Types of variables and types of models

Types of variables and relationships

I. Variables
   A. Observed
   B. Latent
II. Relationships
   A. Observed - Observed
   B. Observed - Latent
   C. Latent - Latent

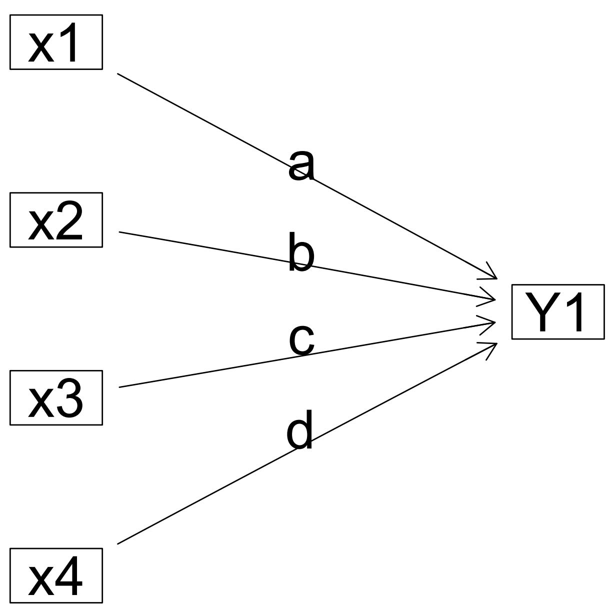
Observed-Observed

I. Kerchoff (1974), Kenny (1979) predicting attainment as a function of background variables

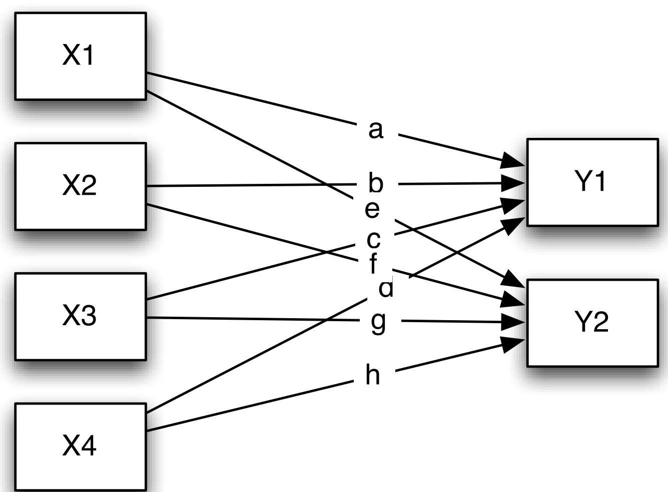
   A. (adapted from the LISREL manual, example 4.5)

Multiple regression

Regression model



Bivariate regression



Conceptual model

I. Background variables

   A. Intelligence, number of Siblings, Father’s Education, Father’s Occupation

II. Intermediate variables

   A. Grades, Educational expectation

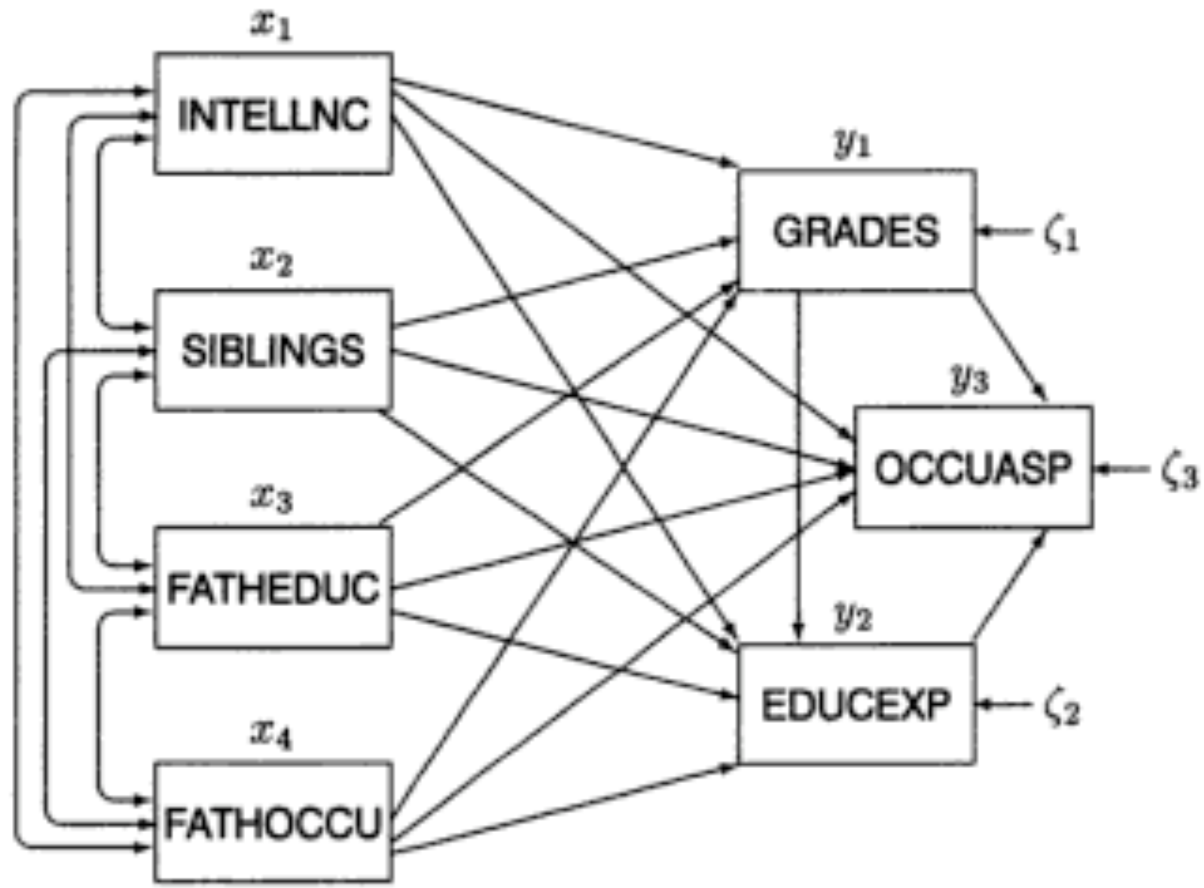
III. Final outcomes

   A. Occupational aspiration

Kerchoff/Kenny example

```
> R.kerch
             Intelligence Siblings FatherEd FatherOcc Grades EducExp OccupAsp
Intelligence        1.000   -0.100    0.277     0.250  0.572   0.489    0.335
Siblings           -0.100    1.000   -0.152    -0.108 -0.105  -0.213   -0.153
FatherEd            0.277   -0.152    1.000     0.611  0.294   0.446    0.303
FatherOcc           0.250   -0.108    0.611     1.000  0.248   0.410    0.331
Grades              0.572   -0.105    0.294     0.248  1.000   0.597    0.478
EducExp             0.489   -0.213    0.446     0.410  0.597   1.000    0.651
OccupAsp            0.335   -0.153    0.303     0.331  0.478   0.651    1.000
```

Graphical model



Matrix Regression

I. Most regression examples use raw data

A. $Y = Xß + \partial$

B. $\beta = (X'X)^{-1} X'Y$

C. lm(y~x)

II. Regression is just solving the matrix equation

A. $ß = R^{-1} r_{xy}$

B. mat.regress(R,x,y)

Simple regression: 4 predictors, 3 criteria

```
> mat.regress(R.kerch,c(1:4),c(5:7))
$beta
             Grades EducExp OccupAsp
Intelligence   0.53    0.37     0.25
Siblings      -0.03   -0.12    -0.09
FatherEd       0.12    0.22     0.10
FatherOcc      0.04    0.17     0.20

$R
  Grades  EducExp OccupAsp
    0.59     0.61     0.44

$R2
  Grades  EducExp OccupAsp
    0.35     0.38     0.19
```

More complicated regression

```
> mat.regress(R.kerch,c(1:5),c(6:7))
$beta
             EducExp OccupAsp
Intelligence    0.16     0.05
Siblings       -0.11    -0.08
FatherEd        0.17     0.05
FatherOcc       0.15     0.18
Grades          0.41     0.38

$R
 EducExp OccupAsp
    0.70     0.54

$R2
 EducExp OccupAsp
    0.48     0.29
```

Try it as a sem model with fixed x

```
> mod.kk
      path                              parameter start value
 [1,] "Intelligence -> Grades"    "a"       "NA"
 [2,] "Siblings -> Grades"        "b"       "NA"
 [3,] "FatherEd -> Grades"        "c"       "NA"
 [4,] "FatherOcc -> Grades"       "d"       "NA"
 [5,] "Intelligence -> EducExp"   "e"       "NA"
 [6,] "Siblings -> EducExp"       "f"       "NA"
 [7,] "FatherEd -> EducExp"       "g"       "NA"
 [8,] "FatherOcc -> EducExp"      "h"       "NA"
 [9,] "Intelligence -> OccupAsp" "i"       "NA"
[10,] "Siblings -> OccupAsp"      "j"       "NA"
[11,] "FatherEd -> OccupAsp"      "k"       "NA"
[12,] "FatherOcc -> OccupAsp"     "l"       "NA"
[13,] "Grades <-> Grades"         "m"       "NA"
[14,] "EducExp <-> EducExp"       "n"       "NA"
[15,] "OccupAsp <-> OccupAsp"     "o"       "NA"
```

sem of fixed x

```
sem.kk <- sem(mod.kk,R.kerch,737,fixed.x
=c('Intelligence','Siblings','FatherEd','FatherOcc'))
```

```
> summary(sem.kk)

 Model Chisquare =  411.72   Df =  3 Pr(>Chisq) = 0
 Chisquare (null model) =  1664.3   Df =  21
 Goodness-of-fit index =  0.85747
 Adjusted goodness-of-fit index =  -0.33031
 RMSEA index =  0.43024   90% CI: (0.39572, 0.46581)
 Bentler-Bonnett NFI =  0.75262
 Tucker-Lewis NNFI =  -0.74103
 Bentler CFI =  0.75128
 SRMR =  0.09976
 BIC =  391.91
```

