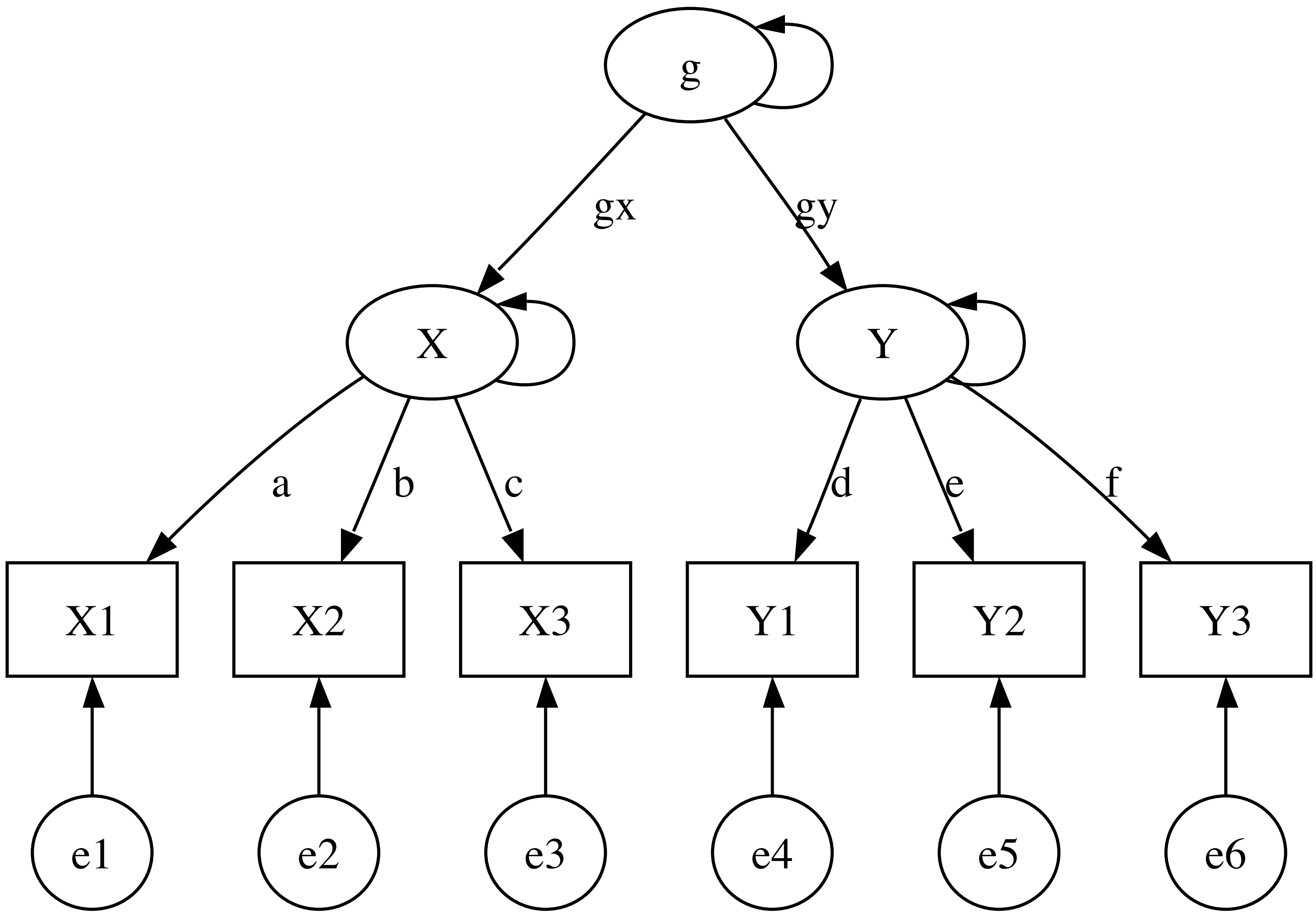
T1 <-> T2

Model Chisquare = 5.39 Df = 8 Pr(>Chisq) = 0.715
Chisquare (null model) = 2206 Df = 15
Goodness-of-fit index = 0.998
Adjusted goodness-of-fit index = 0.995
RMSEA index = 0 90% CI: (NA, 0.0278)
Bentler-Bonnett NFI = 0.998
Tucker-Lewis NNFI = 1.00
Bentler CFI = 1
BIC = -49.9

T1-> T2

Model Chisquare = 5.39 Df = 8 Pr(>Chisq) = 0.715
Chisquare (null model) = 2206 Df = 15
Goodness-of-fit index = 0.998
Adjusted goodness-of-fit index = 0.995
RMSEA index = 0 90% CI: (NA, 0.0278)
Bentler-Bonnett NFI = 0.998
Tucker-Lewis NNFI = 1.00
Bentler CFI = 1
BIC = -49.9

