Latent Variable Methods Workshop
Longitudinal Data Analysis Models

Tue Sep 2 - Thu Sep 4 2025

Day 1 : Overview, Path Analysis, Factor Analysis, Structural Equation Modeling, and Effectively Using Mplus Software

lvmworkshop.org

Brown University

Zoom September 2, 2025

Day 1

- Orientation, overview and objectives
- Longitudinal data analysis (LDA)
- Structural equation modeling (SEM)
- Using Mplus & R/lavaan
- Lab exercise: CFA in Mplus & R/lavaan

Workshop Objectives

1. Introduce LDA in LVM (Mplus, R/lavaan) framework
   - A range of questions addressed by LDA
   - How approached using Mplus, R/lavaan
2. Experience applied LDA using Mplus, or R/lavaan
   - A range of models
     - Latent growth curve
     - latent change score models
     - survival analysis (Mplus only)
     - Mixture models (Mplus only)

Workshop Structure

- Morning Sessions
    - Lecture
    - Lab sessions
    - Run through examples
    - Ask questions at any time
- Afternoon Sessions
    - More lecture
    - More time to run through examples
    - Or with your own data
    - Individual assistance is the goal, emphasis on provided examples

Content Covered - Day 1

- Introduction to Longitudinal Data Analysis
  - Overview
  - Data Management for LDA
  - Introduction to Structural Equation Modeling (with Mplus)
  - Nuts and Bolts of Using Mplus
    - How to communicate about models
    - How to get data into Mplus
    - The Statistical model in general (arrays)

Content Covered - Day 2

- Example data set
- Latent Growth Curve Models (LGCM)
  - Unconditional LGCM
  - Conditional LGCM
  - Non-Linear LGCM
- Random Effects Models for growth with Mplus
- Multilevel model approach for growth modeling with Mplus
- Growth Mixture Models
- Multiple indicator LGCM
- Modeling Retest, Practice and other methods artifacts

Content Covered - Day 3

- Advanced Topics
  - LCGM
    - Multiple Group Models
    - Randomized Controlled Trials
    - Bivariate, Parallel, and Sequential Process Models
  - Latent Change Score (Dual Change Score Models)
    - Dual Change Score Models (DSCM), Bivariate DSCM
  - Bayesian Data Analysis and LGCM
  - Survival Analysis with Mplus
    - Discrete time, Continuous time
    - Simultaneous survival and growth modeling
  - Modeling with complex sample weights

Where to go for more information

- see links at
- https://www.statmodel.com/courses_openenroll.shtml
- The Center for Statistical Training by Curran-Bauer Analytics https://centerstat.org/
- https://www.lvmworkshop.org/home/--deprecated-pages/usefullinks

Other Resources

- What is longitudinal data analysis?
  - Singer JD & Willett JB. Applied longitudinal data analysis: Modeling change and event occurrence. 2003, New York: Oxford University Press.
- How do I do latent growth curve modeling?
  - Duncan TE, Duncan SC, & Strycker LA. An introduction to latent variable growth curve modeling: concepts, issues and applications. Second ed. 2006, Mahwah, New Jersey: LEA, Inc.
  - Newsom, J., Jones, R., & Hofer, S. (Eds.). (2011). Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health and Social Sciences. New York: Routledge.
- Tell me more about the math behind latent curve methods
  - Bollen KA & Curran PJ. Latent curve models: a structural equation perspective. Wiley series in probability and statistics. 2006, Hoboken, N.J.: Wiley-Interscience.
- More about structural equation modeling
  - Bryne, B. (2011). Structural Equation Modeling with Mplus: Basic Concepts, Applications and Programming. New York: Routledge.

Other Key References

- McArdle J. Five steps in latent curve modeling with longitudinal life-span data. In: Levy R, Ghisletta P, Le Gorr J, Spini D, Widmer E, eds. Towards an Interdisciplinary Perspective on the Life Course (10). San Diego, CA: JAI Press; 2005
- McArdle J, Epstein D. Latent growth curves within developmental structural equation models. Child Development. 1987;58:110-133.
- McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. Annual review of psychology. 2009;60:577-605.
- Curran, PJ, & Muthén, BO. Testing developmental theories in intervention research: Latent growth analysis and power estimation. American Journal of Community Psychology. 1999; 27:567-595.
- Jung T, & Wickrama, KAS. An introduction to latent class growth analysis and growth mixture modeling. Social and Personality Psychology Compass. 2008;2:302-317.

Introduction

Longitudinal Data Analysis

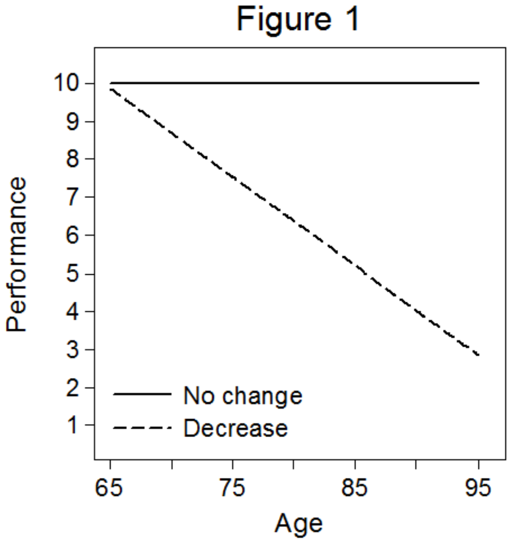
This Workshop

- Applied statistical data analysis of longitudinal data
  - Latent growth curve (LGC) or latent growth model (LGM)
  - Survival models
  - Latent change score models
  - Mixture models
    - growth mixture
    - change score mixture

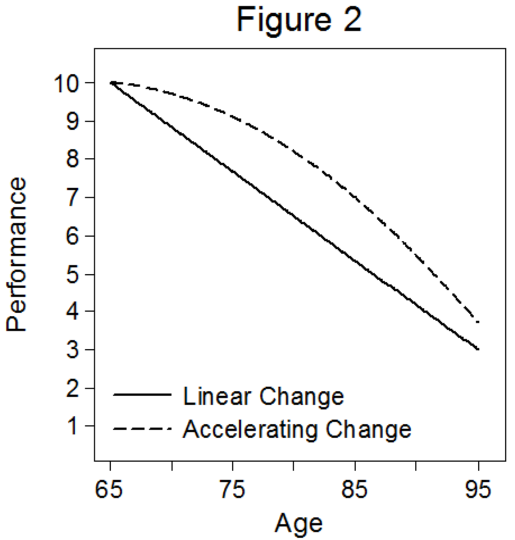
The Motivating Questions

- How does Y change over time? Y could be
  - cognitive functioning, neuropsych performance
  - mood, depression
  - weight
  - number of falls
  - etc.
- What factors predict who changes slower or faster on Y

About Change



About Change



The M in LGM

Why modeling is important

Change Scores

Since life-span research is, in essence, concerned with change over time, it is tempting to assume that the basic psychometric datum with which one works is the change score. Thus, if one is making comparisons of intelligence test scores at two points in time (symbolized by X_1 and X_2), it is assumed that the interest is in $(X_2 - X_1)$ and that such scores should be computed as one of the first steps in performing analyses. Actually, both the history of the problem and the logic of investigation indicate that the last thing one wants to do is to think in terms of or compute such change scores unless the problem makes it absolutely necessary. (p87, emphasis added)

Nunnally, J. C. (1973). Research Strategies and Measurement Methods for Investigating Human Development. In J. R. Nesselroade & H. W. Reese (Eds.), Life-Span Developmental Psychology: Methodological Issues (pp. 87-88). New York: Academic Press.

The Problems with $Y_t - Y_{(t-1)}$

- Classical
  - Reliability of Change Scores
  - Regression to the Mean
  - "Law" of Initial Values
- Empirical
  - VAR(observed change) $>$ VAR(true change)
  - Floor/Ceiling Effects
  - Limited information / assumption of linear change

Change Scores

- Difference between current and prior state
- Easy to interpret
- Easy to communicate meaning
- Really only useful in completely randomized designs when comparison and control groups are (or can be assumed to be) equivalent at baseline
- ...and even still may not be most powerful approach

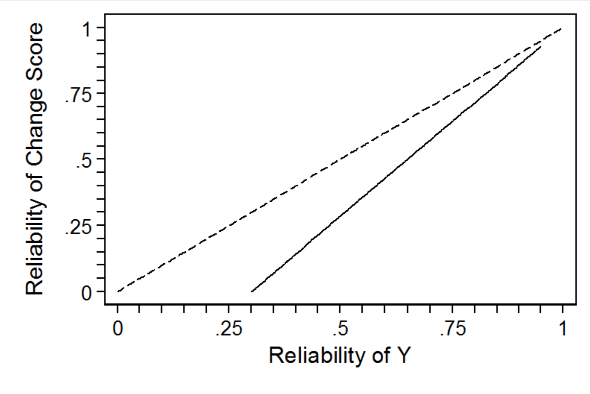
(Un)Reliability of Change Scores

- The first line of criticism of change scores
  - Cronbach, L., & Furby, L. (1970). How we should measure "change" - or should we? Psyc Bull 74; 68-80.
  - Lord FM. The measurement of growth. Educ Psychol Meas 1956;16:421-437.
  - Lord FM. Further problems in the measurement of growth. Educ Psychol Meas 1958;18:437-454.
- Aging research example
  - van Belle, G., Uhlmann, R. F., Hughes, J. P., & Larson, E. B. (1990). Reliability of estimates of changes in mental status test performance in senile dementia of the Alzheimer type. J Clin Epi, 43(6), 589-95.
- See Also
  - Galton, F. (1886). Regression towards mediocrity in hereditary stature. Journal of the Anthropological Institute, 15, 246-263.

(Un)Reliability of Change Scores

- Premise is not universally accepted.
  - Rogosa D: Myths about longitudinal research. In Schaie K, Cambell R, Meredith W, Rawlings S (eds), Methodological issues in aging research. New York, Springer Publishing Company, 1988, 171-209
  - Nesselroade JR: Application of multivariate strategies to problems of measuring and structuring long-term change. In Goulet LR, Baltes PB (eds), Life-span developmental psychology: Research and theory. New York, Academic Press, 1970, 193-207
- Detailed review
  - Wittmann, W. (1997). The reliability of change scores: many misinterpretations of Lord and Cronbach by many others; revisiting some basics for longitudinal research. Methodology conference: Evaluation of Change in Longitudinal Data, Nürnberg. `http://www.psychologie.uni-mannheim.de/psycho2/publi/papers/reliab.pdf`