Not modeling the DV correlations

Residuals show the effect of not modeling

```
> round(residuals(sem.kk),2)
             Intelligence Siblings FatherEd FatherOcc Grades EducExp OccupAsp
Intelligence            0        0        0         0   0.00    0.00     0.00
Siblings                0        0        0         0   0.00    0.00     0.00
FatherEd                0        0        0         0   0.00    0.00     0.00
FatherOcc               0        0        0         0   0.00    0.00     0.00
Grades                  0        0        0         0   0.00    0.26     0.25
EducExp                 0        0        0         0   0.26    0.00     0.38
OccupAsp                0        0        0         0   0.25    0.38     0.00
```

Paths match regression

```
 Parameter Estimates
  Estimate Std Error z value Pr(>|z|)
a  0.526   0.031      16.9    0.0e+00  Grades <--- Intelligence
b -0.030   0.030      -1.0    3.2e-01  Grades <--- Siblings
c  0.119   0.038       3.1    1.9e-03  Grades <--- FatherEd
d  0.041   0.038       1.1    2.8e-01  Grades <--- FatherOcc
e  0.373   0.031      12.2    0.0e+00  EducExp <--- Intelligence
f -0.124   0.030      -4.2    2.7e-05  EducExp <--- Siblings
g  0.221   0.037       5.9    3.6e-09  EducExp <--- FatherEd
h  0.168   0.037       4.6    5.3e-06  EducExp <--- FatherOcc
i  0.249   0.035       7.2    7.7e-13  OccupAsp <--- Intelligence
j -0.092   0.034      -2.7    6.3e-03  OccupAsp <--- Siblings
k  0.099   0.043       2.3    2.0e-02  OccupAsp <--- FatherEd
l  0.198   0.042       4.7    2.4e-06  OccupAsp <--- FatherOcc
m  0.651   0.034      19.2    0.0e+00  Grades <--> Grades
n  0.624   0.033      19.2    0.0e+00  EducExp <--> EducExp
o  0.807   0.042      19.2    0.0e+00  OccupAsp <--> OccupAsp
```

fixed sem = regression

```
 > round(sem.kk$coeff,3)
      a      b      c      d      e      f      g      h 
 i      j      k      l      m      n      o 
  0.526 -0.030  0.119  0.041  0.373 -0.124  0.221  0.168 
 0.249 -0.092  0.099  0.198  0.651  0.624  0.807 
```

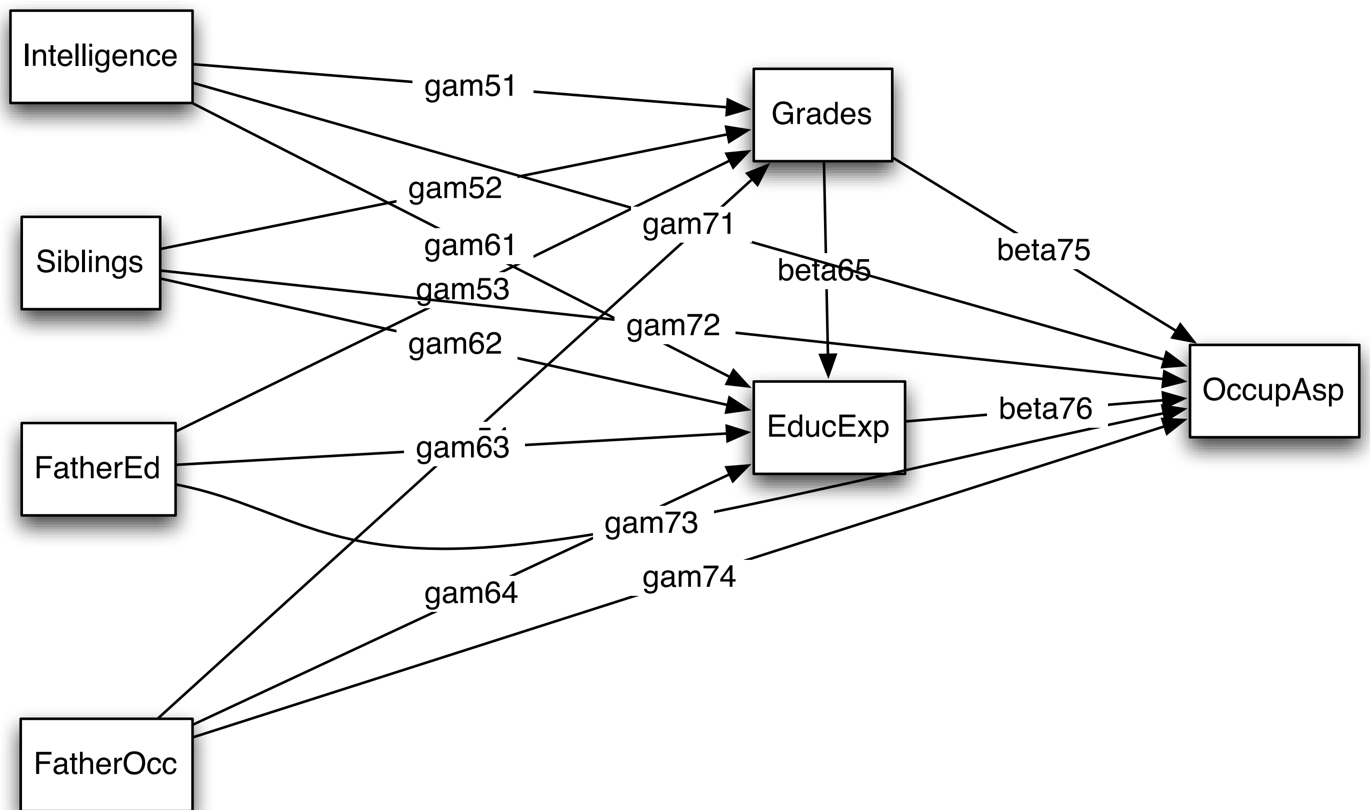
```
> mr.kk <- mat.regress(R.kerch,c(1:4),c(5:7),digits=3)
> as.vector(mr.kk$beta)
 [1]  0.526 -0.030  0.119  0.041  0.373 -0.124  0.221  0.168  0.249
-0.092  0.099  0.198
```

More complicated regression

I. Able to model the intercorrelations of the Y variables

II. Able to add some Ys to regression of other Ys

Kerchoff/Kenny example



sem model

```
> model.kerch
   Path                         Parameter
1  Intelligence -> Grades       gam51
2  Siblings -> Grades           gam52
3  FatherEd -> Grades           gam53
4  FatherOcc -> Grades          gam54
5  Intelligence -> EducExp      gam61
6  Siblings -> EducExp          gam62
7  FatherEd -> EducExp          gam63
8  FatherOcc -> EducExp         gam64
9  Grades -> EducExp            beta65
10 Intelligence -> OccupAsp     gam71
11 Siblings -> OccupAsp         gam72
12 FatherEd -> OccupAsp         gam73
13 FatherOcc -> OccupAsp        gam74
14 Grades -> OccupAsp           beta75
15 EducExp -> OccupAsp          beta76
16 Grades <-> Grades            psi5
17 EducExp <-> EducExp          psi6
18 OccupAsp <-> OccupAsp        psi7
```

```
> sem.kerch <- sem.kerch <- sem(model.kerch, R.kerch, 737,
fixed.x=c('Intelligence','Siblings','FatherEd','FatherOcc'))


> summary(sem.kerch,digits=2)

 Model Chisquare =  3.3e-13   Df =  0 Pr(>Chisq) = NA
 Chisquare (null model) =  1664   Df =  21
 Goodness-of-fit index =  1
 BIC =  3.3e-13

 Normalized Residuals
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-1.4e-15  0.0e+00  0.0e+00  4.9e-16  0.0e+00  5.2e-15
```

Parameters

```
Parameter Estimates
       Estimate Std Error z value Pr(>|z|)
gam51   0.526   0.031      16.87   0.0e+00  Grades <--- Intelligence
gam52  -0.030   0.030      -0.99   3.2e-01  Grades <--- Siblings
gam53   0.119   0.038       3.11   1.9e-03  Grades <--- FatherEd
gam54   0.041   0.038       1.07   2.8e-01  Grades <--- FatherOcc
gam61   0.160   0.033       4.90   9.6e-07  EducExp <--- Intelligence
gam62  -0.112   0.027      -4.16   3.2e-05  EducExp <--- Siblings
gam63   0.173   0.034       5.03   4.8e-07  EducExp <--- FatherEd
gam64   0.152   0.034       4.51   6.6e-06  EducExp <--- FatherOcc
beta65  0.405   0.033      12.34   0.0e+00  EducExp <--- Grades
gam71  -0.039   0.035      -1.14   2.5e-01  OccupAsp <--- Intelligence
gam72  -0.019   0.028      -0.67   5.0e-01  OccupAsp <--- Siblings
gam73  -0.041   0.036      -1.14   2.5e-01  OccupAsp <--- FatherEd
gam74   0.100   0.035       2.81   5.0e-03  OccupAsp <--- FatherOcc
beta75  0.158   0.037       4.22   2.5e-05  OccupAsp <--- Grades
beta76  0.550   0.038      14.36   0.0e+00  OccupAsp <--- EducExp
psi5    0.651   0.034      19.18   0.0e+00  Grades <--> Grades
psi6    0.517   0.027      19.18   0.0e+00  EducExp <--> EducExp
psi7    0.557   0.029      19.18   0.0e+00  OccupAsp <--> OccupAsp
```