Hierarchical models



Model g

```
     path                 label initial estimate
 [1,] "theta1 -> V1"     "a"   NA
 [2,] "theta1 -> V2"     "b"   NA
 [3,] "theta1 -> V3"     "c"   NA
 [4,] "theta2 -> V4"     "d"   NA
 [5,] "theta2 -> V5"     "e"   NA
 [6,] "theta2 -> V6"     "f"   NA
 [7,] "g -> theta1"      NA    "1"
 [8,] "g -> theta2"      "g2"  NA
 [9,] "V1 <-> V1"        "u"   NA
[10,] "V2 <-> V2"        "v"   NA
[11,] "V3 <-> V3"        "w"   NA
[12,] "V4 <-> V4"        "x"   NA
[13,] "V5 <-> V5"        "y"   NA
[14,] "V6 <-> V6"        "z"   NA
[15,] "theta1 <-> theta1" NA   "1"
[16,] "theta2 <-> theta2" NA   "1"
[17,] "g <-> g"          NA    "1"
```

Hierarchical model

Model Chisquare = 5.39 Df = 8 Pr(>Chisq) = 0.715
Chisquare (null model) = 2206 Df = 15
Goodness-of-fit index = 0.998
Adjusted goodness-of-fit index = 0.995
RMSEA index = 0 90% CI: (NA, 0.0278)
Bentler-Bonnett NFI = 0.998
Tucker-Lewis NNFI = 1.00
Bentler CFI = 1
BIC = -49.9