(Un)Reliability of Change Scores



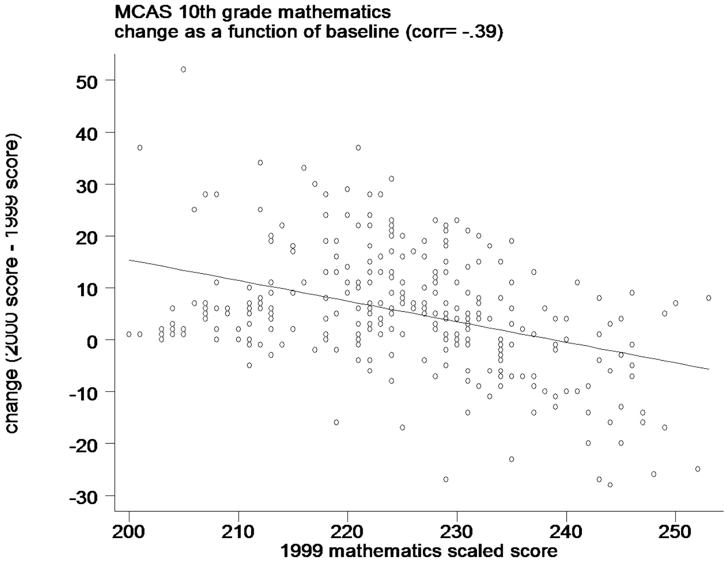
$$r_\Delta = \frac{r_{Y(t)} + r_{Y(t-1)} - 2 \times r_{Y(t),Y(t-1)}}{2 - 2 \times r_{Y(t),Y(t-1)}}$$

r_{YY} = reliability = 0.80; $r_{Y(t),Y(t-1)} = 0.3$

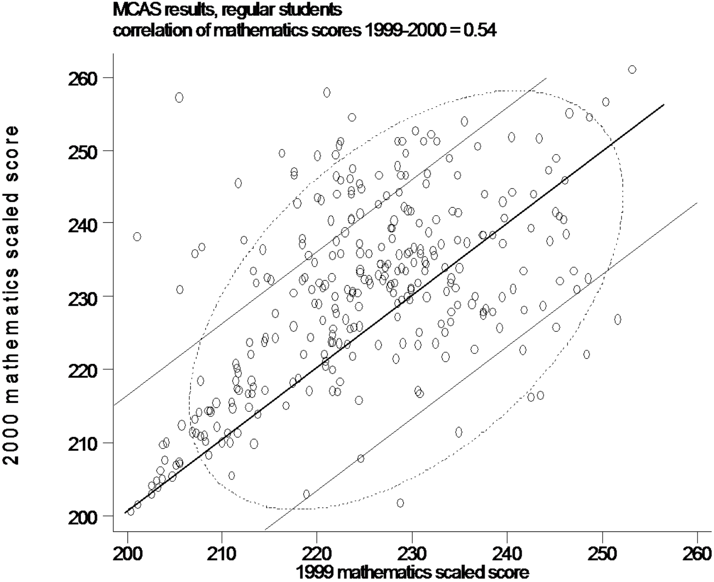
Streiner DL & Norman GR (1995). *Health Measurement scales*. New York. Oxford.

What Is Regression to the Mean?

- The only thing consistent about regression to the mean is *the absence of explicit, defensible definitions of the phenomenon* (p186).
- Rogosa, D. (1988). Myths about longitudinal research. In K. Schaie, R. Campbell, W. Meredith, & S. Rawlings (Eds.), Methodological issues in aging research (pp. 171-209). New York: Springer Publishing Company.



$COR(y_2, y_1) = 0.54;\ COR(y_2 - y_1, y_1) = -0.39$



$COR(y_2, y_1) = 0.54;\ COR(y_2 - y_1, y_1) = -0.39$

Some top-scoring schools faulted

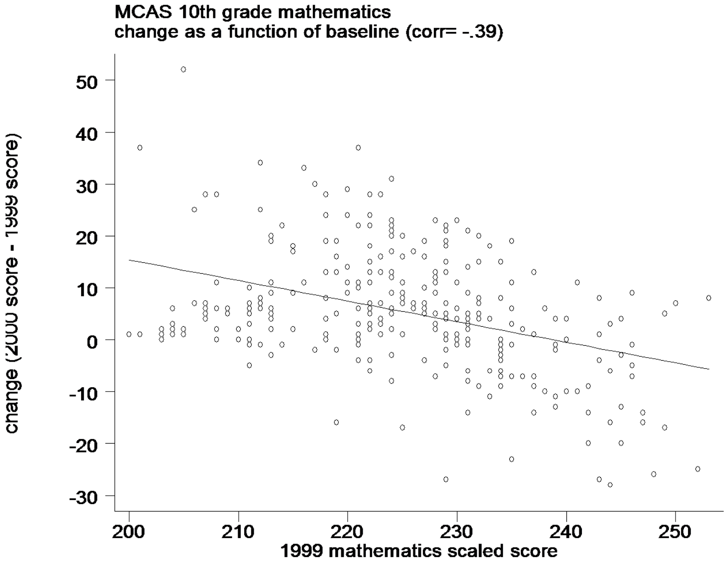
Many questioning MCAS assessment
By Anand Vaishnav, Globe Staff, 1/10/2001

A state report rating public schools on three years of MCAS results shows some of Massachusetts's highest-performing schools "failed to meet" expectations, a description that angered many educators across the Commonwealth.

In the state's first attempt to give schools targets to reach on the Massachusetts Comprehensive Assessment System exams, 56 percent of 1,539 schools did not improve enough to meet those goals. The remaining 44 percent "exceeded, met, or approached" their growth targets, according to the report released yesterday.

The results largely bode well for urban districts that had the most ground to make up. But the state's top public schools - in wealthier districts such as Harvard, Newton, Wellesley - performed so well in 1998 that boosting scores any higher was nearly impossible. As a result, many of those schools received an "F," for "failed to meet" expectations.

Boston Globe.



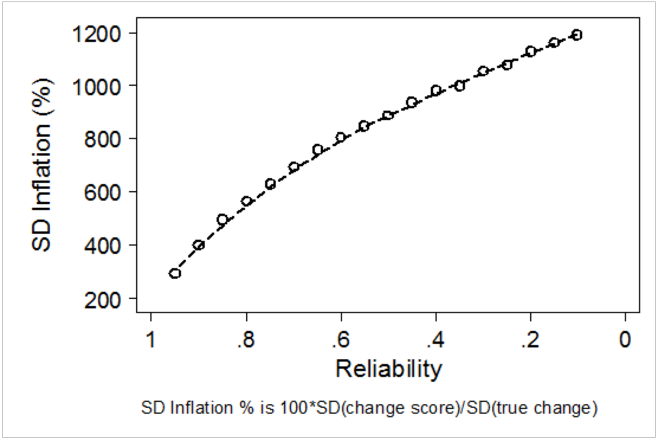
$COR(y_2, y_1) = 0.54; \; COR(y_2 - y_1, y_1) = -0.39$

VAR(observed change) > VAR(true change)

$$
\begin{aligned}
Y &= T + e \\
Y_2 - Y_1 &= (T_2 - T_1) + (e_2 - e_1) \\
VAR(Y_2 - Y_1) &= VAR(Y_2) + VAR(Y_1) - 2COV(Y_2, Y_1) \\
&= VAR(T_2) + VAR(T_1) - 2COV(T_2, T_1) + \\
&\quad VAR(e_2) + VAR(e_1) - 2COV(e_2, e_1)
\end{aligned}
$$

- Even if errors are uncorrelated, the variance of observed change score compounds measurement errors
- The lower the $COR(Y_t, Y_{t-1})$, the more variable is the change score
- Observed change score variance can be greater than baseline variance

VAR(observed change) > VAR(true change)

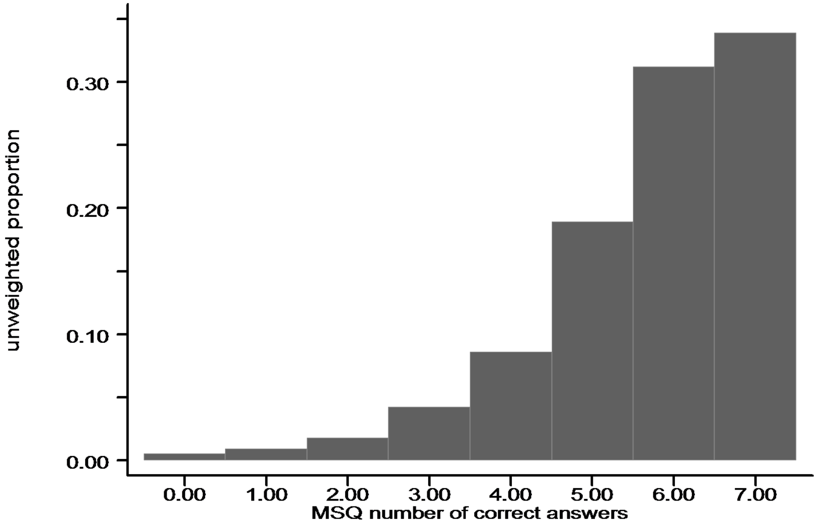


Change is random, additive, and correlated with baseline (r = -0.40)
SD(true change) = 0.1SD(baseline)

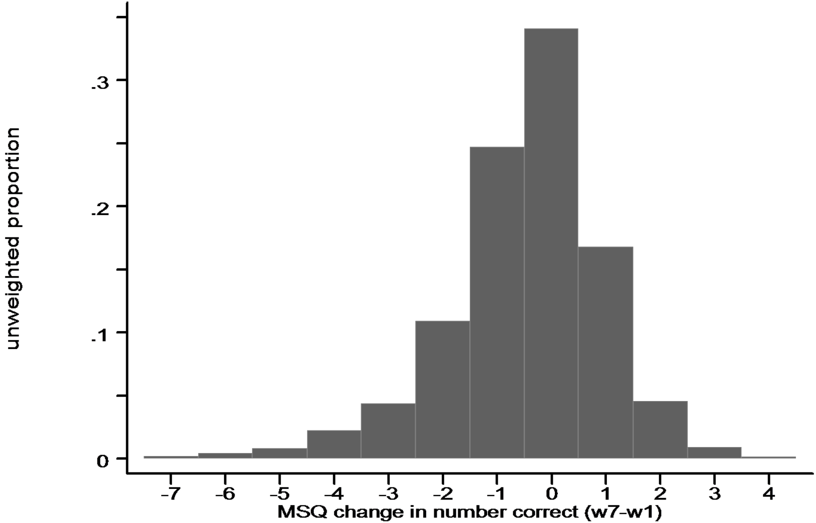
Floor/Ceiling Effect

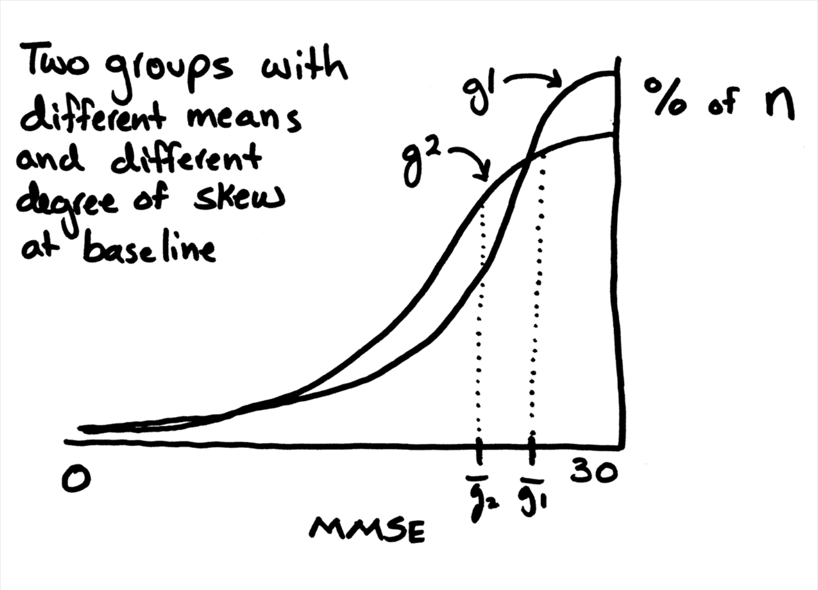
- Many measures in health research suffer floor/ceiling effects: a large proportion of the sample score at lowest or highest possible level
- For Example
  - Cognitive function
  - Depressive symptoms
  - Physical functioning