Latent variable model

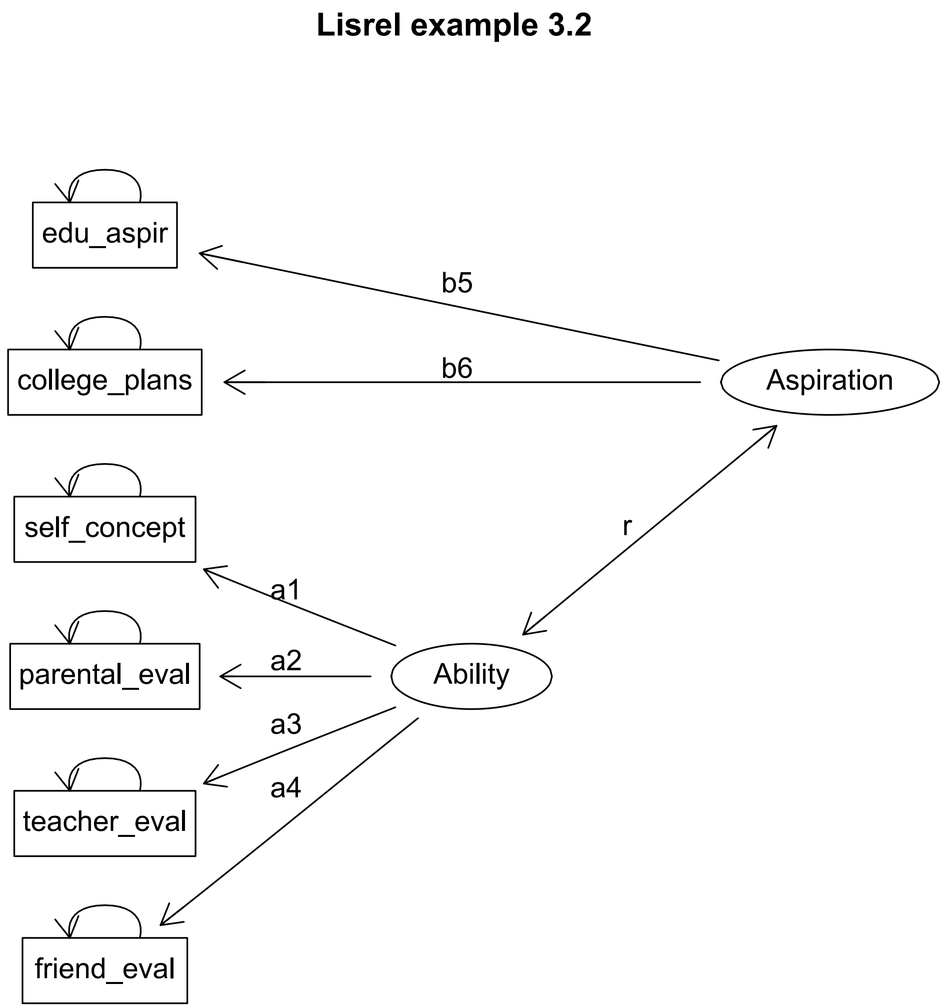
I. Latent
   A. Educational Ability and Aspirations
II. Observed
   A. evaluations of ability
   B. Educational aspirations

Educational Attainment and aspirations

I. Data from Caslyn and Kenny (1977) as cited in the LISREL User's Reference Guide

```
> ability
               self parent teacher friend edu_asp college
self_concept   1.00   0.73    0.70   0.58    0.46    0.56
parental_eval  0.73   1.00    0.68   0.61    0.43    0.52
teacher_eval   0.70   0.68    1.00   0.57    0.40    0.48
friend_eval    0.58   0.61    0.57   1.00    0.37    0.41
edu_aspir      0.46   0.43    0.40   0.37    1.00    0.72
college_plans  0.56   0.52    0.48   0.41    0.72    1.00
```

Educational attainment



Making the model

```
> fx <- structure.list(6,list(c(1:4),c(5:6)),item.labels =
rownames(ability),f.labels=c("Ability","Aspiration"))
>
> fx
              Ability Aspiration
self_concept  "a1"    "0"
parental_eval "a2"    "0"
teacher_eval  "a3"    "0"
friend_eval   "a4"    "0"
edu_aspir     "0"     "b5"
college_plans "0"     "b6"

mod.edu <- structure.graph(fx,"r",title="Lisrel example 3.2
```

sem model

```
> mod.edu
      Path                                  Parameter Value
 [1,] "Ability->self_concept"               "a1"      NA
 [2,] "Ability->parental_eval"              "a2"      NA
 [3,] "Ability->teacher_eval"               "a3"      NA
 [4,] "Ability->friend_eval"                "a4"      NA
 [5,] "Aspiration->edu_aspir"               "b5"      NA
 [6,] "Aspiration->college_plans"           "b6"      NA
 [7,] "self_concept<->self_concept"         "x1e"     NA
 [8,] "parental_eval<->parental_eval"       "x2e"     NA
 [9,] "teacher_eval<->teacher_eval"         "x3e"     NA
[10,] "friend_eval<->friend_eval"           "x4e"     NA
[11,] "edu_aspir<->edu_aspir"               "x5e"     NA
[12,] "college_plans<->college_plans"       "x6e"     NA
[13,] "Aspiration<->Ability"                "rF2F1"   NA
[14,] "Ability<->Ability"                   NA        "1"
[15,] "Aspiration<->Aspiration"             NA        "1"
```

sem results

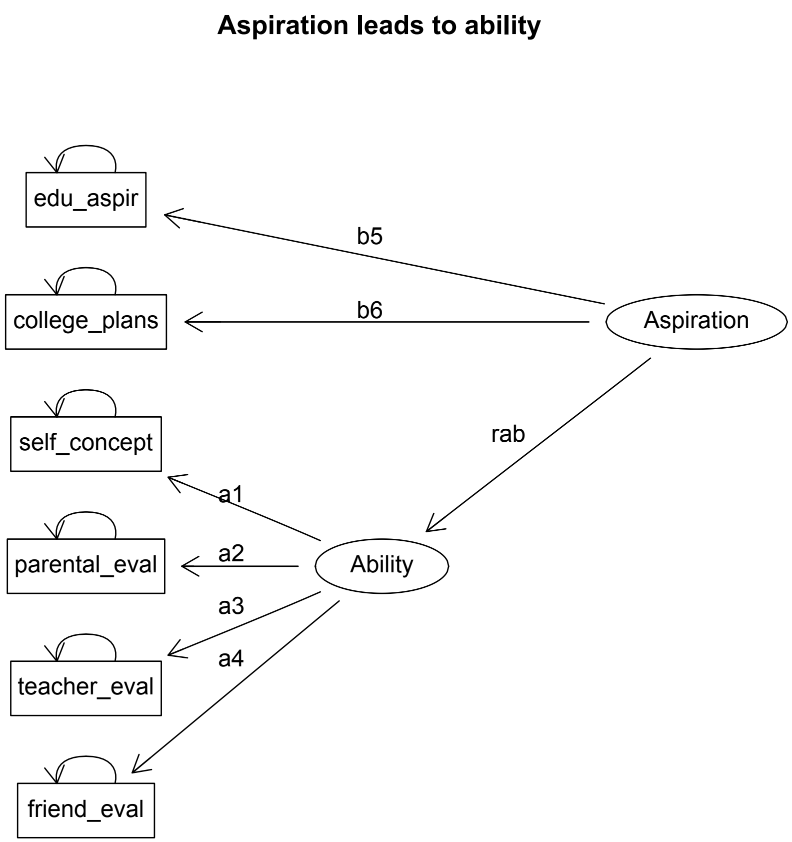
```
> colnames(ability) <-rownames(ability)
> isSymmetric(ability)
[1] TRUE
> sem.edu <- sem(mod.edu,ability,556)
> summary(sem.edu,digits=2)

 Model Chisquare =  9.3   Df =  8 Pr(>Chisq) = 0.32
 Chisquare (null model) =  1832   Df =  15
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  0.99
 RMSEA index =  0.017   90% CI: (NA, 0.054)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1
 Bentler CFI =  1
 SRMR =  0.012
 BIC =  -41
```

Parameters match the LISREL standardized

```
  Parameter Estimates
      Estimate Std Error z value Pr(>|z|)
a1    0.86     0.035     24.5    0.0000   self_concept <--- Ability
a2    0.85     0.035     23.9    0.0000   parental_eval <--- Ability
a3    0.81     0.036     22.1    0.0000   teacher_eval <--- Ability
a4    0.70     0.039     18.0    0.0000   friend_eval <--- Ability
b5    0.78     0.040     19.2    0.0000   edu_aspir <--- Aspiration
b6    0.93     0.039     23.6    0.0000   college_plans <--- Aspiration
x1e   0.25     0.024     10.8    0.0000   self_concept <--> self_concept
x2e   0.28     0.024     11.5    0.0000   parental_eval <--> parental_eval
x3e   0.35     0.027     13.1    0.0000   teacher_eval <--> teacher_eval
x4e   0.52     0.035     14.8    0.0000   friend_eval <--> friend_eval
x5e   0.40     0.038     10.4    0.0000   edu_aspir <--> edu_aspir
x6e   0.14     0.044      3.1    0.0016   college_plans <--> college_plans
rF2F1 0.67     0.031     21.5    0.0000   Ability <--> Aspiration
 Iterations =  28
>
>
```

But what if aspiration causes ability?




```
> phi <- phi.list(2,c(2))
> phi
   F1    F2
F1 "1"   "0"
F2 "rab" "1"
mod.edu <-
structure.graph(fx,phi,title
="Aspiration leads to
ability")
```