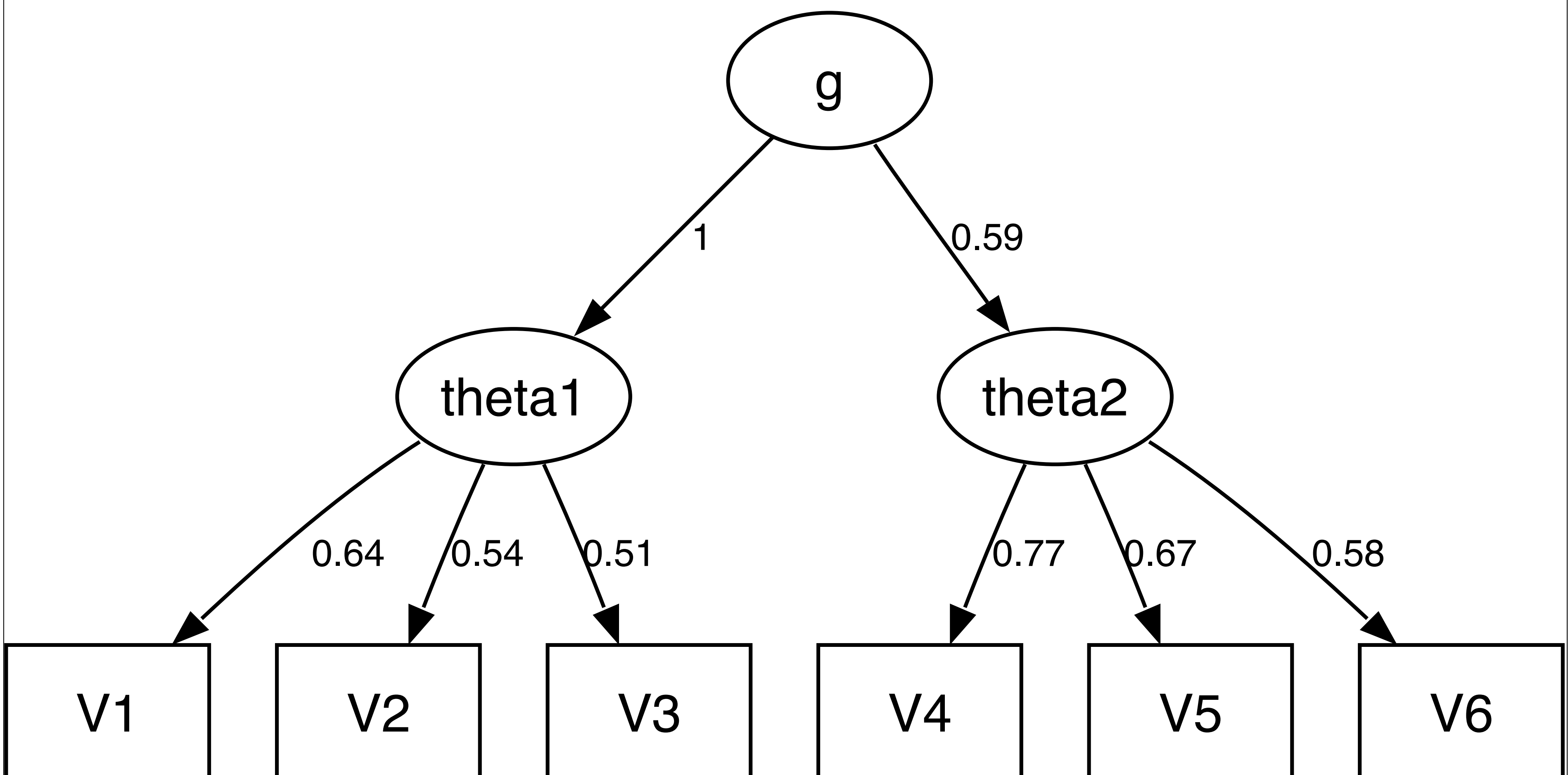
Parameters

a	0.636	0.0198	32.10	0.00e+00	V1 <--- theta1
b	0.544	0.0198	27.46	0.00e+00	V2 <--- theta1
c	0.513	0.0212	24.20	0.00e+00	V3 <--- theta1
d	0.767	0.0361	21.21	0.00e+00	V4 <--- theta2
e	0.669	0.0345	19.36	0.00e+00	V5 <--- theta2
f	0.585	0.0337	17.33	0.00e+00	V6 <--- theta2
g2	0.590	0.0747	7.90	2.89e-15	theta2 <--- g
u	0.201	0.0256	7.86	4.00e-15	V1 <--> V1
v	0.354	0.0239	14.83	0.00e+00	V2 <--> V2
w	0.513	0.0280	18.34	0.00e+00	V3 <--> V3
x	0.315	0.0378	8.34	0.00e+00	V4 <--> V4
y	0.536	0.0361	14.82	0.00e+00	V5 <--> V5
z	0.700	0.0380	18.42	0.00e+00	V6 <--> V6

standardized

		Std. Estimate	
a	a	0.89472	V1 <--- theta1
b	b	0.79114	V2 <--- theta1
c	c	0.71163	V3 <--- theta1
d	d	0.84578	V4 <--- theta2
e	e	0.72766	V5 <--- theta2
f	f	0.63007	V6 <--- theta2
		0.70711	theta1 <--- g
g2	g2	0.50800	theta2 <--- g