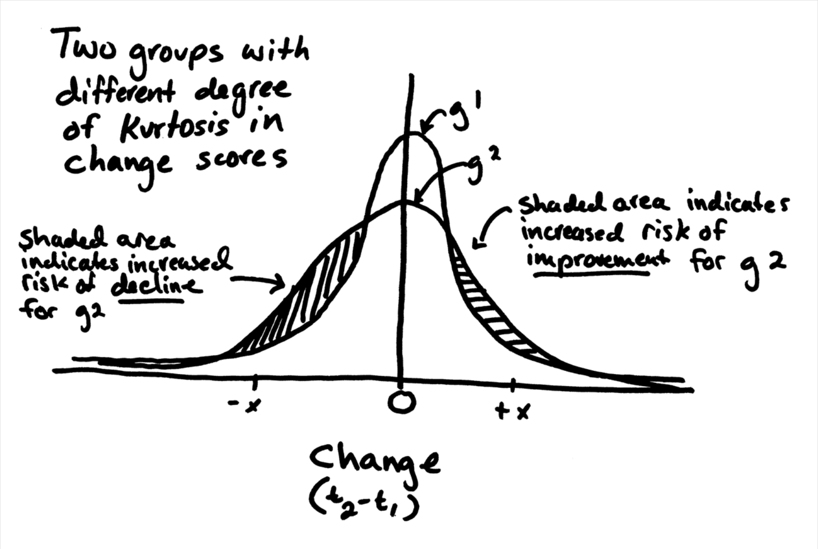
Distribution of Cognitive Functioning in EPESE

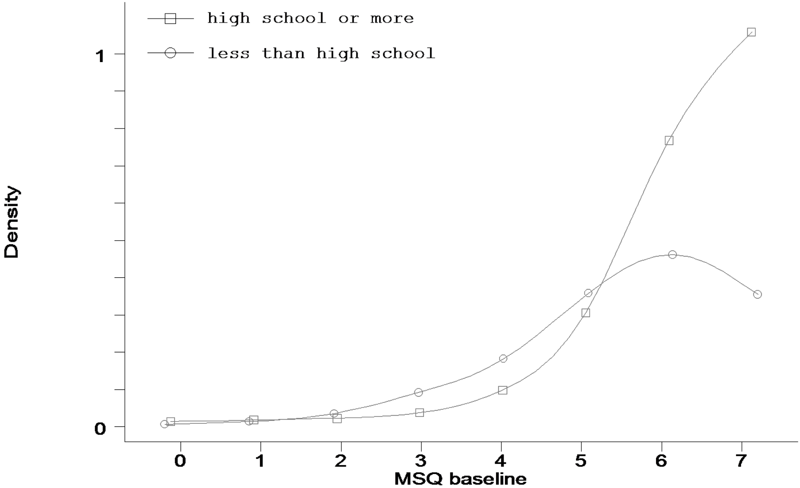


Distribution of Cognitive Functioning (w7-w1) in EPESE

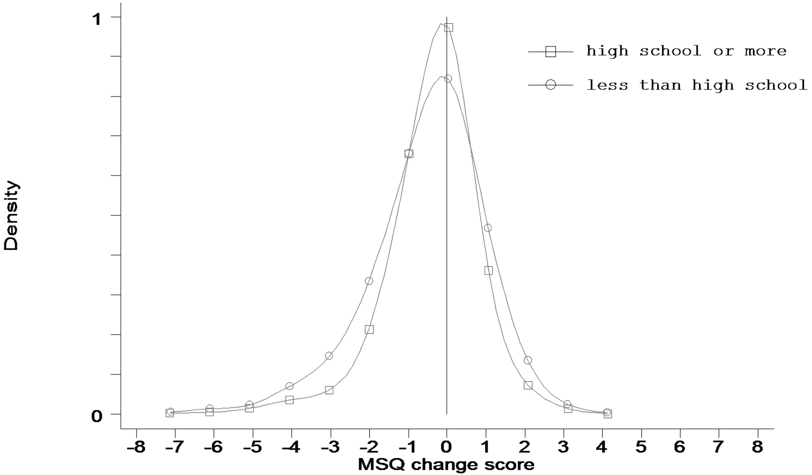



Two groups with different means and different degree of skew at baseline
g1
g2
% of n
0
30
g2
g1
MMSE


Two groups with different degree of Kurtosis in change scores
g1
g2
Shaded area indicates increased risk of improvement for g2
Shaded area indicates increased risk of decline for g2
-x
0
+x
Change
(t2-t1)



Density plot. MSQ scores at year 1. EPESE (N=13,697) by education group

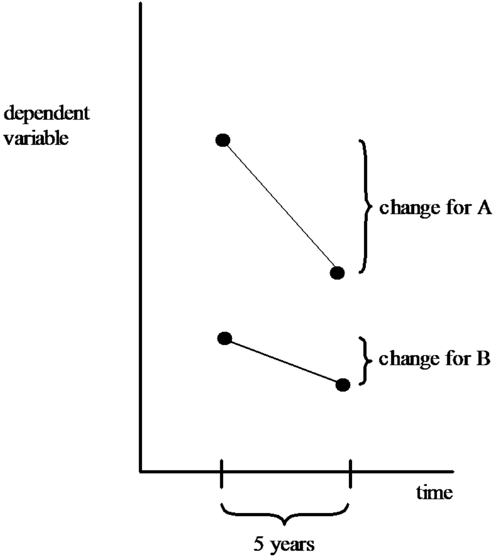


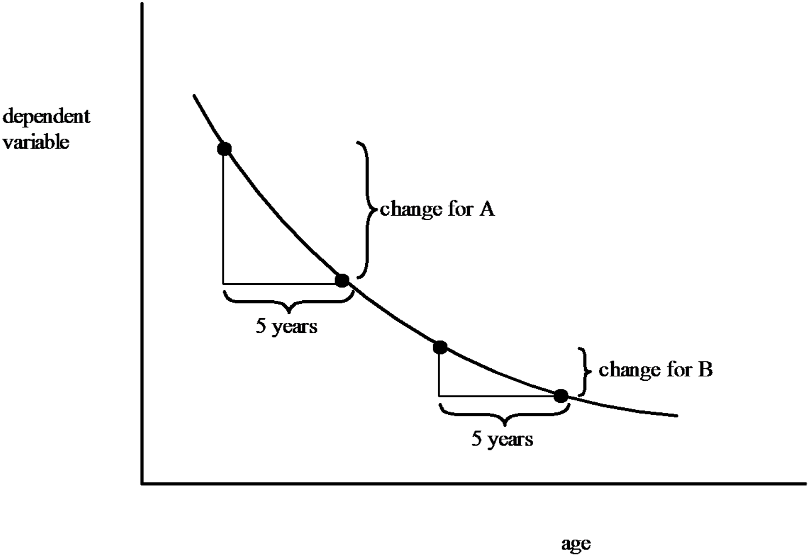
Density plot. MSQ change scores year 7 - year 1. EPESE (N=7,939) by education group

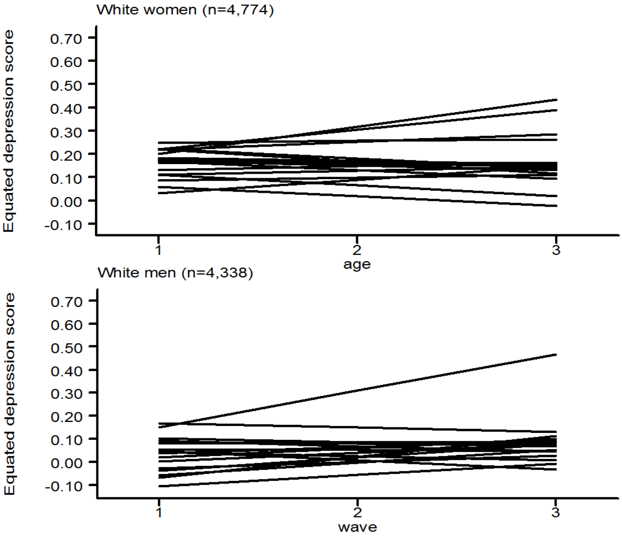
Table. Cognitive decline (2 or more points lower at follow-up) and cognitive improvement (2 or more points higher at 6th annual follow-up) among EPESE participants, by level of educational attainment.

Level of Education (years)	%	OR
Cognitive Decline		
less than 12	22	1.7 (1.5, 2.0)
12 of more	14	1.0
Cognitive Improvement		
less than 12	6	1.7 (1.4, 2.1)
12 of more	4	1.0

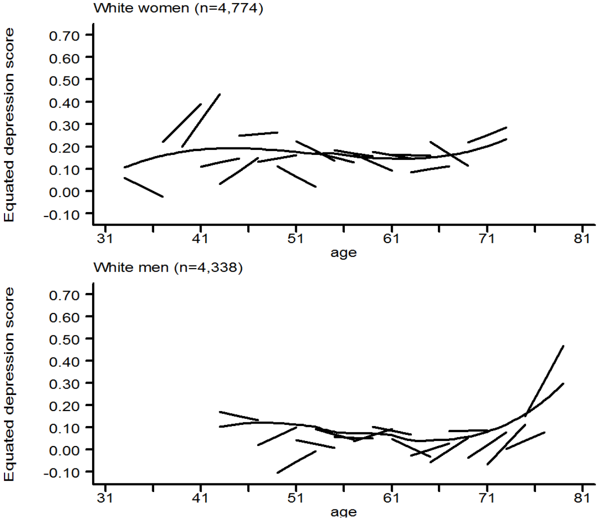
Assumption of Linear Change




dependent variable
change for A
5 years
change for B
5 years
age



Birth cohort-specific depression symptom trajectories, Health and Retirement Study waves 1,2,3



Birth cohort-specific depression symptom trajectories, Health and Retirement Study waves 1,2,3

An Approach

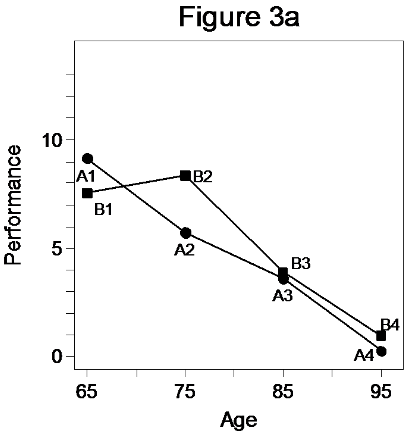
Model, Rather than Calculate, Change

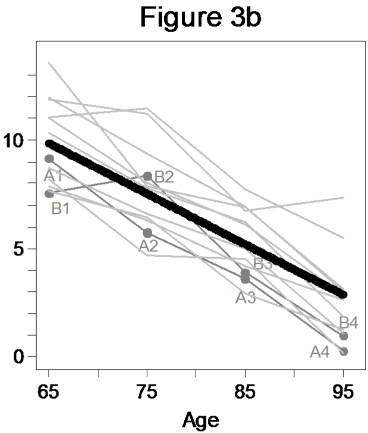
Latent Growth Curve Models

- Appropriate when outcome
  - changes systematically over time, on a time interval relevant to the study
  - when there at least three, preferably four repeated observations
  - is discrete or continuous
- Alternatives
  - Survival models (for "absorbing events": events that can only be observed once)
  - Latent change score models
  - Autoregressive models

Latent Growth Curve Models

- Sometimes equivalent to
  - mixed effect models
  - random coefficient regression models
  - hierarchical or multilevel models
  - HLM, Mx, PROC Mixed, xtmixed
- Can be different
  - how time is handled (data vs. model parameter)
  - how variances are handled
  - how flexible the modeling framework is





Observed scores and an overall sample mean trajectory

Panel 3a shows two persons (A and B) and their performance at four ages (A1=person A at age 65, A4=person A at age 95). Panel B shows how person A and B and their performance trajectory are distributed in a larger sample of persons. Panel 3b also shows a group average trajectory (heavy line).