Change to causal

```
> mod.edu1 <- edit(mod.edu)
> mod.edu1
      Path                                   Parameter Value
 [1,] "Ability->self_concept"                "a1"      NA
 [2,] "Ability->parental_eval"               "a2"      NA
 [3,] "Ability->teacher_eval"                "a3"      NA
 [4,] "Ability->friend_eval"                 "a4"      NA
 [5,] "Aspiration->edu_aspir"                "b5"      NA
 [6,] "Aspiration->college_plans"            "b6"      NA
 [7,] "self_concept<->self_concept"          "x1e"     NA
 [8,] "parental_eval<->parental_eval"        "x2e"     NA
 [9,] "teacher_eval<->teacher_eval"          "x3e"     NA
[10,] "friend_eval<->friend_eval"            "x4e"     NA
[11,] "edu_aspir<->edu_aspir"                "x5e"     NA
[12,] "college_plans<->college_plans"        "x6e"     NA
[13,] "Aspiration ->Ability"                 "rF2F1"   NA
[14,] "Ability<->Ability"                    NA        "1"
[15,] "Aspiration<->Aspiration"              NA        "1"
```

Identical fits

```
> sem.edu.1 <- sem(mod.edu1,ability,556)
> summary(sem.edu.1,digits=2)

 Model Chisquare =  9.3   Df =  8 Pr(>Chisq) = 0.32
 Chisquare (null model) =  1832   Df =  15
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  0.99
 RMSEA index =  0.017   90% CI: (NA, 0.054)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1
 Bentler CFI =  1
 SRMR =  0.012
 BIC =  -41
```

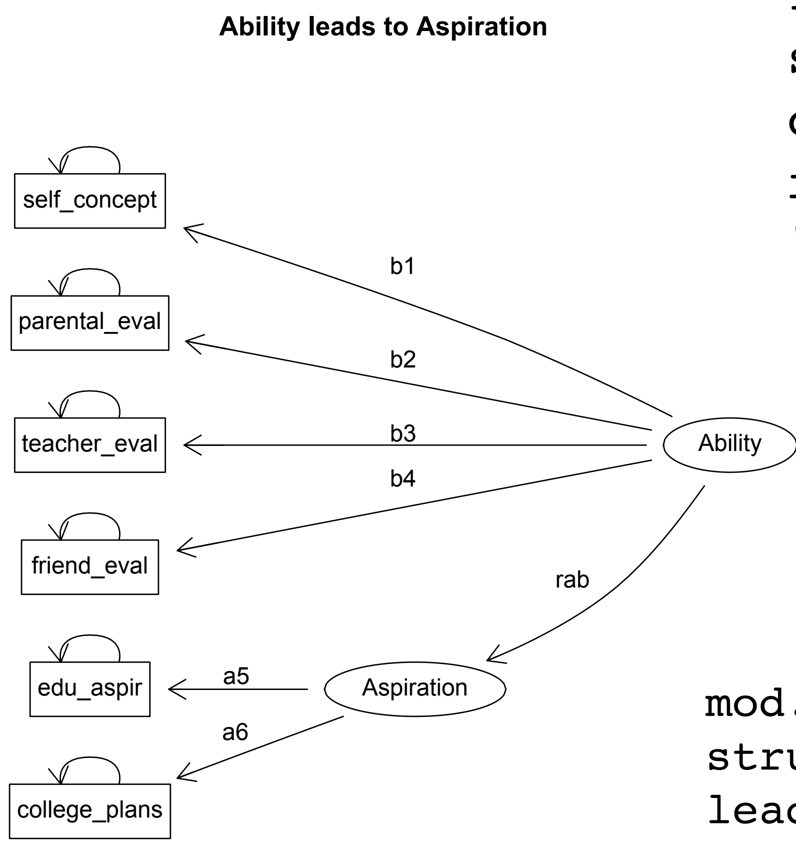
But paths are different

```
 Parameter Estimates
      Estimate Std Error z value Pr(>|z|)
a1     0.64     0.030     21.2    0.0000   self_concept <--- Ability
a2     0.63     0.031     20.5    0.0000   parental_eval <--- Ability
a3     0.60     0.031     19.3    0.0000   teacher_eval <--- Ability
a4     0.52     0.032     16.4    0.0000   friend_eval <--- Ability
b5     0.78     0.040     19.2    0.0000   edu_aspir <--- Aspiration
b6     0.93     0.039     23.6    0.0000   college_plans <--- Aspiration
x1e    0.25     0.024     10.8    0.0000   self_concept <--> self_concept
x2e    0.28     0.024     11.5    0.0000   parental_eval <--> parental_eval
x3e    0.35     0.027     13.1    0.0000   teacher_eval <--> teacher_eval
x4e    0.52     0.035     14.8    0.0000   friend_eval <--> friend_eval
x5e    0.40     0.038     10.4    0.0000   edu_aspir <--> edu_aspir
x6e    0.14     0.044      3.1    0.0016   college_plans <--> college_plans
rF2F1 0.89      0.075     12.0    0.0000   Ability <--- Aspiration
```

Reverse cause

```
> mod.edu2 <- edit(mod.edu)
> mod.edu2
      Path                                Parameter Value
 [1,] "Ability->self_concept"            "a1"      NA
 [2,] "Ability->parental_eval"           "a2"      NA
 [3,] "Ability->teacher_eval"            "a3"      NA
 [4,] "Ability->friend_eval"             "a4"      NA
 [5,] "Aspiration->edu_aspir"            "b5"      NA
 [6,] "Aspiration->college_plans"        "b6"      NA
 [7,] "self_concept<->self_concept"      "x1e"     NA
 [8,] "parental_eval<->parental_eval"    "x2e"     NA
 [9,] "teacher_eval<->teacher_eval"      "x3e"     NA
[10,] "friend_eval<->friend_eval"        "x4e"     NA
[11,] "edu_aspir<->edu_aspir"            "x5e"     NA
[12,] "college_plans<->college_plans"    "x6e"     NA
[13,] "Aspiration<-Ability"              "rF2F1"   NA
[14,] "Ability<->Ability"                NA        "1"
[15,] "Aspiration<->Aspiration"          NA        "1"
```

Ability is causal




```
fx <-
structure.list(6,list(c(5,6),
c(1:4)),item.labels =
rownames(ability),f.labels=c(
"Aspiration","Ability"))
```

```
mod.edu <-
structure.graph(fx,phi,title="Ability
leads to Aspiration")
```

Fits are the same (again)

```
> sem.mod.edu2 <- sem(mod.edu2,ability,556)
> summary(sem.mod.edu2,digits=2)

 Model Chisquare =  9.3   Df =  8 Pr(>Chisq) = 0.32
 Chisquare (null model) =  1832   Df =  15
 Goodness-of-fit index =  1
 Adjusted goodness-of-fit index =  0.99
 RMSEA index =  0.017   90% CI: (NA, 0.054)
 Bentler-Bonnett NFI =  1
 Tucker-Lewis NNFI =  1
 Bentler CFI =  1
 SRMR =  0.012
 BIC =  -41
```

Compare paths

```
> edu <- data.frame(correlated=sem.edu$coeff,asp=sem.edu.
1$coeff,abil=sem.mod.edu2$coeff)
> round(edu,2)
      correlated  asp abil
a1          0.86 0.64 0.86
a2          0.85 0.63 0.85
a3          0.81 0.60 0.81
a4          0.70 0.52 0.70
b5          0.78 0.78 0.58
b6          0.93 0.93 0.69
x1e         0.25 0.25 0.25
x2e         0.28 0.28 0.28
x3e         0.35 0.35 0.35
x4e         0.52 0.52 0.52
x5e         0.40 0.40 0.40
x6e         0.14 0.14 0.14
rF2F1       0.67 0.89 0.89
```

Paths differ as a function of presumed direction of influence

Implications of arrows

I. Need to fit alternative models

II. Need to consider alternative representations

III. Are there external variables that allow one to choose between models?

IV. Confirmation that a model fits does not confirm theoretical adequacy.

Types of variables

I. Observed variables can be ‘reflective’ of the latent variable. They are ‘effect indicators’.

II. Observed variables can be ‘causal indicators’ or ‘formative indicators’ that directly effect the latent variable

Formative indicators (Bollen, 2002)

I. Time spent with friends, time spent with family, time spent with coworkers as indicators of time spent in social interaction.

II. Formative indicators correlational structure is independent of loadings on a factor. They are not locally independent