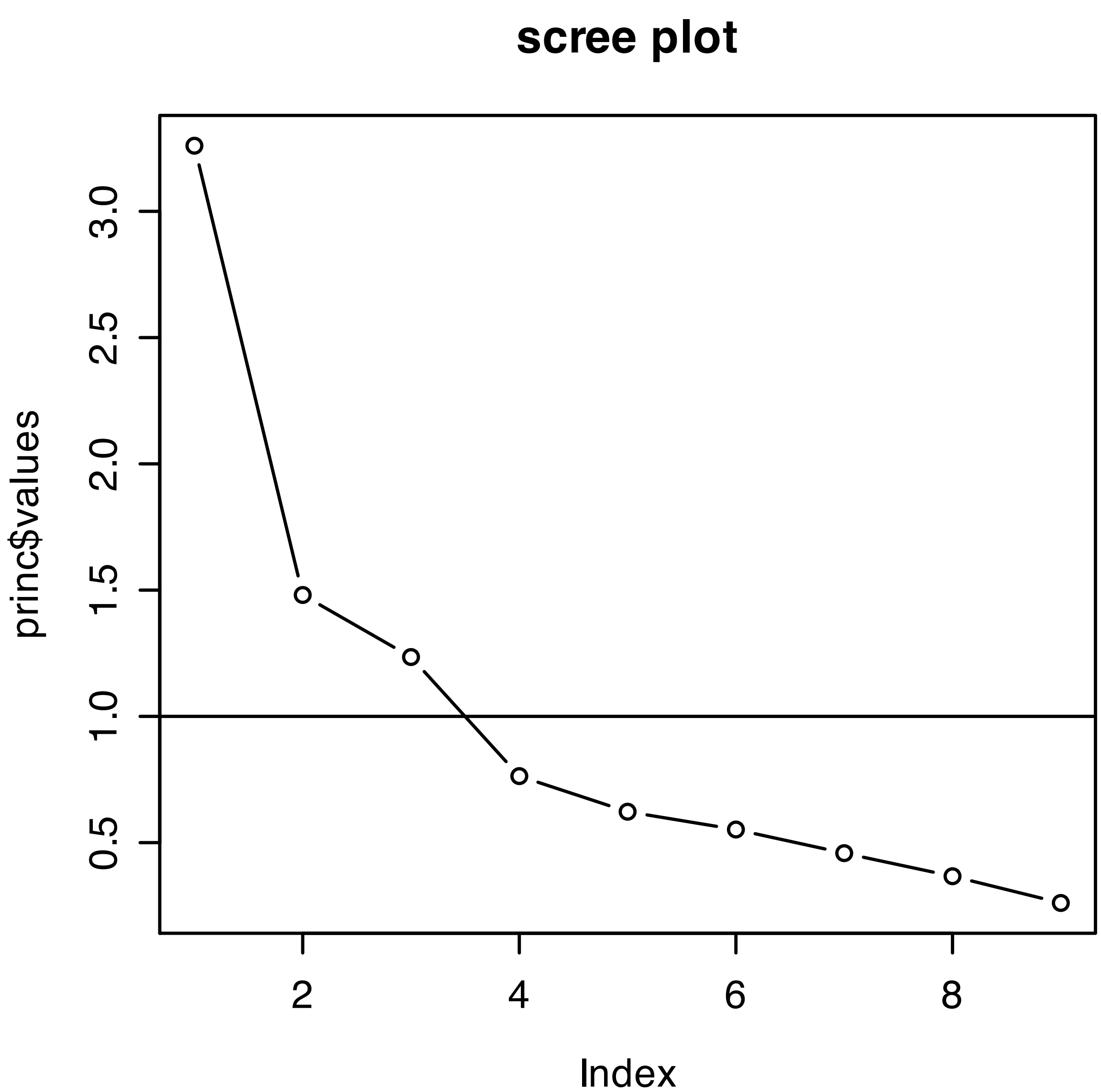
Path diagram



More complicated data

	V1	V2	V3	V4	V5	V6	V7	V8	V9
V1	1.00	0.73	0.61	0.38	0.28	0.28	0.14	0.16	0.14
V2	0.73	1.00	0.57	0.31	0.24	0.27	0.14	0.14	0.12
V3	0.61	0.57	1.00	0.27	0.21	0.23	0.12	0.13	0.11
V4	0.38	0.31	0.27	1.00	0.60	0.54	0.26	0.20	0.19
V5	0.28	0.24	0.21	0.60	1.00	0.43	0.17	0.13	0.14
V6	0.28	0.27	0.23	0.54	0.43	1.00	0.22	0.16	0.09
V7	0.14	0.14	0.12	0.26	0.17	0.22	1.00	0.37	0.32
V8	0.16	0.14	0.13	0.20	0.13	0.16	0.37	1.00	0.25
V9	0.14	0.12	0.11	0.19	0.14	0.09	0.32	0.25	1.00