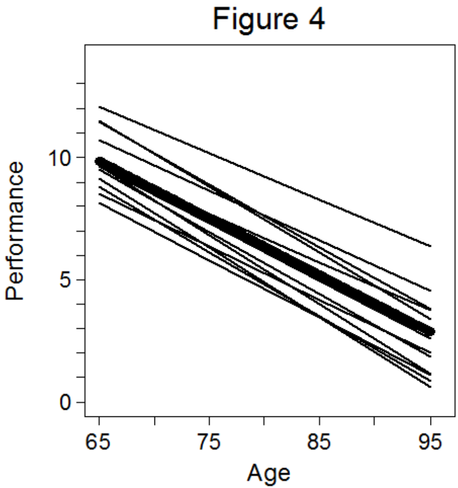
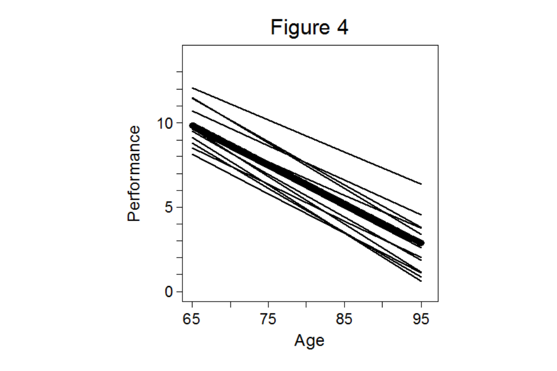
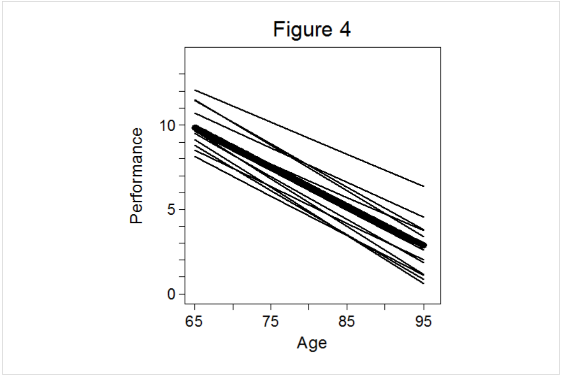


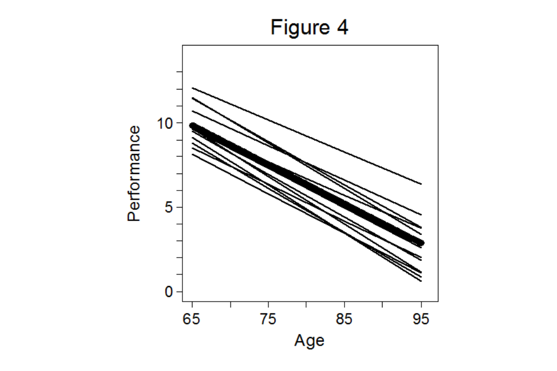
Figure 4
Performance
10
5
0
65
75
85
95
Age



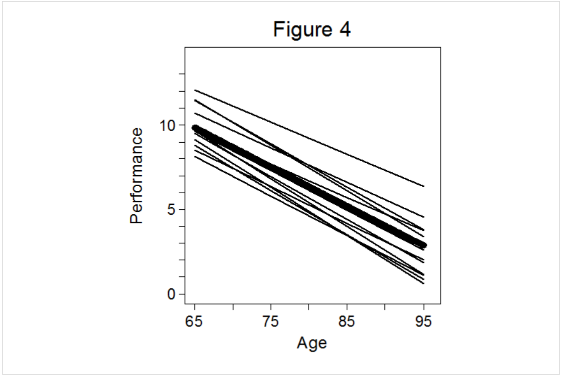
$$\text{Performance} = \text{Intercept} + \text{Slope} \times \text{Age} + \text{Residual}$$
$$\text{Intercept} = \text{Mean}_{65} + \text{Disturbance}$$
$$\text{Slope} = \text{Mean Change} + \text{Disturbance}$$



$$y_{it} = \eta_{i1} + \eta_{i2} \times \text{Age}_{it} + \epsilon_{it}$$
$$\eta_{i1} = \alpha_1 + \zeta_{i1}$$
$$\eta_{i2} = \alpha_2 + \zeta_{i2}$$



$$y_{it} = (\alpha_1 + \zeta_{i1}) + (\alpha_2 + \zeta_{i2}) \times \text{Age}_{it} + \epsilon_{it}$$



$$y_{it} = (\alpha_1 + \alpha_2 \times \text{Age}_{it}) + (\zeta_{i1} + \zeta_{i2} \times \text{Age}_{it} + \epsilon_{it})$$
$$y_{it} = (\textit{fixed part}) + (\textit{random part})$$

Time

- Time as data (Age_{it})
  - Random effects models, Mixed effect models, Multilevel/hierarchical models
  - Accommodates irregular assessment schedule
  - Time varying covariate effects vary over people
- Time as model parameter (Age_t)
  - LGCM
  - Assumes regular reassessments
    - ⋆ not necessarily evenly spaced
  - Allows flexible curve shape
  - Time varying covariate effects vary over time
- Mplus will do both

Summary

- Change scores have issues
- Modeling doesn't remove those issues but can accommodate (estimate) them
- There are many paths to enlightenment (hierarchical, multilevel, mixed, random, latent growth)
- Different approaches can be equivalent, mostly depending upon how time is handled

Don't get too hung up on the validity/invalidity of one model over another

- LGCM $\approx$ Mixed/Random Hierarchical
  - Curran PJ: Have Multilevel Models Been Structural Equation Models All Along? Multivariate Behavioral Research 38:529-569, 2003
- LGCM $\approx$ paired t-tests, RM-ANOVA
  - Voelkle MC: Latent growth curve modeling as an integrative approach to the analysis of change. Psychology Science 49:375, 2007
- But, mixed, random, hierarchical, change score, repeated measures ANOVA, all can be viewed as special cases of SEM models of change
- So SEM model is a flexible, extensible and general approach to change and worth mastering.

Questions

Structural Equation Modeling

Communicating and expressing modeling ideas using latent variables

Learning Note

- Gist
  - What am I doing?
  - How do I explain it to my colleagues?
- Statistical concepts
  - What am I really doing?
  - How do I try to explain it to my statistician colleagues?
- Person-machine interface
  - How do I do this?

Covariance Structure Model

Regressions among dependent (or involving latent variables)	(Continuous) Latent Variables **Yes**	**No**
Yes	Structural Equation Modeling	Path Analysis
No	Factor Analysis	Multivariate Regression

How to Learn about SEM

How

- Brown T. Confirmatory factor analysis for applied research. (2006). Guilford Press
- Jï¿œreskog and Sï¿œrbom, LISREL user's reference guide. (from version 7)
- Raykov and Marcoulides, A first course in structural equation modeling. 2006
- Raykov T, Tomer A, Nesselroade JR. Reporting structural equation modeling results in Psychology and Aging: Some proposed guidelines. Psychol Aging. 1991;6:499-503.
- Resources to help you use Mplus. UCLA: Institute for Digital Research & Education (was: Academic Technology Services, Statistical Consulting Group). https://stats.idre.ucla.edu/mplus/

Why

- De Stavola B et al. Statistical issues in life course epidemiology. Am J Epidemiol. 2006;163:84-96.
- Hoyle RH, Smith GT. Formulating clinical research hypotheses as structural equation models. J Consult Clin Psychol. 1994;62:429-440.
- Muthï¿œn B. Psychometric modeling in epidemiology with an emphasis on longitudinal analysis of alcohol use. Alcoholic Beverage Medical Research Foundation. 1994.

Latent Variable Models

Latent variables are variables that are not directly observed

- Presumed to exist
- Presence inferred from observed variables
- Account for covariation among observed variables

Latent and observed variables can be continuous or categorical

Ridiculously Brief and Highly Selected History of LV Models

- Spearman (1904) - Factor Analysis
- Wright (1934) - Path Analysis
- Lazarsfeld (1950) - Latent Class
- Lord (1952) - Item Response Theory
- Jöreskog (1973) - LISREL model
- Muthén (1987) - LISCOMP model
- Muthén (1998) - General LVM (Mplus)
- Asparouhov + Muthén (2008) - Exploratory SEM
- Rosseel (2010) - R/lavaan

Other Latent Variables

- Measurement errors
- Random effects
- Frailties, liabilities
- Variance components
- Missing data

see www.statmodel.com/download/Topic 3.pdf

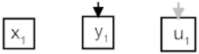
Path diagrams

Path analysis was invented by the geneticist Sewall Wright in 1918. Wright also made use of path diagrams. However, the diagrams used here follow a set of rules [loosely] for drawing path diagrams that has been described as the LISREL model (specifically, LISREL submodel 2). LISREL was developed in the 1960s and released widely in the 1970s by Karl G. Joreskog.

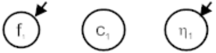
Challenges with path diagrams in LDA is that conventions for describing means and thresholds of observed (and latent) variables are not well developed. This ows to the historical fact that most early applications of covariance structure analysis (path analysis, factor analysis) concern the covariances among observed data, and variables were often standardized or centered (to mean 0).

The most widely used convention for describing threshold and mean structures is McArdle's RAM notation (Reticular Action Model).

Finally, it is worth mentioning that no specific knowledge of or ability to use and interpret path diagrams is required for using Mplus.



Observed variables in rectangles
Exogenous variables without residuals: covariate, cause
Endogenous variables (y, u) usually with residuals (errors): item, indicator, outcome
Sometimes y implies continuous, u implies discrete



Latent variables in ovals
Sometimes f and η imply continuous latent variables, c categorical
Continuous latent variables with disturbance terms
factors, random effects, classes, clusters, mixtures

Constant (=1)
Used to indicate means, thresholds, intercepts in regression relationships. Several other conventions are used for this purpose.

Straight arrows indicate regression relationships including regression and factor loadings. Sometimes labeled with a regression coefficient ($b, \beta, \gamma, \kappa, \lambda$, *; * implies freely estimated). If not labeled and others in model are, might be constrained to be equal to 1. If labeled, subscripts in the order to-from (b_{yx}).

Curved arrows indicate covariance (correlation) relationships.
Often omitted from among exogenous variables, even when included in the model.

Notes:

There is no formal and universally accepted language or syntax rules for path diagrams.

Different authors have their own conventions.

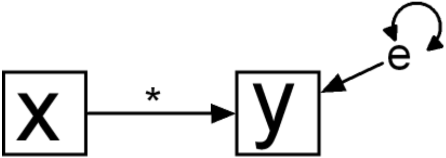
Path diagrams in publications are rarely complete images of the statistical model.

Path diagrams do not always translate directly into the syntax of an estimation program.

Path diagrams can be helpful tools. They can also be confusing and distracting.

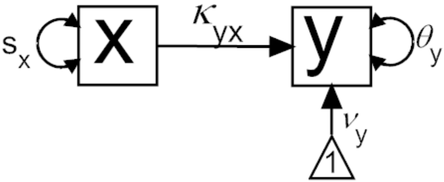
Path diagrams as used in SEM are in general not the same as directed acyclic graphs (DAGs).