Effect (reflective) indicators

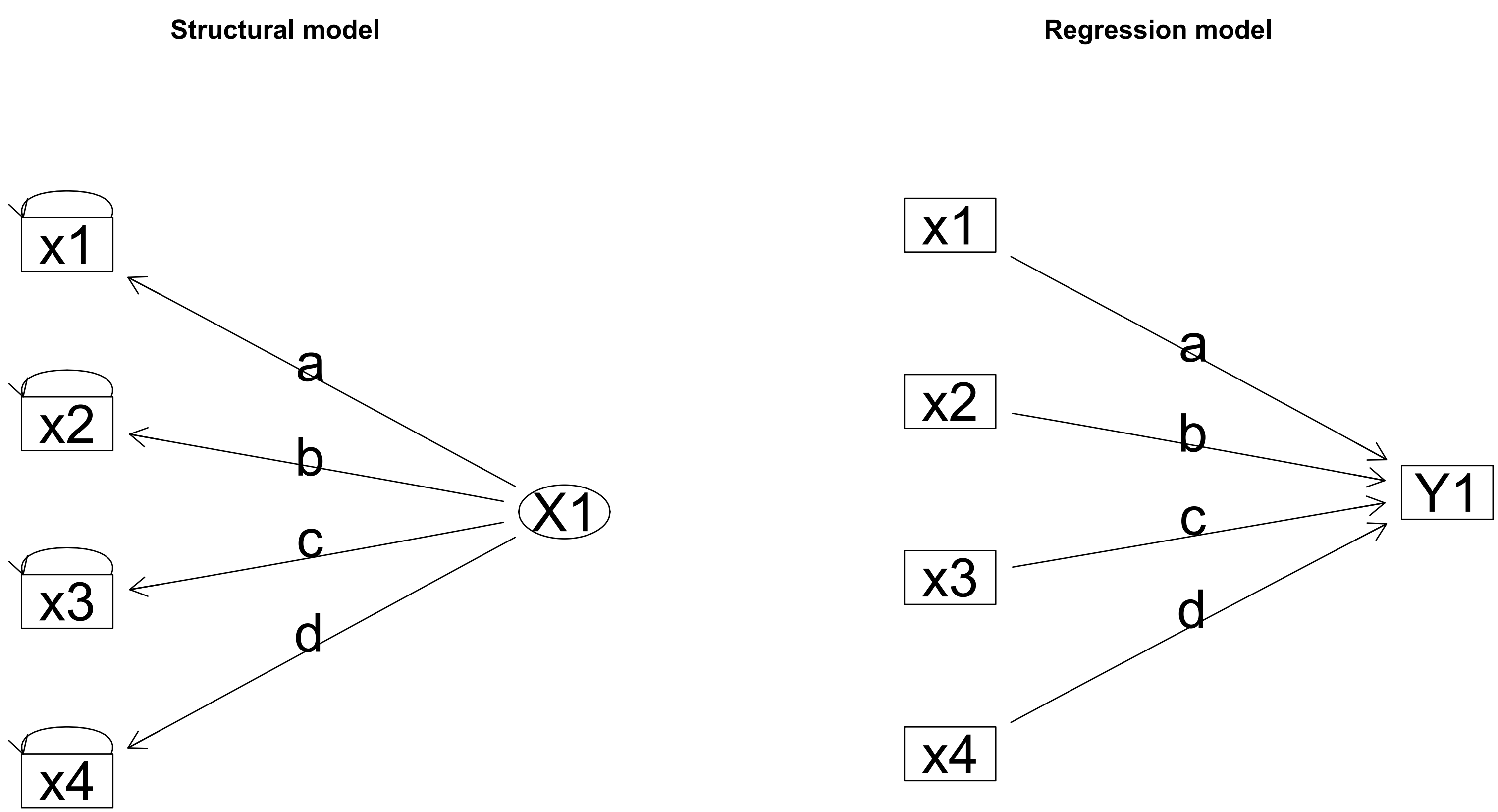
I. test scores on various quantitative tests as effect indicators of ability

II. feelings of self worth as effect indicators of self esteem.

III. Correlational structure is a function of path coefficients with latent variable

IV. Values are locally independent (uncorrelated when latent is partialled out).

Type of indicator and direction of the arrows



between X correlations not shown

McArdle (2009)

I. McArdle, J. J. Latent variable modeling of differences and changes with longitudinal data. Annual Review of Psychology, 60, 577-605.

II. http://arjournals.annualreviews.org/doi/pdf/10.1146/annurev.psych.60.110707.163612

Change models

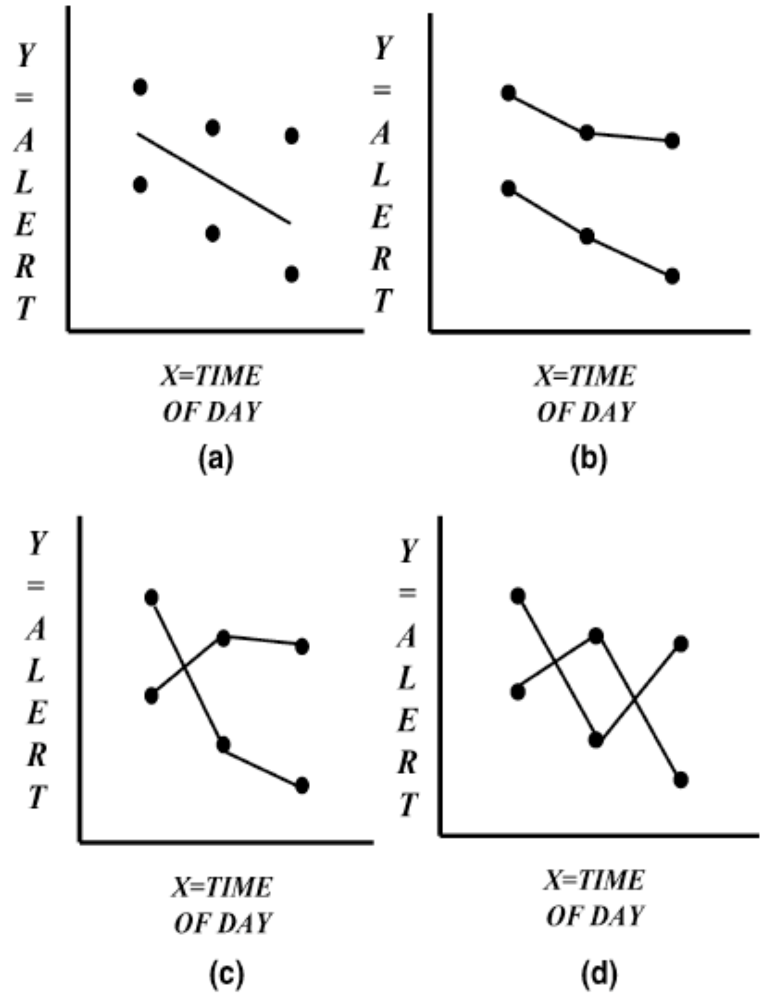
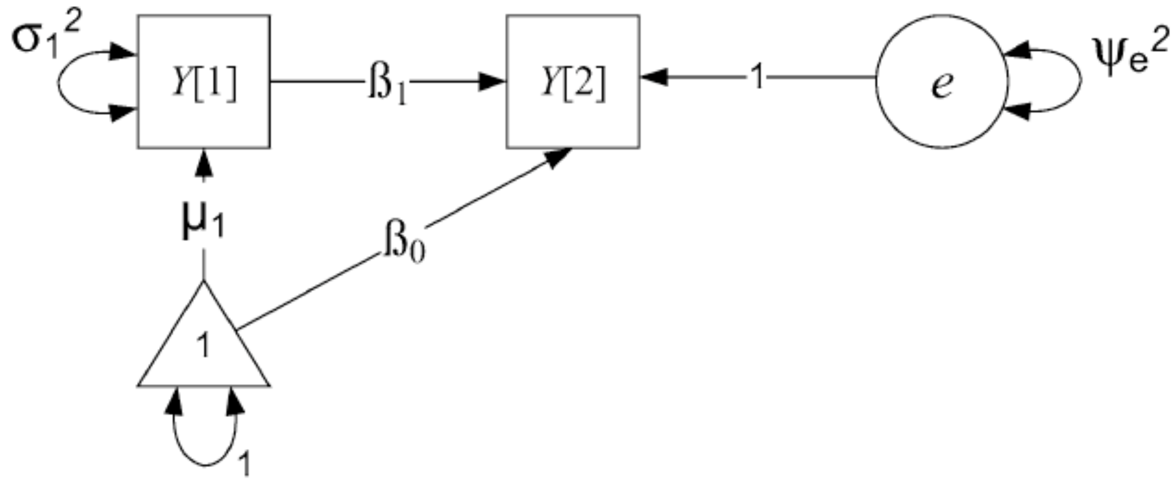


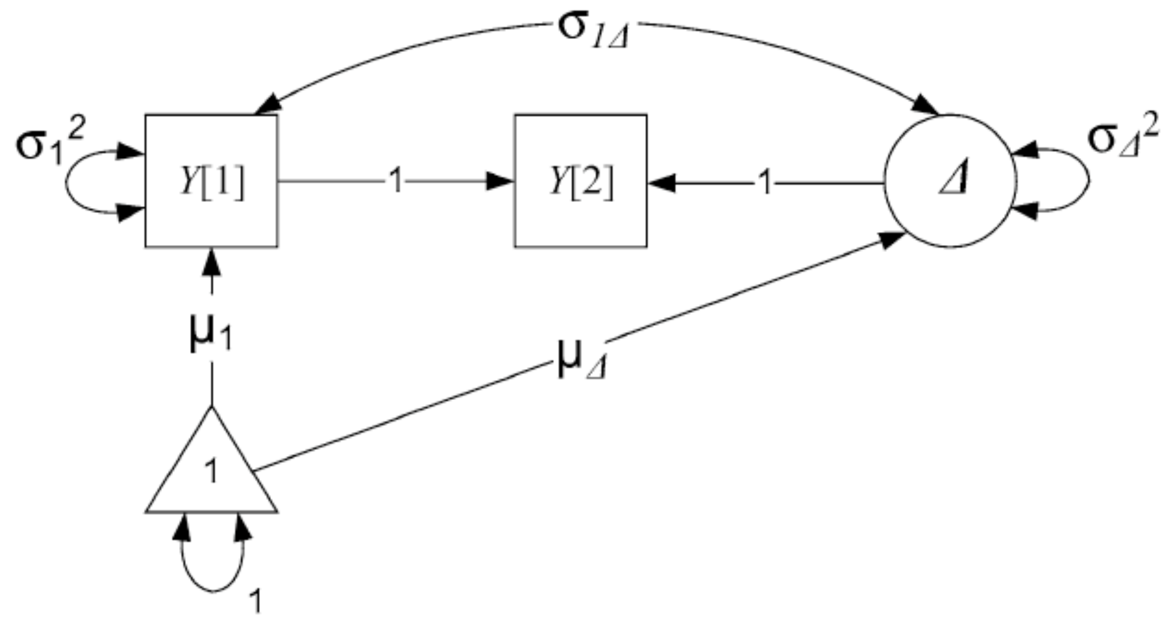
Figure 1

Alternative plots of cross-sectional and longitudinal data. (*a*) Cross-sectional measurements, (*b*) longitudinal measurements, (*c*) one longitudinal alternative, and (*d*) another longitudinal alternative.