Scree plot



Very Simple Structure

I. VSS as a quasi confirmatory model

   A. Fit the factor model as interpreted

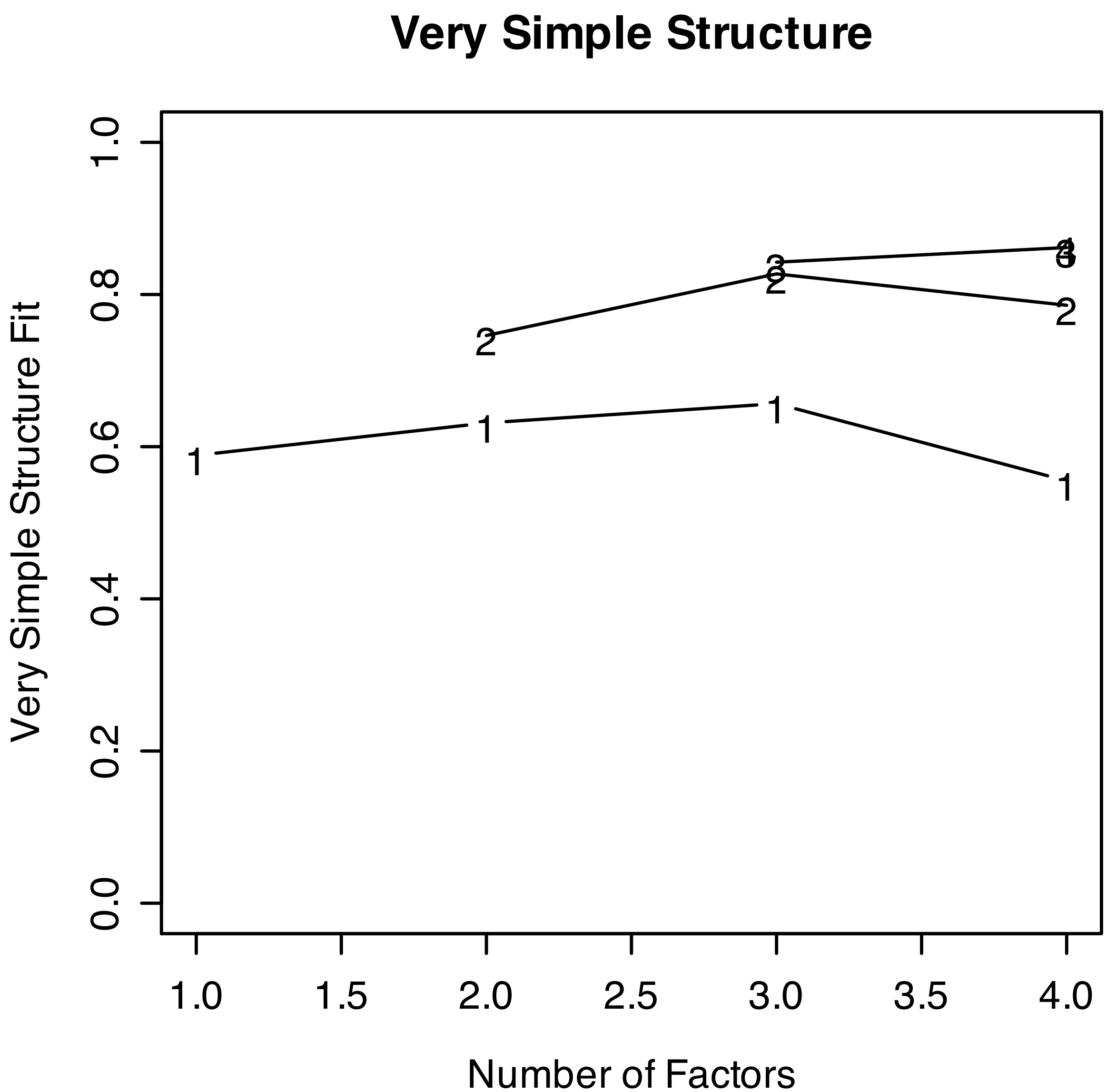
      1. keep big loadings

      2. reject small loading

   B. Try alternative number of factors

II. $R \approx F F' + U^2$ $\qquad$ $R \approx F_s F_s' + U^2$

Very Simple Structure



Extract 3 factors

Loadings:

	Factor1	Factor2	Factor3
V1	0.809	-0.358	
V2	0.737	-0.362	
V3	0.624	-0.295	
V4	0.692	0.523	-0.128
V5	0.530	0.419	-0.143
V6	0.513	0.349	
V7	0.297	0.242	0.579
V8	0.263	0.142	0.445
V9	0.230	0.137	0.381

	Factor1	Factor2	Factor3
SS loadings	2.820	1.016	0.720
Proportion Var	0.313	0.113	0.080
Cumulative Var	0.313	0.426	0.506

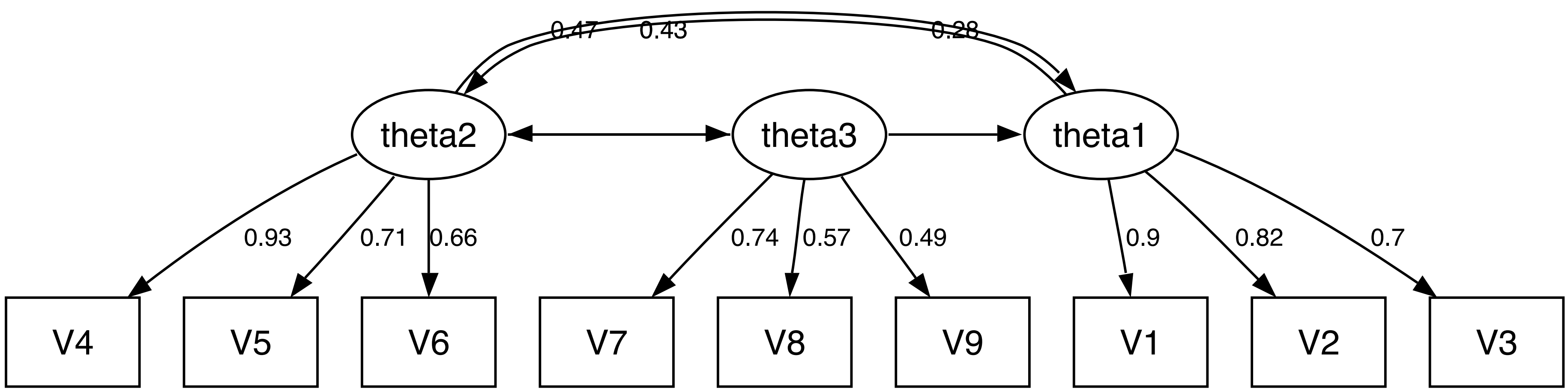
Varimax rotation

	Factor1	Factor2	Factor3
V1	0.86	0.20	0.06
V2	0.80	0.15	0.06
V3	0.67	0.14	0.06
V4	0.23	0.84	0.13
V5	0.17	0.67	0.06
V6	0.19	0.58	0.12
V7	0.08	0.19	0.66
V8	0.12	0.13	0.51
V9	0.09	0.12	0.44