Simple Regression



$$y = \nu + kx + \epsilon$$
$$E(y) = \nu + kx$$

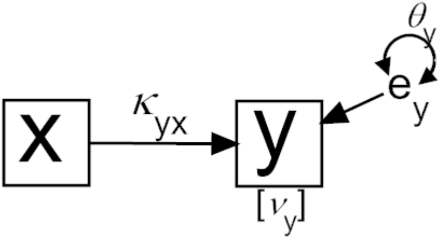
Simple Regression



$$y = \nu + kx + \epsilon$$
$$E(y) = \nu + kx$$

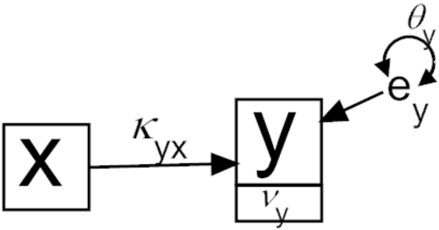
RAM path diagram conventions described (James Steiger, Vanderbilt University)

Simple Regression



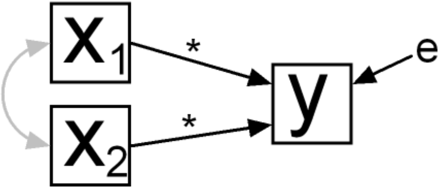
$$y = \nu + kx + \epsilon$$
$$E(y) = \nu + kx$$

Simple Regression



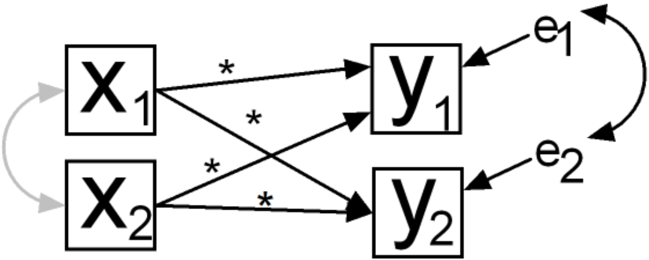
$$y = \nu + kx + \epsilon$$
$$E(y) = \nu + kx$$

Multivariable Regression



$$y = \nu + k_1 x_1 + k_2 x_2 + \epsilon$$

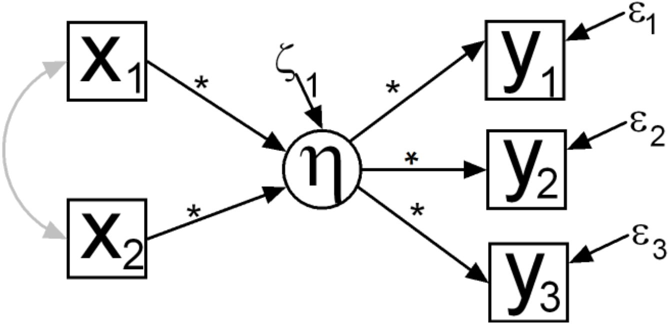
Multivariate Regression



$$y = Kx + \epsilon$$

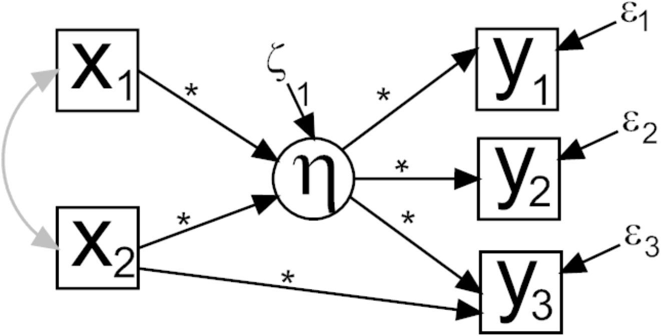
(Confirmatory) Factor Analysis with Covariates

(MIMIC = Multiple Indicators Multiple Causes)

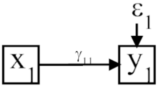


$$\mathrm{y} = \boldsymbol{\nu} + \Lambda\boldsymbol{\eta} + \mathrm{K}\mathrm{x} + \boldsymbol{\epsilon}$$
$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \Gamma\mathrm{x} + \boldsymbol{\zeta}$$

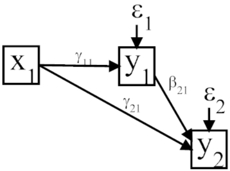
MIMIC Model with Direct Effect



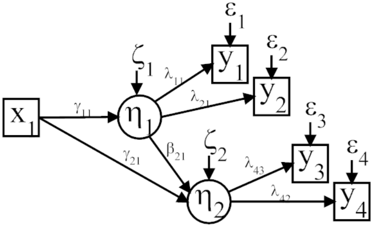
$$y = \boldsymbol{\nu} + \Lambda\boldsymbol{\eta} + Kx + \boldsymbol{\epsilon}$$
$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \Gamma x + \boldsymbol{\zeta}$$



regression

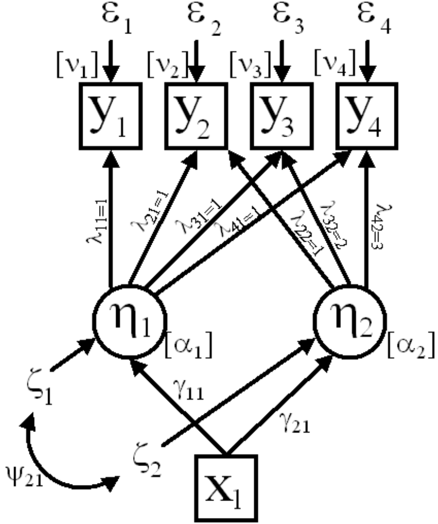


path analysis



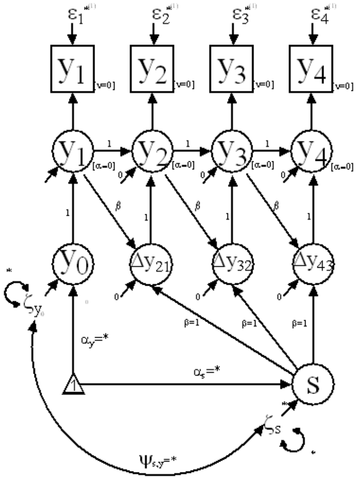
Structural Equation Model

Latent Growth Model (Linear Change)



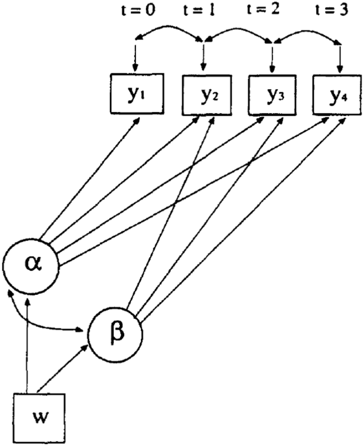
$$
\begin{aligned}
y_{it} &= \nu + \eta_1 \lambda_{t1} + \eta_2 \lambda_{t2} + \epsilon_{it} \\
\eta_1 &= \alpha_1 + \gamma_{11} x_{1i} + \zeta_1 \\
\eta_2 &= \alpha_2 + \gamma_{21} x_{1i} + \zeta_2
\end{aligned}
$$

Dual Change Score Model



$$y_{it} = y_0 + \Delta y_{t(t-1)} + \epsilon$$
$$\Delta y_{t(t-1)} = \beta y_{(t-1)} + s$$
$$y_0 = \alpha_y + \zeta_{y_0}$$
$$s = \alpha_s + \zeta_s$$

FIGURE 3
Graphical representation of a growth model for four time points.



Muthén BO, Khoo ST. Longitudinal studies of achievement growth using latent variable modeling. Learning and individual differences. 1998;10(2):73-101.

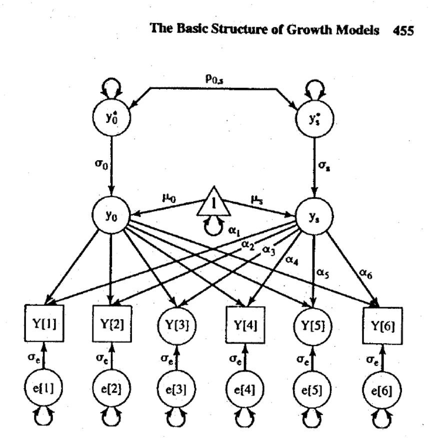


Figure 18.7 The basic latent growth structural model as a path diagram from McArdle & Epstein (1987) and McArdle & Hamagami (1992).

McArdle J, Nesselroade J. Growth curve analysis in contemporary psychological research. In: Schinka JA, Velicer W, eds. Comprehensive Handbook of Psychology, Volume Two: Research Methods in Psychology. New York: Wiley; 2002.

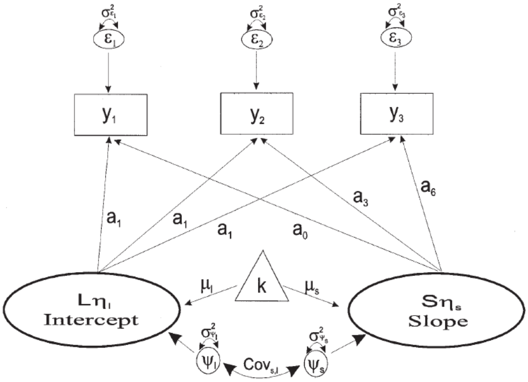


FIGURE 1 A latent growth model for studying change in well-being through Tai Chi.

Li et al. Enhancing the psychological well-being of elderly individuals through Tai Chi exercise: A latent growth curve analysis. Structural Equation Modeling. January 2001;8(1):53-83.

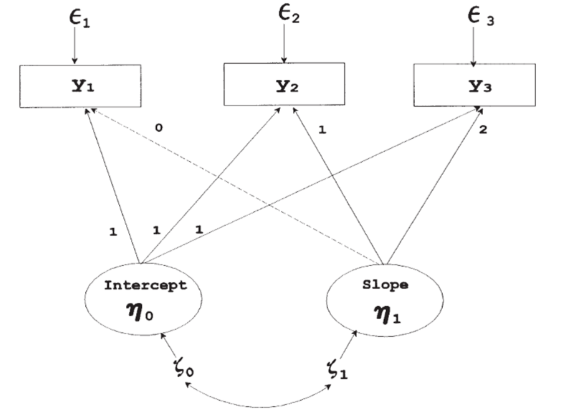
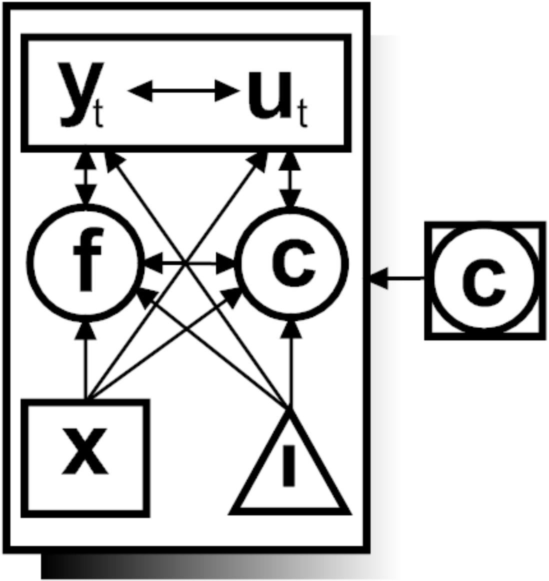


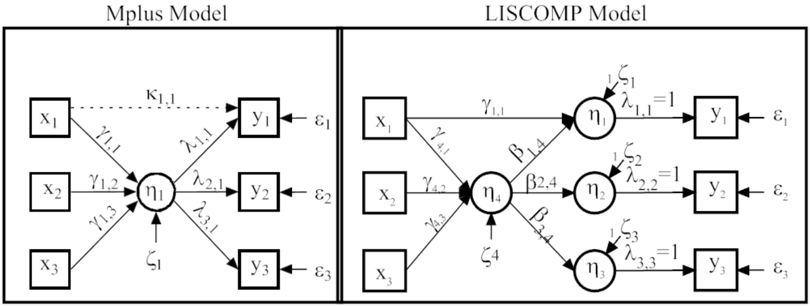
Figure 1: Simple Linear Latent Growth Model for Three Time Points
NOTE: The ys are observed repeated measures and the Greek letters are model parameters.