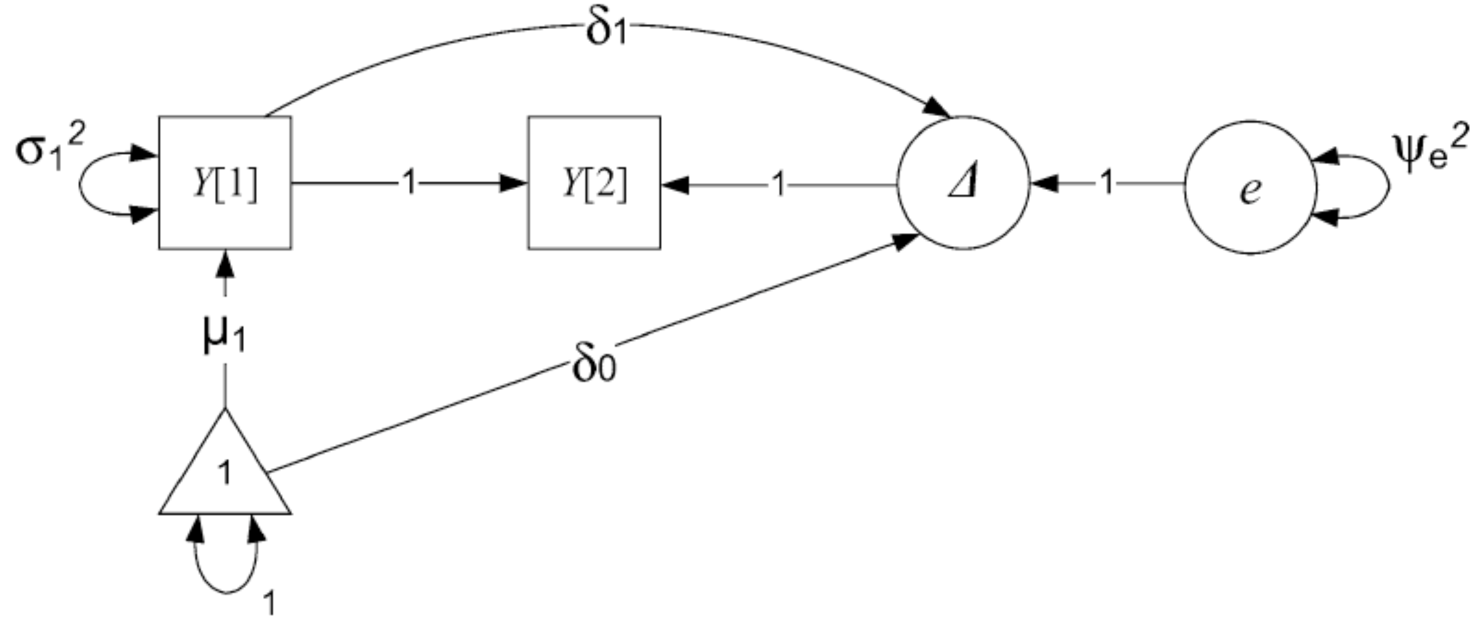
Traditional regression



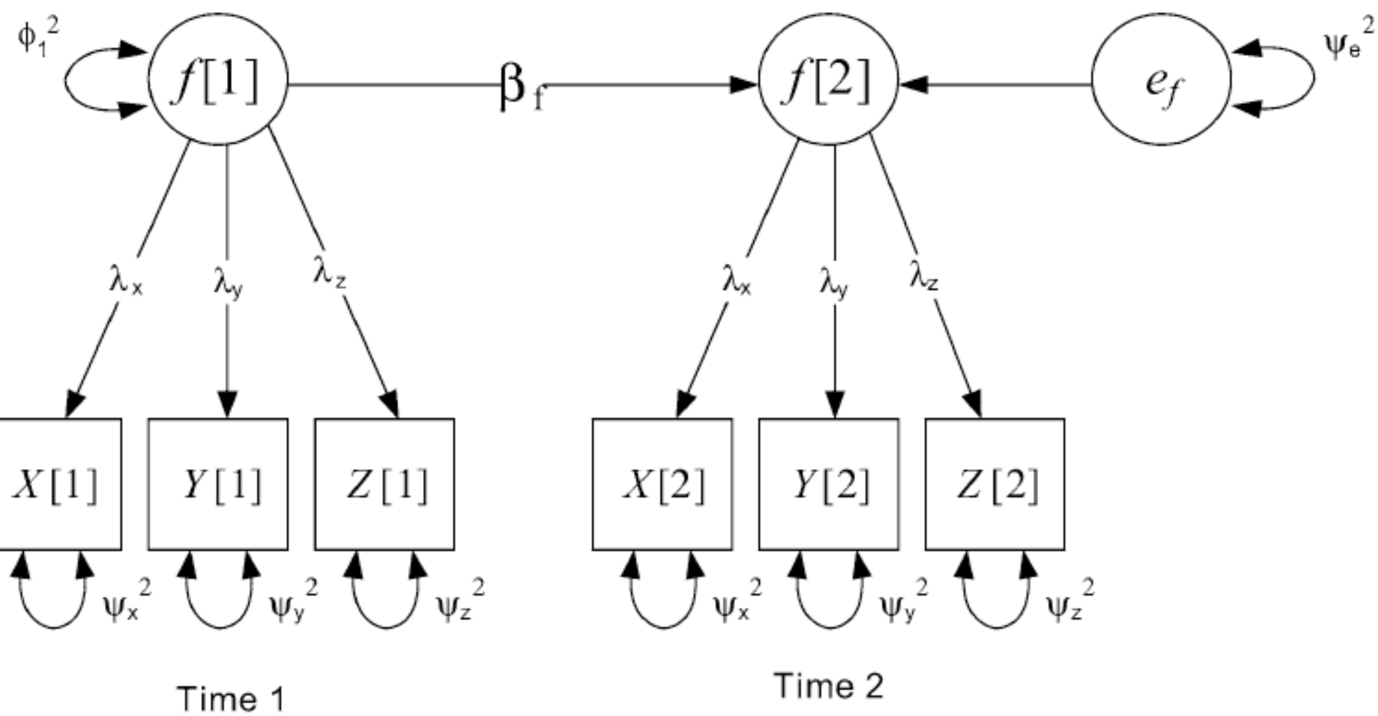
Latent Change scores



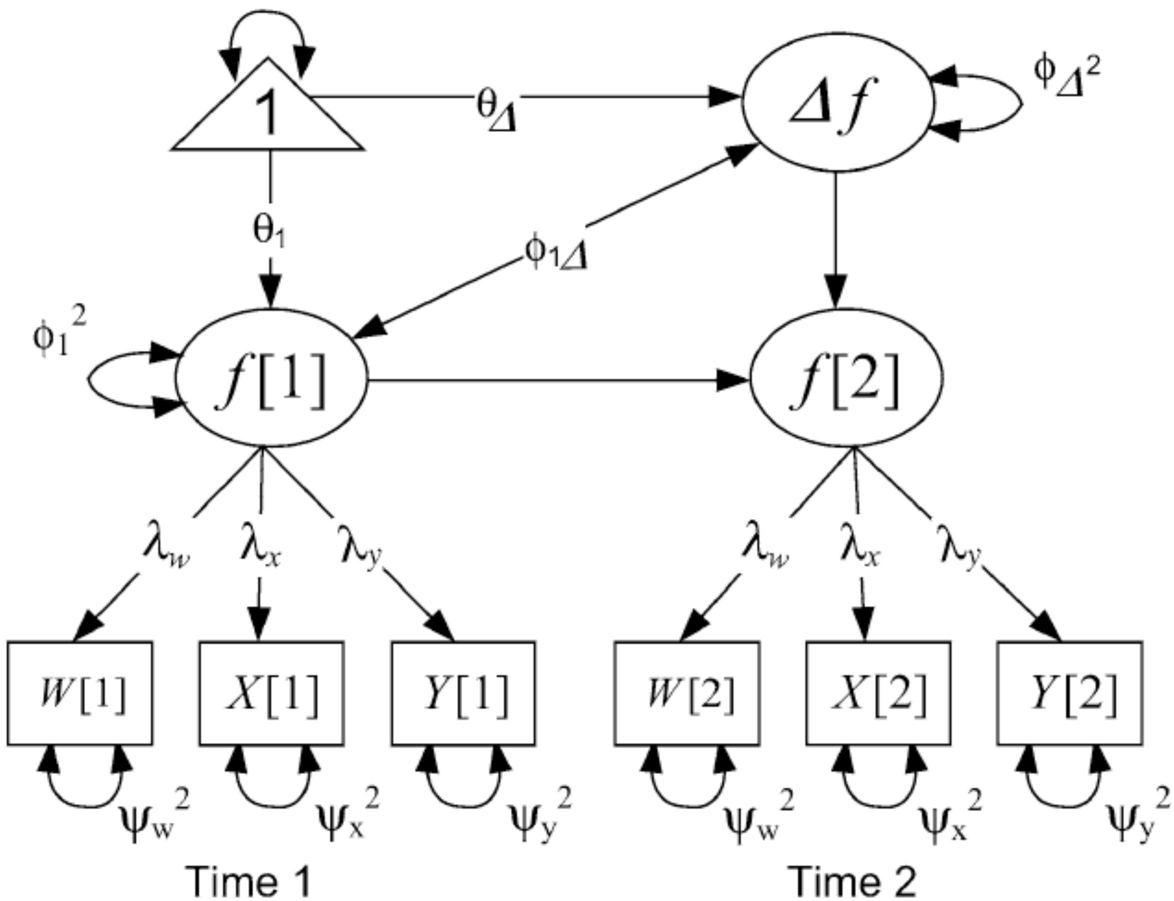
Change Regression



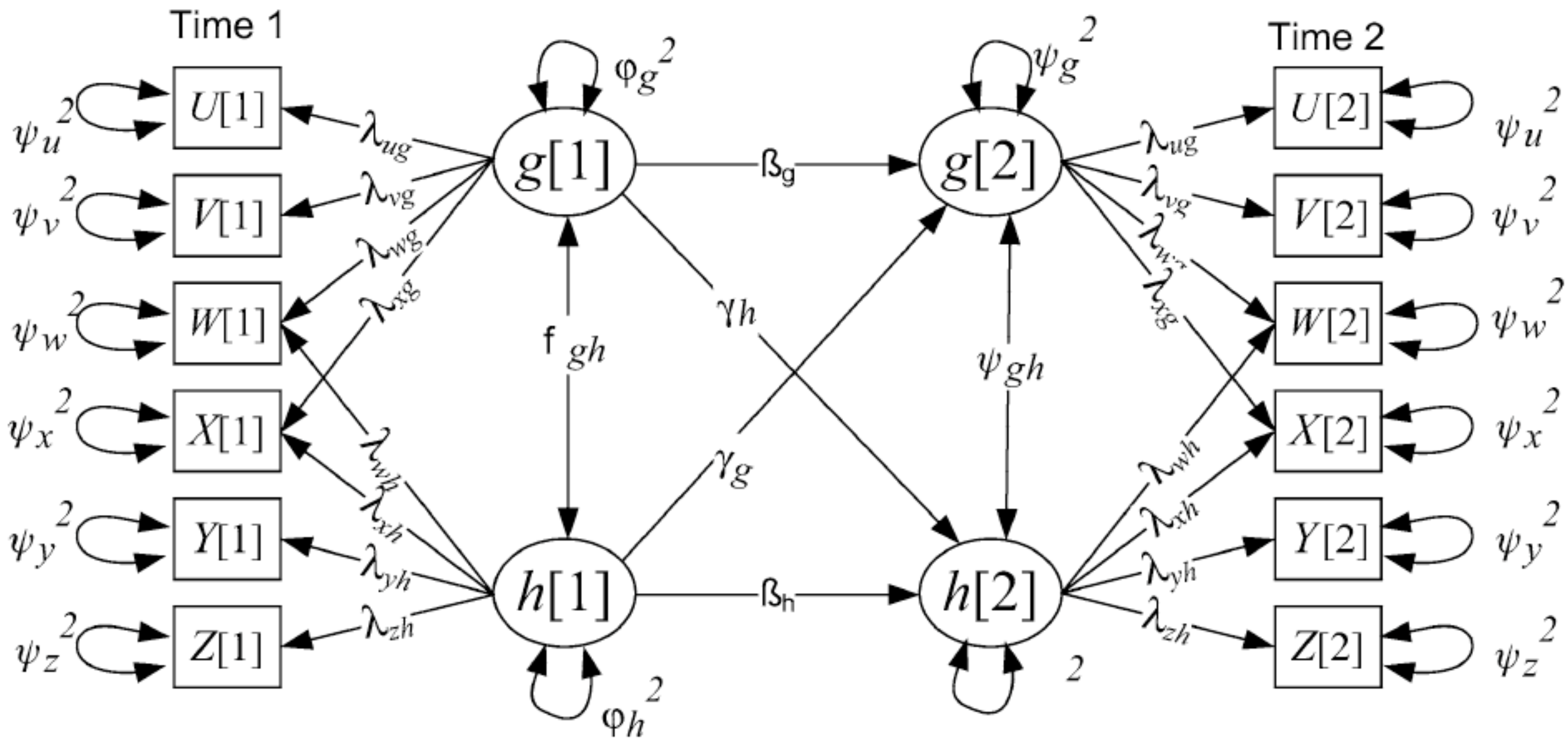