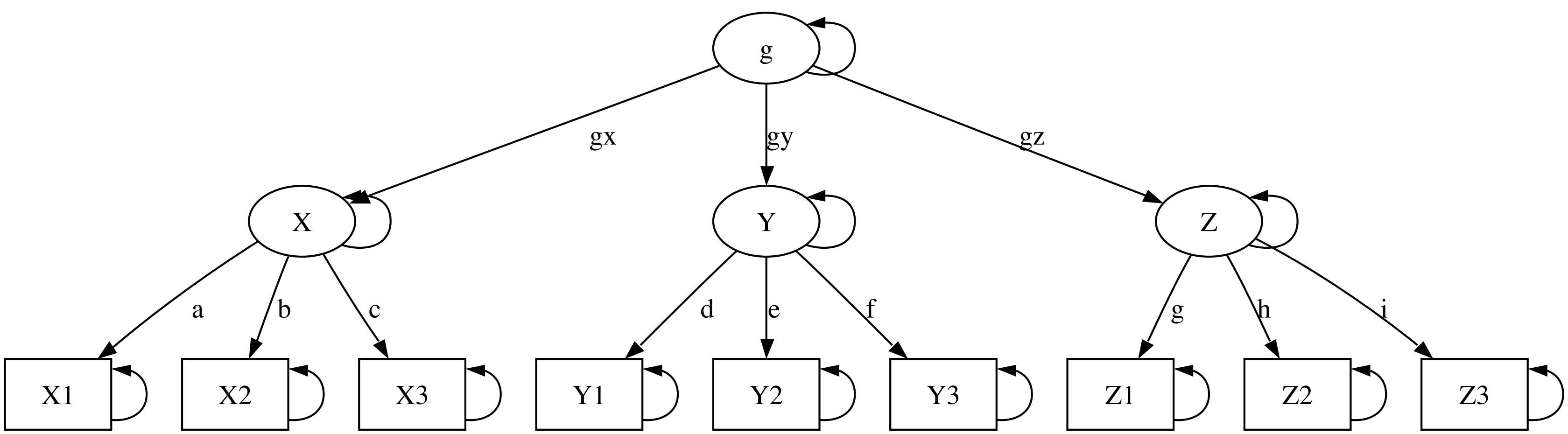
Oblique (oblimin) transformation

	Factor1	Factor2	Factor3
V1	0.88	0.02	-0.01
V2	0.83	-0.02	0.00
V3	0.69	-0.01	0.01
V4	0.00	0.87	0.01
V5	-0.02	0.71	-0.04
V6	0.03	0.59	0.03
V7	-0.03	0.01	0.70
V8	0.04	-0.02	0.53
V9	0.03	0.00	0.46