Wang et al. Evaluation of Hiv Risk Reduction Intervention Programs Via Latent Growth Model. Eval Rev 1999 23: 648-662

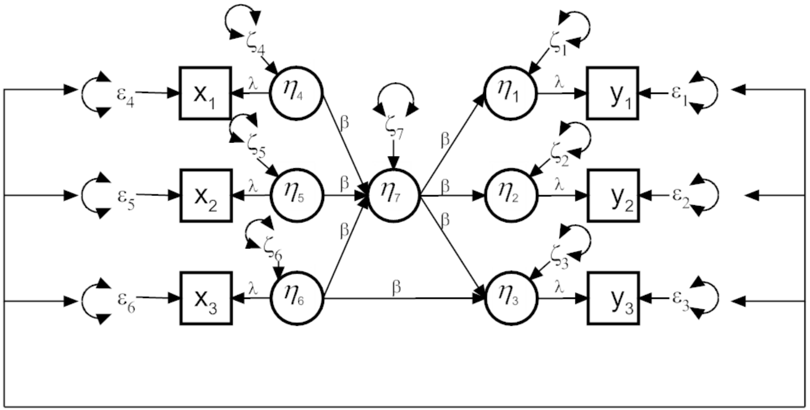
Mplus Framework



y_t
u_t
f
c
c
x
i


Mplus Model
LISCOMP Model

Mplus Model, All Y continuous



Mplus and LISCOMP Parameter Arrays

Mplus Variable and Parameter Arrays

K (kappa) a $p \times q$ matrix of slopes for the regressions of the p observed response variables (y's) on the q x-variables

Measurement Model

$$y^* = \nu + \Lambda\eta + Kx + \varepsilon$$

Structural Regression model:

$$\eta = \alpha + B\eta + \Gamma x + \zeta$$

LISCOMP Variable and Parameter Arrays

- x (exogenous variables) a q-vector, i.e., $x_1, x_2, x_3 \ldots x_q,$
- y (observed response variables) a p-vector i.e., $y_1, y_2 \ldots y_p$
- τ (tau) a vector of length (c-1), equal to the number of thresholds; where c is number of categories
- ν (nu) a $p \times 1$ vector of measurement intercepts, where p is the number of response variables ($y_1, y_2 \ldots y_p$)
- Λ (lambda) a p x m matrix of measurement slopes/loadings, where p is equal to the number of y's, and m is equal to the number of latent constructs (η's).
- I (identity matrix), a $m \times m$ identity matrix
- Θ (theta) a $p \times p$ covariance matrix for the residuals in the measurement relations
- α (alpha) a $m \times 1$ vector of the structural regression latent variable (η's) means and intercepts. *May be fixed to zero in the for categorical variables.*
- B (beta) a $m \times m$ matrix of slopes for the regressions among the latent variables (η's) with zeros as it's diagonal elements.
- Γ (gamma) a $m \times q$ matrix of slopes for the regressions of the m latent variables (η's) on the q x-variables
- Ψ (psi) a $m \times m$ covariance matrix for the latent variables (η's) and the residuals in the latent variable relations
- Δ (delta) diagonal of a $p \times p$ matrix of scaling factors. Only the p diagonal elements are parameters.

Measurement Model

$$y^* = \nu + \Lambda\eta + \varepsilon$$

Structural Regression model:

$$\eta = \alpha + B\eta + \Gamma x + \zeta$$

Link to PDF handout version of this slide

→ https://s3.amazonaws.com/lvmworkshop/Jones-2013-Understanding-LISCOMP-and-Mplus-Parameter-Arrays.pdf

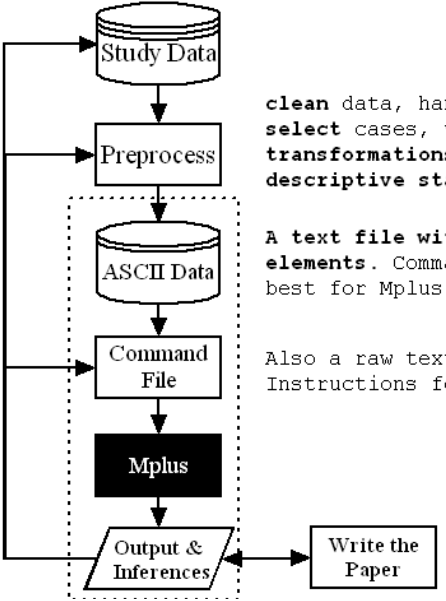
$$\mathrm{y} = \boldsymbol{\nu} + \Lambda\boldsymbol{\eta} + \mathrm{Kx} + \boldsymbol{\epsilon}$$
$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathrm{B}\boldsymbol{\eta} + \Gamma\mathrm{x} + \boldsymbol{\zeta}$$
$$V(\mathrm{y}) = \Lambda\Psi\Lambda^{\intercal} + \Theta$$
$$V(\boldsymbol{\epsilon}) = \Theta$$
$$V(\boldsymbol{\zeta}) = \Psi$$
$$V(\boldsymbol{\eta}) = \ddot{\mathrm{B}}^{-1}(\Gamma\Phi\Gamma^{\intercal} + \Psi)\ddot{\mathrm{B}}^{\intercal-1}$$

where
$\ddot{\mathrm{B}} = (\mathrm{I} - \mathrm{B})$
$\Phi = COV(\mathrm{x})$

Person-Machine Interface

How do I do this?

SAS, SPSS,
R/S-Plus
STATA



clean data, handle missingness
select cases, variables
transformations
descriptive statistics

A text file with selected data elements. Comma delimited works best for Mplus

Also a raw text (ASCII) file
Instructions for a single analysis

How to Write a Mplus Command File

- Read the Users Manual
  - (at least once all the way through)
- Keep it close to your computer
- Find a similar example
  - (Mplus provides input programs and data for all examples)
- Hack the example to suit your problem

Mplus Commands

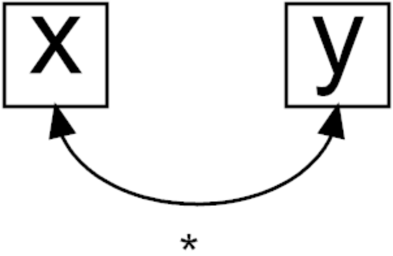
- Title
- Data
- Variable
- Analysis
- Model
- Output
- Define
- Plot
- Savedata
- Montecarlo

How to Write a Mplus Command File

- Read the Users Manual
  - (at least once all the way through)
- Keep it close to your computer
- Find a similar example
  - (Mplus provides input programs and data for all examples)
- Hack the example to suit your problem

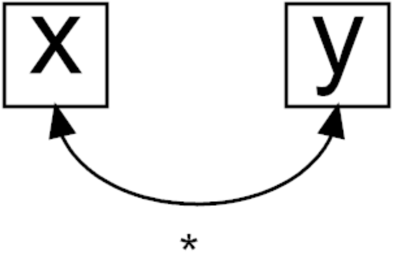
Understanding Mplus Syntax

Correlation



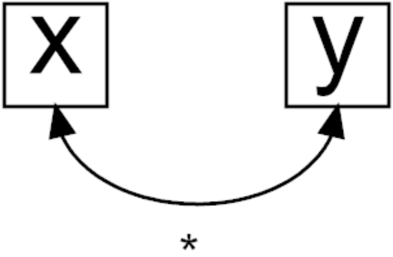

```
MODEL: x with y ;
```

Covariance



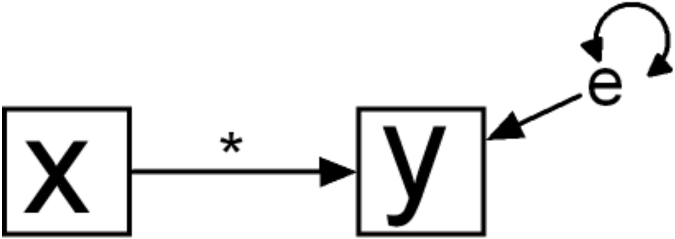

```
TITLE:    A Covariance
DATA:     FILE = temp.dat ;
VARIABLE: NAMES = x y ;
MODEL:    x with y ;
```

Correlation



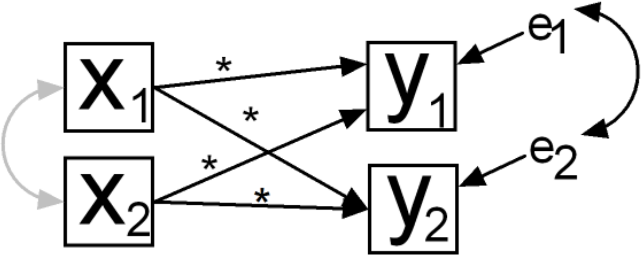

```
TITLE:    A Covariance
DATA:     FILE = temp.dat ;
VARIABLE: NAMES = x y ;
OUTPUT:   STANDARDIZED ;
MODEL:    x with y ;
```

Simple Regression



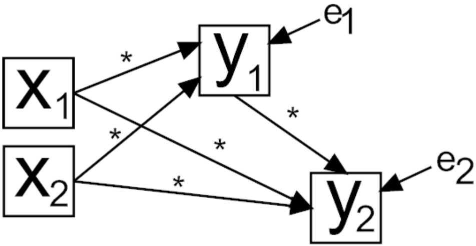

```
MODEL:  y on x * ;
        y *   ;  ! error (residual) term
       [y*]   ;  ! and intercept are
                 ! estimated by default
                 ! so "MODEL: y on x* ;" would be
                 ! sufficient
```

Multivariate Regression



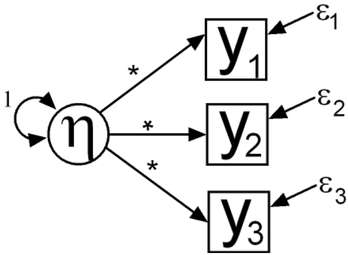

```
MODEL: y1 y2 on x1 * x2 * ;
! default intercepts estimated -> [y1 * y2*] ;
! default resid. var. estimated -> y1 * y2 * ;
! default resid. cov. estimated -> y2 with y1 * ;
! default covariance of x's -> x1 with x2 * ;
```

Path Analysis



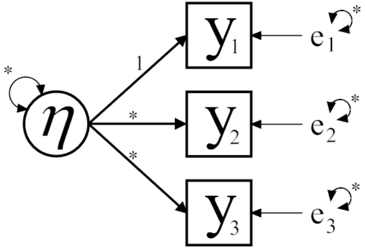

```
MODEL:  y1 y2 on x1 x2 * ;
        y2 on y1 * ;
```

Confirmatory Factor Analysis




```
TITLE:    Confirmatory factor analysis
          Alternative specification
DATA:     FILE = trash.dat ;
VARIABLE: NAMES = y1-y3 ;
OUTPUT:   STANDARDIZED ;
MODEL:    eta by y1-y3*  ;
          eta@1 ;
```

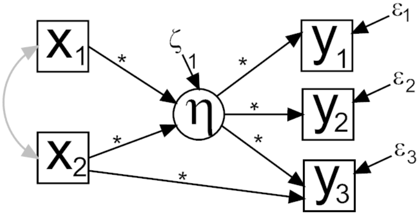
Confirmatory Factor Analysis




```
TITLE:    Confirmatory factor analysis
          Default specification
DATA:     FILE = trash.dat ;
VARIABLE: NAMES = y1-y3 ;
OUTPUT:   STANDARDIZED ;
MODEL:    eta by y1-y3  ; ! NB: no *
```

MIMIC

Multiple Indicators Multiple Causes

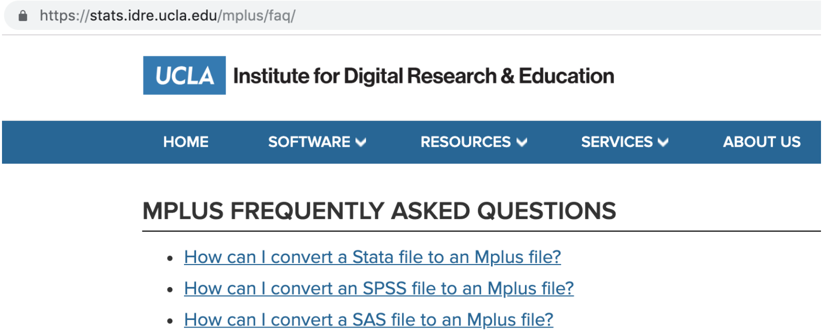



```
TITLE:    MIMIC
DATA:     FILE = trash.dat ;
VARIABLE: NAMES = y1-y3 x1 x2 ;
MODEL:    eta by y1-y3*  ;
          eta@1 ;
          eta on x1-x2 ;  ! indirect effects
          y3 on x2 ;      ! direct effect
```