Common factor regression



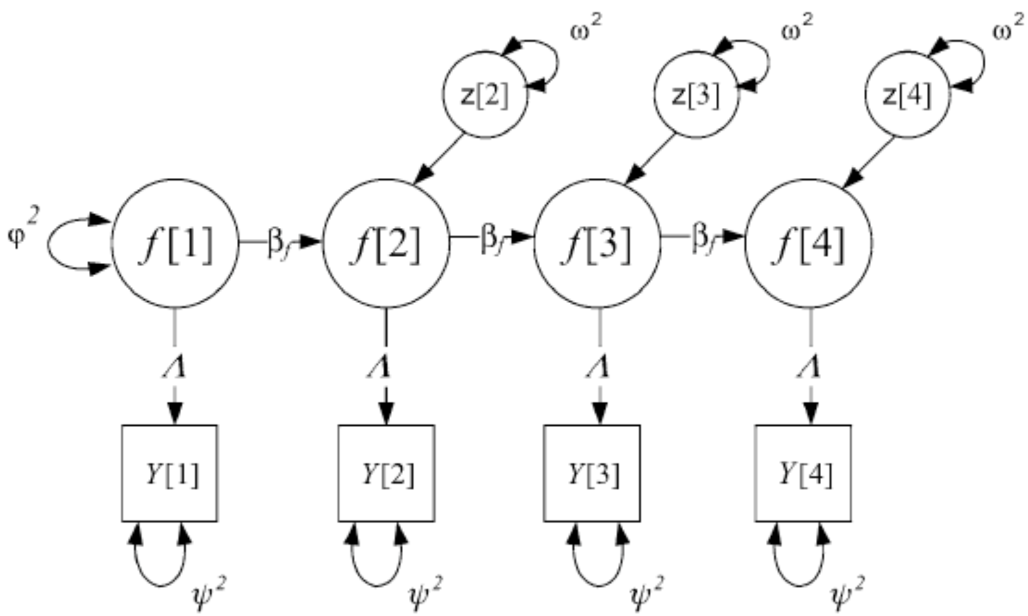
Common factor latent change score



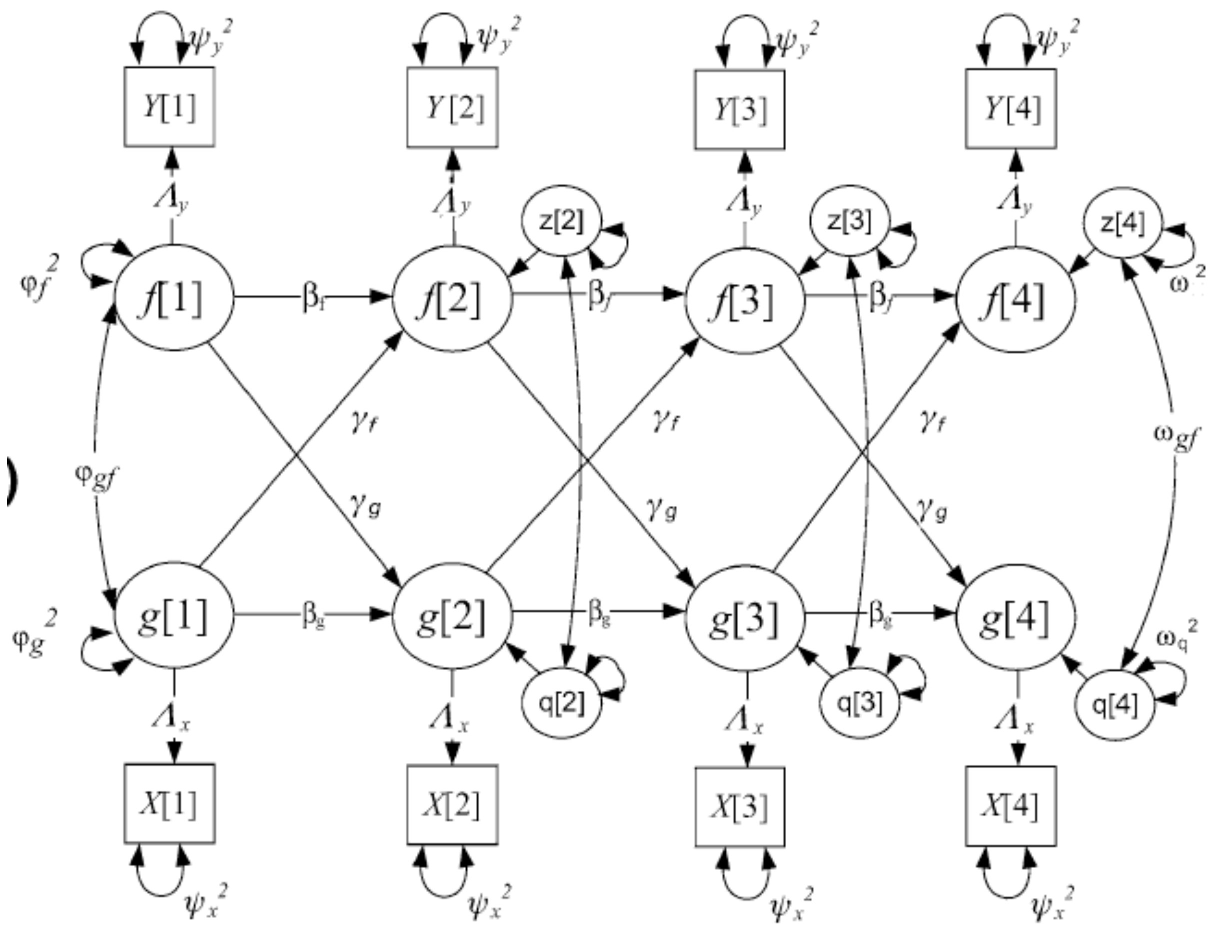
multiple common factors crossed lagged regression