3 correlated factors



3 correlated factors accounted for by a g factor



SEM CFA : Model

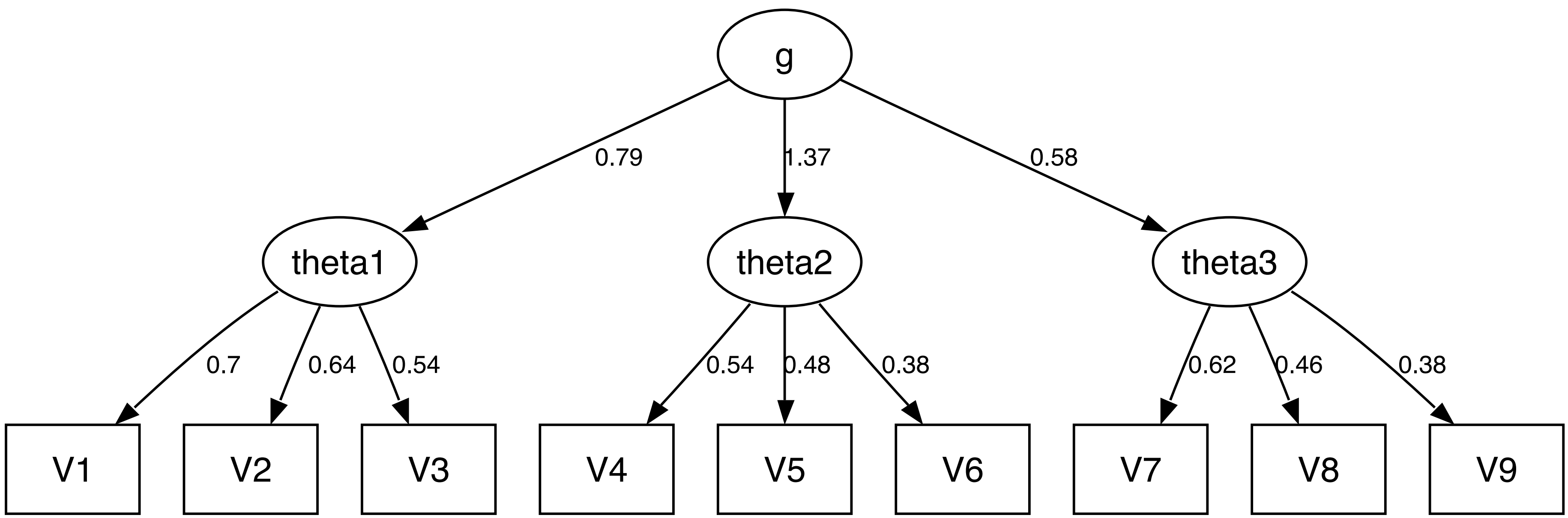
```
      path                  label initial estimate
 [1,] "theta1 -> V1"        "a"   NA
 [2,] "theta1 -> V2"        "b"   NA
 [3,] "theta1 -> V3"        "c"   NA
 [4,] "theta2 -> V4"        "d"   NA
 [5,] "theta2 -> V5"        "e"   NA
 [6,] "theta2 -> V6"        "f"   NA
 [7,] "theta3 -> V7"        "g"   NA
 [8,] "theta3 -> V8"        "h"   NA
 [9,] "theta3 -> V9"        "i"   NA
[10,] "g -> theta1"         "g1"  NA
[11,] "g -> theta2"         "g2"  NA
[12,] "g -> theta3"         "g3"  NA
[13,] "V1 <-> V1"           "u"   NA
[14,] "V2 <-> V2"           "v"   NA
[15,] "V3 <-> V3"           "w"   NA
[16,] "V4 <-> V4"           "x"   NA
[17,] "V5 <-> V5"           "y"   NA
[18,] "V6 <-> V6"           "z"   NA
[19,] "V7 <-> V7"           "s"   NA
[20,] "V8 <-> V8"           "t"   NA
[21,] "V9 <-> V9"           "r"   NA
[22,] "theta1 <-> theta1"   NA    "1"
[23,] "theta2 <-> theta2"   NA    "1"
[24,] "theta3 <-> theta3"   NA    "1"
[25,] "g <-> g"             NA    "1"
```

3 correlated latent traits

Model Chisquare = 20.5 Df = 24 Pr(>Chisq) = 0.665
Chisquare (null model) = 2647 Df = 36
Goodness-of-fit index = 0.995
Adjusted goodness-of-fit index = 0.991
RMSEA index = 0 90% CI: (NA, 0.0211)
Bentler-Bonnett NFI = 0.992
Tucker-Lewis NNFI = 1.00
Bentler CFI = 1
BIC = -145

Path coefficients

```
   Estimate Std Error z value Pr(>|z|)
a  0.699    0.0380    18.40   0.00e+00 V1 <--- theta1
b  0.642    0.0361    17.77   0.00e+00 V2 <--- theta1
c  0.542    0.0327    16.55   0.00e+00 V3 <--- theta1
d  0.543    0.0744     7.29   3.02e-13 V4 <--- theta2
e  0.482    0.0664     7.26   3.79e-13 V5 <--- theta2
f  0.379    0.0535     7.09   1.34e-12 V6 <--- theta2
g  0.618    0.0485    12.72   0.00e+00 V7 <--- theta3
h  0.461    0.0392    11.77   0.00e+00 V8 <--- theta3
i  0.377    0.0379     9.95   0.00e+00 V9 <--- theta3
g1 0.788    0.0985     8.00   1.11e-15 theta1 <--- g
g2 1.370    0.2804     4.89   1.02e-06 theta2 <--- g
g3 0.583    0.0758     7.69   1.47e-14 theta3 <--- g
u  0.204    0.0238     8.54   0.00e+00 V1 <--> V1
v  0.375    0.0251    14.92   0.00e+00 V2 <--> V2
w  0.503    0.0268    18.78   0.00e+00 V3 <--> V3
x  0.342    0.0364     9.39   0.00e+00 V4 <--> V4
y  0.524    0.0356    14.71   0.00e+00 V5 <--> V5
z  0.702    0.0365    19.21   0.00e+00 V6 <--> V6
s  0.575    0.0609     9.43   0.00e+00 V7 <--> V7
t  0.781    0.0475    16.44   0.00e+00 V8 <--> V8
r  0.925    0.0480    19.26   0.00e+00 V9 <--> V9
```