Getting Data into Mplus

What Mplus Eats

- Individual-level data
- Summary Data
  - correlations, covariances,
  - means, standard deviations
- ASCII Data
  - Raw text
  - Fixed, Free format



https://stats.idre.ucla.edu/mplus/faq/

General Idea

1. Use a general purpose stat package for EDA and data preparation
2. Always use scripts/syntax/programs
3. Understand your data before you put move to Mplus
4. Never overwrite original data
5. Handle missing values
    - Replace with very large negative number (e.g., -9999)
    - Address missing values among INDEPENDENT variables
        - ★ Multiple Imputation (consider Mplus Bayes [Day 3])
        - ★ Declare independent variables dependent variables, assume a distribution, and use ML techniques
6. Output selected variables to ASCII text file (or CSV file)

Converting Microsoft Excel File into raw data file for Mplus

1. Click on "file", select "save as"
2. Under "save as type" select CSV MS-DOS, click save
3. A warning will pop up: click "OK"
    - "The selected file type does not support workbooks..."
4. Then another warning will pop up: click "yes"
    - "*.csv may contain features..."
5. Close out of Excel
6. Find the *.csv file in your directory/file and right click
7. Open with text editor (e.g., Notepad, TextWrangler) and examine
    - First row of variable names and
    - Commas delimiting the variables
8. Copy variable names (for Mplus command file) and delete this line

Stata program

prepare_csv.do

```
use ms q1.dta, clear
summarize
* Missing values in dependent variables (-9999)
foreach var of varlist msq3-msq9 {
      replace `var' = -9999 if inlist(`var',0,1)~=1
}
foreach var of varlist male age educ {
    drop if `var'==. // one way to handle missingness in x's
}
* assume anyone not identified as Black is White
generate white=black~=1
order msq3-msq9 site id male age white educ
keep msq3-msq9 site id male age white educ
outsheet using msq1.csv , nonames noquote nolabel comma
```

Selected lines from data file

msq1.csv

```
2,1,1,1,1,2,3,300065,0,67,1,1
1,1,1,1,1,1,3,300066,0,67,0,4
2,1,1,2,1,2,3,300067,1,87,0,1
1,2,1,1,1,2,3,300068,1,82,0,1
2,1,1,2,1,2,3,300069,1,82,0,2
2,1,1,1,1,2,3,300070,0,77,0,2
2,1,1,2,2,2,3,300071,0,72,0,2
2,1,1,1,1,1,3,300072,1,72,0,4
1,1,1,1,1,1,3,300073,1,67,0,4
2,1,-9999,-9999,2,-9999,3,300074,1,67,0,1
```

SPSS program

prepare_csv.sps

```
get file= 'msq1.sav' .
execute .
compute n = 1 .
desc var all .
do repeat X = male age educ / .
   if missing(X) n=0 .
end repeat .
select if n=1 .
do repeat X = msq3 msq4 msq5 msq6 msq7 msq9 / .
   if missing(X) X=-9999 .
end repeat .
compute white=1 .
if black=1 white=0 .
save outfile ='trash.sav'
   /keep=msq3 msq4 msq5 msq6 msq7 msq9 site id male age white educ .
get file ='trash.sav' .
save translate outfile='msq_from_spss.dat'
   /type=tab /map /replace.
execute .
```

Selected lines from data file

msq_from_spss.dat

2	1	1	1	1	2	3	300065	0	67	1	1
1	1	1	1	1	1	3	300066	0	67	0	4
2	1	1	2	1	2	3	300067	1	87	0	1
1	2	1	1	1	2	3	300068	1	82	0	1
2	1	1	2	1	2	3	300069	1	82	0	2
2	1	1	1	1	2	3	300070	0	77	0	2
2	1	1	2	2	2	3	300071	0	72	0	2
2	1	1	1	1	1	3	300072	1	72	0	4
1	1	1	1	1	1	3	300073	1	67	0	4
2	1	-9999	-9999	2	-9999	3	300074	1	67	0	1

SAS procedure

from http://www.ats.ucla.edu/stat/mplus/faq/sas2mplus.htm

```
data sample2;
set "sample";
array allvars _numeric_ ;
do over allvars;
  if missing(allvars) then allvars = -9999 ;
end;
run;
proc export data=sample2 outfile='sample.dat'
  dbms=dlm replace ;
run;
```

Make your life easy

- Stata

  stata2mplus.ado - format and export data, start an inp. From Stata:

  ```
  ssc install stata2mplus
  ```

  runmplus.ado - format and export data, write inp, run Mplus, display results to Stata results window, save estimation results. Visit http://www.lvmworkshop.org/home/runmplus-stuff

  mplus.ado - similar to runmplus

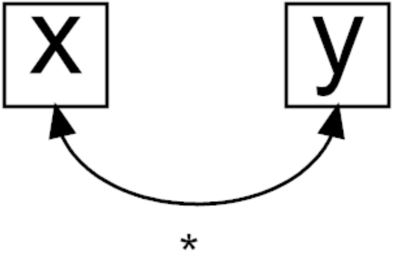
    - `net from http://www.soc-research.net/ado/mplus`

- R
  - MplusAutomation package
- SPSS, SAS
  - Write your own macros

Stata/runmplus and R/MplusAutomation

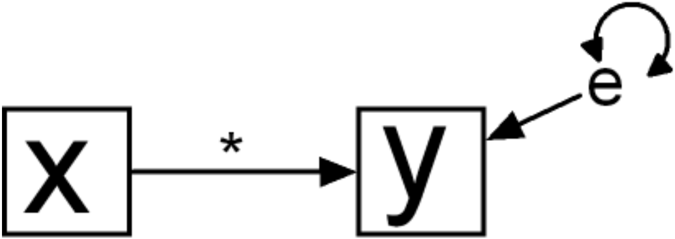
- Integration of Mplus with Stata and R
- Results stored as objects (estimates, fits)
- Facilitates
  - Reproducible data analysis
  - Simulation, bootstrap, multiple imputation
  - Automated model building
  - Learning about models
- Eliminates
  - Copying and pasting of output (and associated errors)
  - Other drudgery (making DAT or CSV files)
  - Clutter (inp, out, and dat, files of unknown origin)
- But,
  - Risk not learning about Mplus software
  - No Mplus plots

Correlation



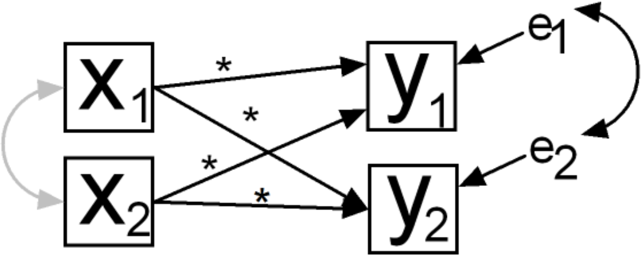

```
runmplus x y , model(x with y;) standardized
```

Simple Regression



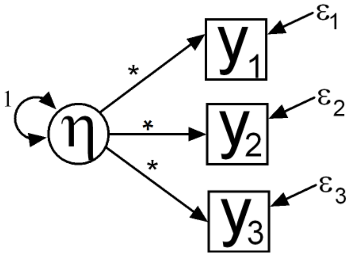

```
runmplus x y , model(y on x *;)
```

Multivariate Regression




```
runmplus y1-y2 x1-x2 , model(y1 y2 on x1 x2; )
```

Confirmatory Factor Analysis




```
runmplus y1-y3 , model(eta by y1-y3* ; eta@1;)
```

R/MplusAutomation

- Run Mplus from R
- Requires separately constructed INP file
- Good parameter and fit statistic extraction
- Nice utilities for SWEAVEing (output to LaTeX)
- Ideally suited for simulations (running lots of INPs sequentially)
- Can be tricked into being a general purpose Mplus launcher

Using Mplus via R

- MplusAutomation R package
- Designed to automate three major aspects of latent variable modeling in Mplus:
  - Creating related groups of models
  - Running Batches
  - Extracting and tabulating model parameters and test statistics
- http://statmodel.com/usingmplusviar.shtml
- http://cran.r-project.org/web/packages/MplusAutomation/index.html
- Written by Michael.Hallquist@gmail.com

Hallquist, M. N., & Wiley, J. F. (2018). Mplus Automation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling*, 25(4), 621-638.

Step 1: Install R Packages

```
# Step 1
# Select your CRAN Mirror
# done already-> chooseCRANmirror()
# Install the MplusviaR package
# done already-> install.packages("MplusAutomation", .Library)
library(MplusAutomation)
```

Step 2: Set your work directory

```
# Step 2
# Set working directory
# This is where your data and input files should be
setwd("C:/work/shortcourse/posted/analysis/learnMPA")
```

Step 3: Getting Data into R

```
# Step 3
# Input data from a .dat file into R
EPESE_data <- read.table("ex01-02.dat",
   header=FALSE, sep=",")

# Declare column names
colnames(EPESE_data) <- c("sad","blues","depress",
   "happy","enjoy", "hopeful","age","male")

# Attach dataframe (allows you to call variables by name)
attach(EPESE_data)
```

Optional Step 4: Create your variables

```
# Step 4
# Create a binary variable for age, which equals
# 0 when a participant’s age is less than the mean
# age in the sample,and 1 if it is more
age_mean <- mean(age)
age_high <- ifelse(age < age_mean, 0, 1)

# Exchange binary age variable for continuous
# in EPESE dataset
EPESE_data_cat <-data.frame(sad,blues,depress,
   happy,enjoy,hopeful,age_high,male)

# Export data from R to Mplus-friendly format
prepareMplusData(EPESE_data_cat,
   "epesecfa.dat")
```

Step 5: Prepare your input file for Mplus

```
# Step 5
# Prepare Mplus INP file
cfa <- mplusObject(
  TITLE = "EX01-02 - CFA with covariates EPESE CESD;",
  VARIABLE = "CATEGORICAL  = sad blues depress
     happy enjoy hopeful;",
  MODEL = "f by sad-hopeful;
     f on age_high male ;",
  OUTPUT = "STDYX;
     MODINDICES (ALL -0);",
  usevariables = c("sad","blues","depress",
                   "happy","enjoy", "hopeful",
                   "age_high","male"),
  rdata = EPESE_data_cat)
```

Step 6: Running the Model via Mplus

```
result <- mplusModeler(cfa,
  "EPESE_data_cat.dat",
  modelout = "ex0102.inp",
  run = 1L)
```

Step 7a: Show output from Mplus in R

```
# Step 7a
# Show all output
model_output <-readModels(
   "C:/work/shortcourse/posted/analysis/learnMPA")
model_output

# Extract summary statistics
summaryStats <- extractModelSummaries(
   "C:/work/shortcourse/posted/analysis/learnMPA")
summaryStats
```

Step 7b: Show table of output from Mplus in R

```
# Step 7b
# Show relevant summary statistics in a table
summary_table <- cbind(summaryStats$Observations,
                       summaryStats$ChiSqM_Value,
                       summaryStats$ChiSqM_DF,
                       summaryStats$ChiSqM_PValue,
                       summaryStats$CFI,
                       summaryStats$TLI,
                       summaryStats$RMSEA_Estimate)

colnames(summary_table) <- cbind(
  "N", "Chi-square", "(DF)", "(p)", "CFI", "TLI", "RMSEA")
summary_table
```