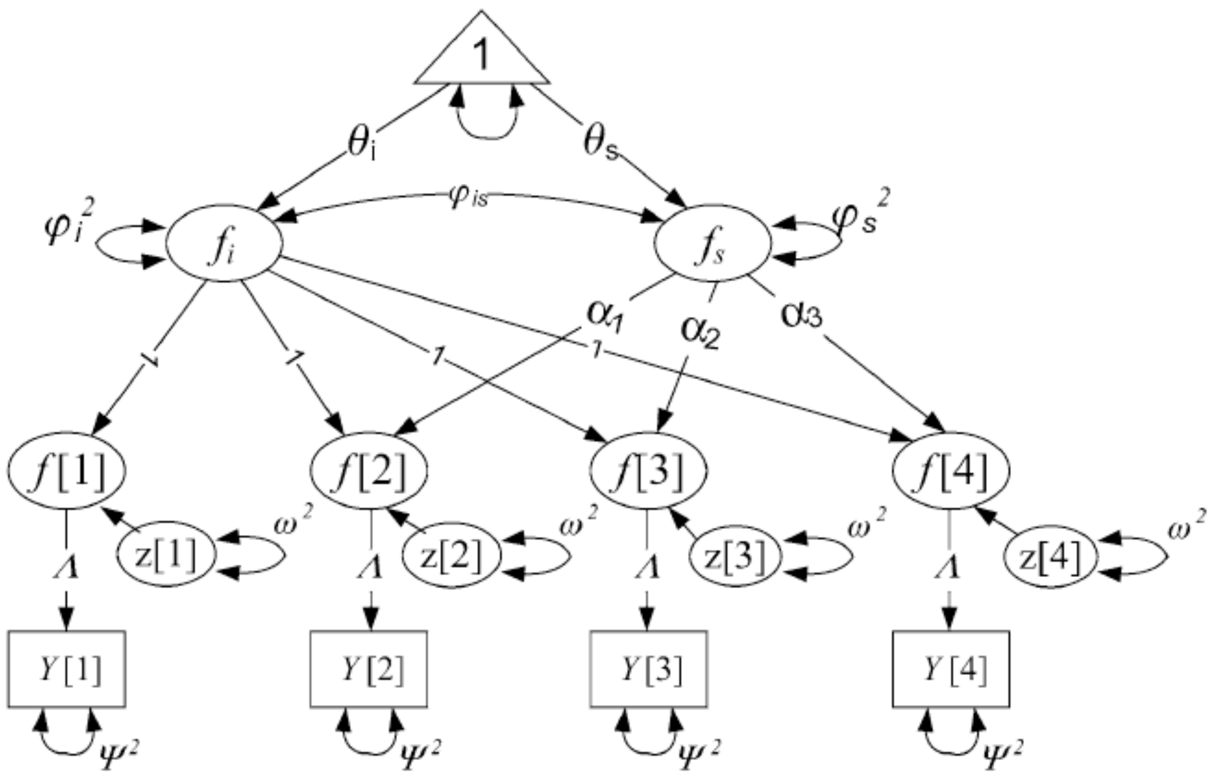
one factor Quasi-Markov simplex



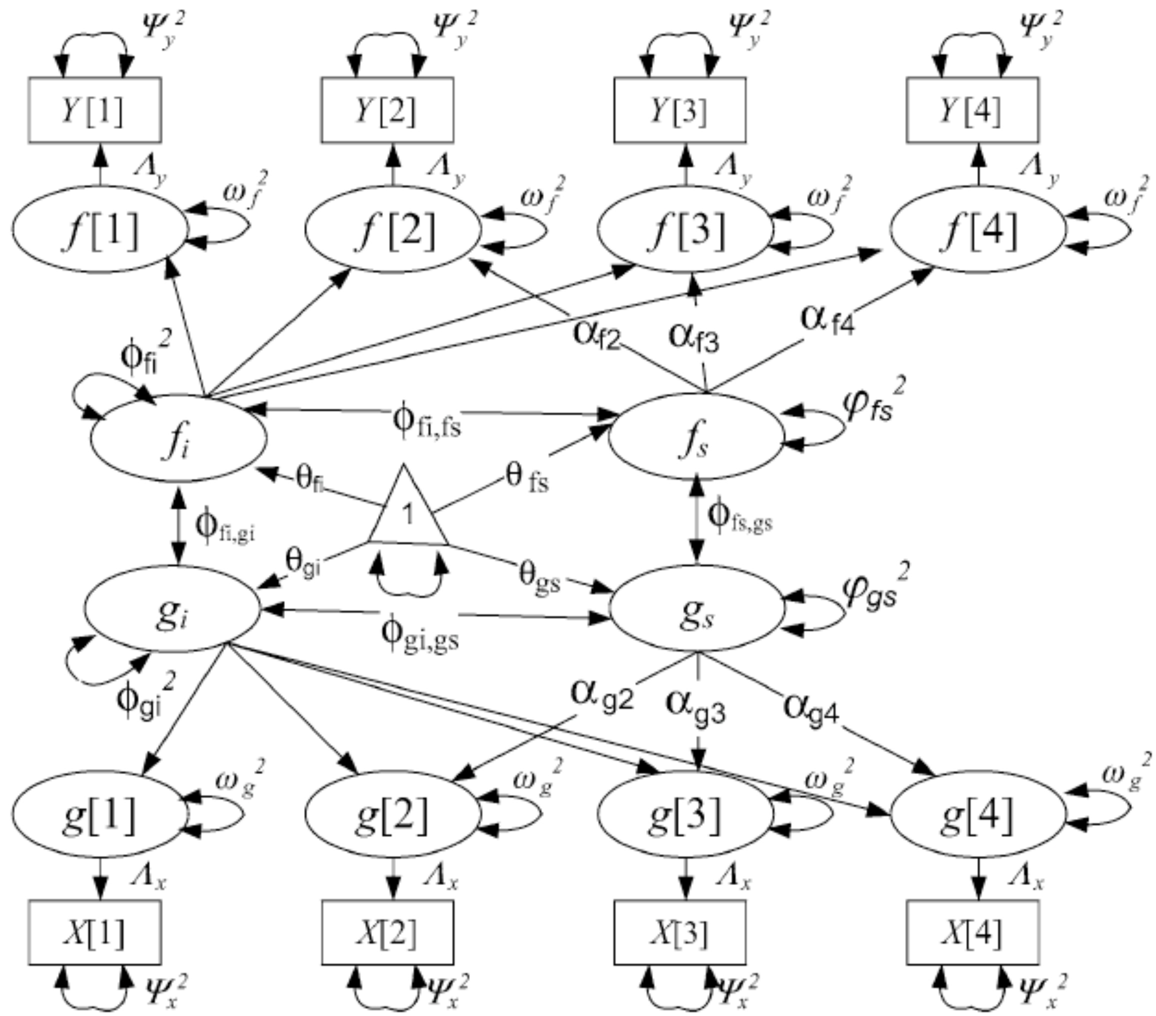
Cross lagged regression over multiple occasions



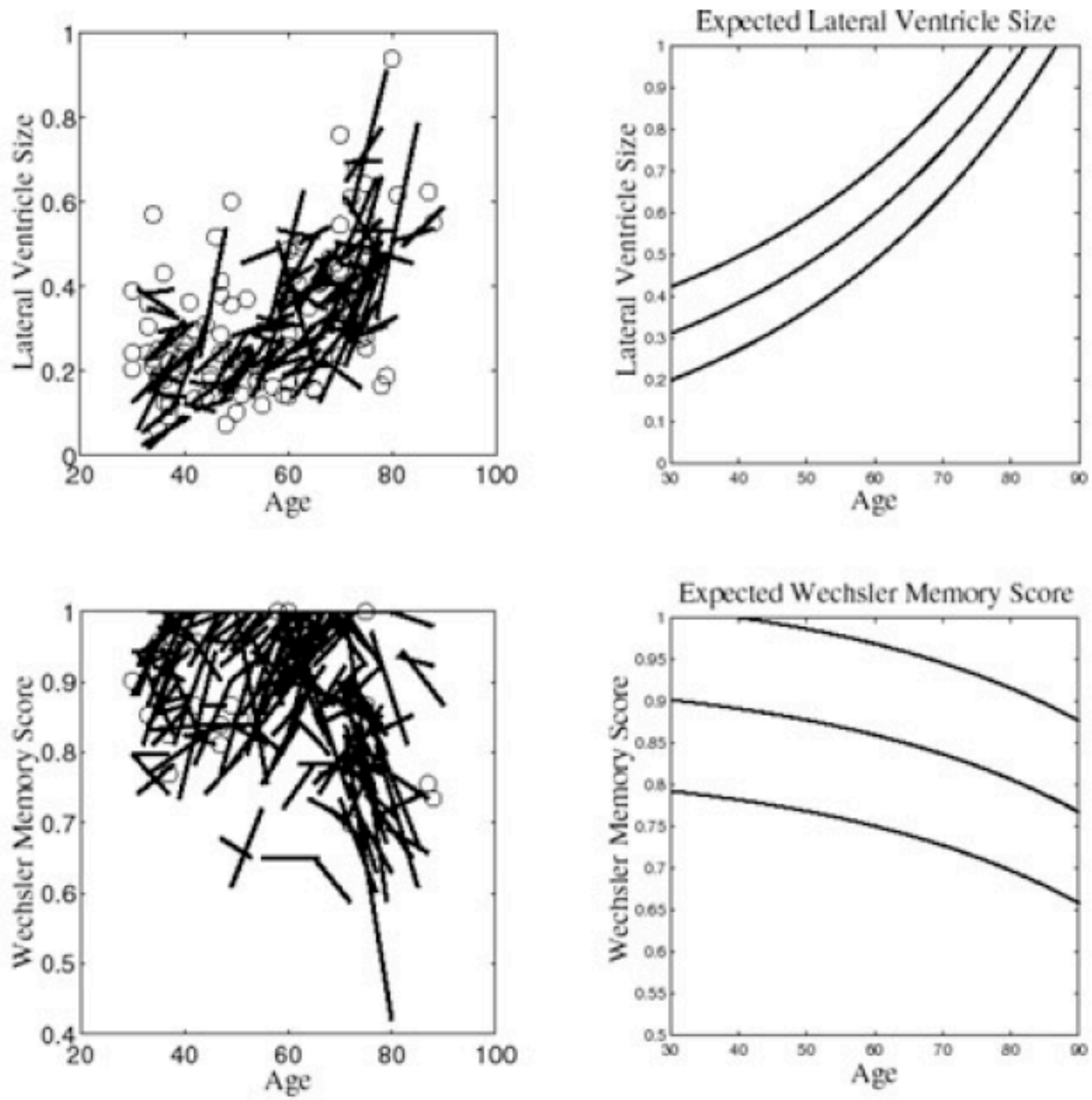
Latent Growth curve models: one factor



Bivariate growth curves



Real and expected change



expected change (vector field)