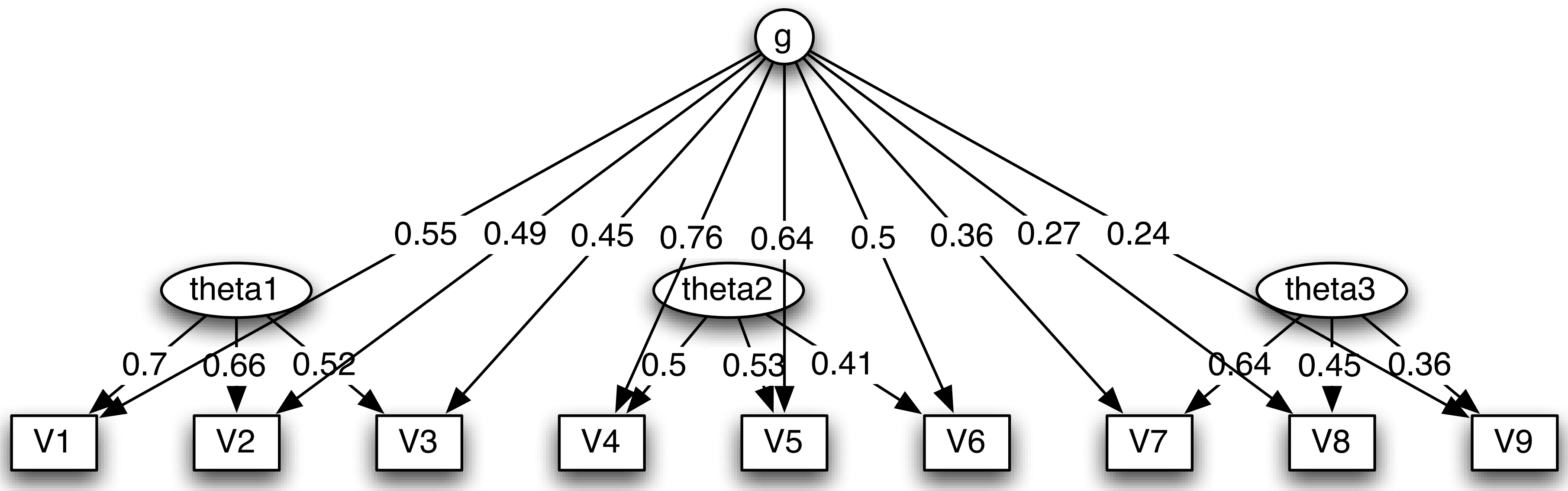


g
0.79
1.37
0.58
theta1
theta2
theta3
0.7
0.64
0.54
0.54
0.48
0.38
0.62
0.46
0.38
V1
V2
V3
V4
V5
V6
V7
V8
V9

Standardized coefficients

		Std. Estimate	
a	a	0.89200	V1 <--- theta1
b	b	0.80019	V2 <--- theta1
c	c	0.69736	V3 <--- theta1
d	d	0.84402	V4 <--- theta2
e	e	0.74887	V5 <--- theta2
f	f	0.60882	V6 <--- theta2
g	g	0.68604	V7 <--- theta3
h	h	0.51711	V8 <--- theta3
i	i	0.41359	V9 <--- theta3
g1	g1	0.61912	theta1 <--- g
g2	g2	0.80777	theta2 <--- g
g3	g3	0.50361	theta3 <--- g

Bi-factor model



From correlated to bifactor: The Schmid Leiman

I. Extract n factors

II. Rotate obliquely

III. Factor their intercorrelations

IV. Extract g from this matrix

V. Remove g from each item (following path rules)

Schmid Leiman

	g factor	Factor1	Factor2	Factor3	h2	u2
V1	0.50	0.73331	0.01326	0.0070	0.77	0.23
V2	0.44	0.69501	0.01116	0.0030	0.69	0.31
V3	0.37	0.57917	0.00523	0.0089	0.48	0.52
V4	0.74	0.00012	0.46614	0.0099	0.76	0.24
V5	0.58	0.01320	0.38105	0.0321	0.51	0.49
V6	0.54	0.02925	0.31604	0.0305	0.35	0.65
V7	0.34	0.02242	0.00607	0.6074	0.49	0.51
V8	0.27	0.03625	0.00940	0.4623	0.28	0.72
V9	0.24	0.02226	0.00044	0.3990	0.21	0.79

Summary

I. Exploratory factor analysis as a way of representing latent variables in the data

II. Confirmatory factor analysis as a way of "testing" model fit of latent variables in the data

III. Multiple representations (hierarchical, bifactor) of similar structures