Step 7c: Show table of other output from Mplus in R

```
# Step 7c
# Show modification indices
mod_indices <-extractModIndices(
   "C:/work/shortcourse/posted/analysis/learnMPA")
mod_indices

# Show parameter estimates
model_results <-extractModelParameters(
   "C:/work/shortcourse/posted/analysis/learnMPA")
model_results
```

R lavaan Package

lavaan

- R package for structural equation modeling and more
- Yves Rosseel, PhD, Ghent University, Belgium
- http://lavaan.ugent.be/
- Replicates Mplus estimators and output
- Supports some models and features (categorical data, sampling weights) but not all (multilevel, mixture, Bayesian) that Mplus does.
- GNU General Public License

lavaan *vs* Stata/runmplus and R/MplusAutomation

- Stata/runmplus and R/MplusAutomation require Mplus
- lavaan is its own software
  - Have to learn new model syntax in lavaan
  - Have to learn something new regardless

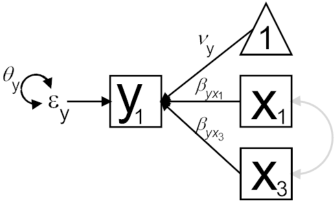
lavaan *vs.* Mplus

M*plus*		R/*lavaan*

TITLE:	Multivariable regression	# Multivariable Regression
		library(lavaan)
		setwd("c:/*whereverdatais* ")

DATA:	FILE = ex3.1.dat ;	data <- read.csv("ex3.1.dat", head=FALSE)
VARIABLE:	NAMES = y1 x1 x3 ;	names(data) <- c("y1","x1","x3")
MODEL:	y1 on x1 x3 ;	fit <- sem('y1 ~ 1 + x1 + x3', std.lv=TRUE, data=data)
OUTPUT:	STANDARDIZED ;	standardizedsolution(fit)

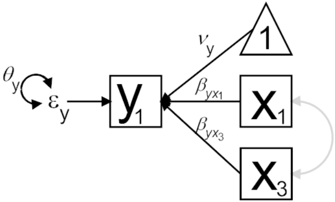
Mplus Users Guide Ex 3.1: Multivariable Regression



Mplus Users Guide Example 3.1

```
TITLE:    Multivariable regression
DATA:     FILE = ex3.1.dat ;
VARIABLE: NAMES = y1 x1 x3 ;
MODEL:    y1 on x1 x3 ;
OUTPUT:   STANDARDIZED ;
```

Mplus Users Guide Ex 3.1: Multivariable Regression (lavaan)



Mplus Users Guide Example 3.1

```
# Mplus Users Guide Ex 3.1: Multivariable Regression
library(lavaan)
setwd("c:/work/shortcourse/private/rnj")
data <- read.csv("ex3.1.dat",head=FALSE)
names(data) <- c("y1","x1","x3")
fit <- sem('y1 ~ 1 + x1 + x3', std.lv=TRUE, data=data )
summary(fit, standardized=TRUE, fit.measures = TRUE)
standardizedsolution(fit)
```

Extra credit

Astute observers will note that Mplus and lavaan provide different loglikelihood and AIC values. This is because Mplus and lavaan allow different parts of the model into the likelihood function. To get Mplus to produce the same loglikelihood, use this model statement:

```
MODEL:
  y1s by y1@1 ;
  x1s by x1@1 ;
  x3s by x3@1 ;
  y1@0 ;
  x1@0 ;
  x3@0 ;
  y1s on x1s x3s ;
  x1s with x3s ;
```

Refer back to the all-y model case (slide 85) to understand why.

Questions

Footnote on Data Management for LDA

Some approaches to LDA require data in "wide" format, others "long" format

LONG
Vertical
Multiple record

id	time	y1	x1
1	1	y11	x11
1	2	y12	x12
1	3	y13	x13
2	1	y21	x21
2	2	y22	x22
2	3	y23	x23
3	1	y31	x31
...			
n	p	ynp	xnp

WIDE
Horizontal
Single record

id	y1	y2	y3	x1	x2	x3
1	y11	y12	y13	x11	x12	x13
2	y21	y22	y23	x21	x22	x23
...						
n	yn1	yn2	yn3	xn1	xn2	xn3

Most Mplus LDA models require WIDE format

Skills Challenge

- Invest time learning how to move from wide to long easily with your favorite general purpose statistics package
- Why
  - Wide
    - Mplus LDA models want wide
  - Long
    - Storage and data management more efficient
    - Better for exploratory graphics
    - Enables checking your growth models with mixed effects or HLM models

Resources

- UCLA IDRE Learning Modules on reshaping data
  - Stata https://stats.idre.ucla.edu/stata/faq/how-can-i-reshape-doubly-or-triply-wide-data-to-long/
  - SAS https://stats.idre.ucla.edu/sas/modules/how-to-reshape-data-long-to-wide-using-proc-transpose/
  - SPSS https://stats.idre.ucla.edu/spss/modules/reshaping-data-wide-to-long/
  - R https://stats.idre.ucla.edu/r/faq/how-can-i-reshape-my-data-in-r/
- Other Resources
  - R (S-plus) https://stat.ethz.ch/pipermail/r-help/2009-February/189684.html
- Mplus will do it with built-in functionality
  - But it is not the right tool for data management

Questions

Examples

Latent growth curve modeling is a highly constrained form of confirmatory factor analysis (CFA). The goal of this session is to provide examples of specifying and estimating a CFA model in Mplus, as a basic Mplus skills-building exercise.

Time permitting, we will look at growth modeling examples from Singer and Willet (2003).

Example Data Set

Established Populations for the Epidemiologic Study of the Elderly

EPESE

- N=14,456
- General population survey of older adults
- Carried out at four US sites:
  - East Boston, MA (N=3,809)
  - Iowa and Washington Counties, IA (N=3,673)
  - New Haven, CT(N=2,812)
  - Durham-Piedmont area of NC (N=4,162)
- Field work began 1981, 1985 in NC
- Data are in the public domain (ICPSR)
- Results we derive have not been published

age	Freq.	Percent	Cum.
67: 65-69	4,457	30.83	30.83
72: 70-74	3,970	27.46	58.29
77: 75-79	2,872	19.87	78.16
82: 80-84	1,878	12.99	91.15
87: 85+	1,279	8.85	100.00
Total	14,456	100.00	

male	Freq.	Percent	Cum.
0	8,960	61.98	61.98
1	5,496	38.02	100.00
Total	14,456	100.00	

Black or African-American

site	0	1	.	Total
1 EB-MA	0	0	3,809	3,809
2 IW-IO	0	0	3,673	3,673
3 NH-CT	2,283	529	0	2,812
4 DP-NC	1,901	2,261	0	4,162
Total	4,184	2,790	7,482	14,456

educ	Freq.	Percent	Cum.
1 <8	4,056	28.06	28.06
2 8-11	5,686	39.33	67.39
3 12	2,552	17.65	85.04
4 >12	1,863	12.89	97.93
.	299	2.07	100.00
Total	14,456	100.00	

Cognition

Short Portable Mental Status Questionnaire

- Conventionally, 10 items
- We will use six items from EPESE
  - msq3 what is date
  - msq4 what is the day of week
  - msq5 who is the president
  - msq6 who was the prior president
  - msq7 what is your Mothers maiden name
  - msq9 Count back from 20 by 3s

EPESE
Mental Status Questionnaire
Wave 1 Item Response Frequency

item	Response: Error 0	NoErr 1	miss .	Total
msq3 what is the date...	3907	9788	761	14456
msq4 day of week........	772	12916	768	14456
msq5 who is president...	1239	12436	781	14456
msq6 who was prior pres.	4302	9355	799	14456
msq7 what is M Maiden...	539	13075	842	14456
msq9 back from 20 by 3..	5194	7768	1494	14456

Depression

- CES-D Centers for Epidemiologic Studies Depression Scale
- We will use six items from New Haven site, which ask if *During the past week* ...

SAD: …I felt sad
BLUES: …I felt I could not shake off the blues
DEPRESS: ... I felt depressed

0: Rarely or None of the time
1: Some of the time
2: Much of the time
3: Most or all of the time

HAPPY: ...I was happy
ENJOY: ...I enjoyed life
HOPEFUL: ...I felt hopeful about the future

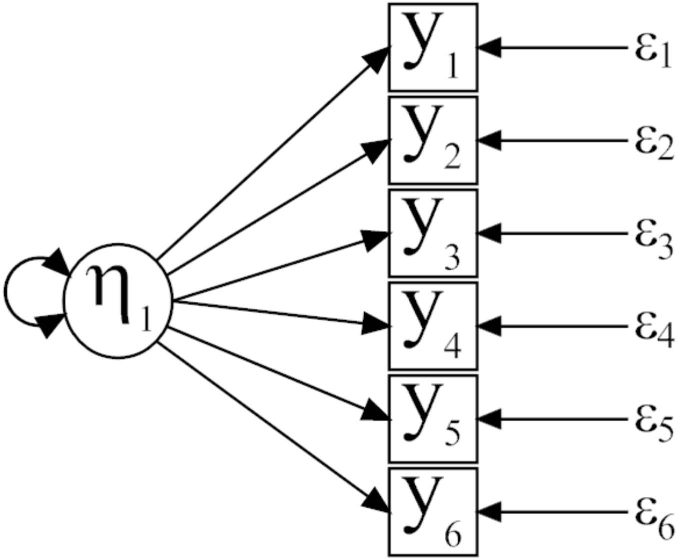
0: Most or all of the time
1: Much of the time
2: Some of the time
3: Rarely or None of the time

Item Response Frequency
(entries are percentages)

item	Response 0	1	2	3	N
01 sad......	66	27	4	4	2442
02 blues....	78	15	3	4	2442
03 depress..	64	26	4	6	2442
04 happy....	67	10	16	7	2442
05 enjoy....	76	8	11	6	2442
06 hopeful..	58	7	16	20	2442

Confirmatory Factor Analysis Example [ex0101.inp]

Confirmatory Factor Analysis
New Haven EPESE
Six-item CES-D
Continuous Factor Indicators


η1
y1
y2
y3
y4
y5
y6
ε1
ε2
ε3
ε4
ε5
ε6

```
Mplus VERSION 8.8 DEMO (Mac)
MUTHEN & MUTHEN
07/11/2022   2:07 PM

INPUT INSTRUCTIONS

  TITLE:    EX0101 CFA EPESE CESD ;

  DATA:     FILE = ex0101.dat ;

  VARIABLE: NAMES =  sad blues depress happy enjoy hopeful age male ;
            USEVARIABLES = sad blues depress happy enjoy hopeful ;
            MISSING ARE ALL (-9999) ;

  OUTPUT:   STANDARDIZED TECH1 ;

  MODEL:    f1 by sad-hopeful;




INPUT READING TERMINATED NORMALLY



EX0101 CFA EPESE CESD ;

SUMMARY OF ANALYSIS

Number of groups                                                 1
Number of observations                                        2442

```

```
Number of dependent variables                                      6
Number of independent variables                                    0
Number of continuous latent variables                              1

Observed dependent variables

  Continuous
   SAD          BLUES        DEPRESS     HAPPY        ENJOY        HOPEFUL

Continuous latent variables
   F1


Estimator                                                         ML
Information matrix                                          OBSERVED
Maximum number of iterations                                    1000
Convergence criterion                                      0.500D-04
Maximum number of steepest descent iterations                     20
Maximum number of iterations for H1                             2000
Convergence criterion for H1                               0.100D-03

Input data file(s)
  ex0101.dat

Input data format  FREE


SUMMARY OF DATA

     Number of missing data patterns            1


```

```
COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value   0.100


     PROPORTION OF DATA PRESENT


           Covariance Coverage
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
 SAD            1.000
 BLUES          1.000         1.000
 DEPRESS        1.000         1.000         1.000
 HAPPY          1.000         1.000         1.000         1.000
 ENJOY          1.000         1.000         1.000         1.000         1.000
 HOPEFUL        1.000         1.000         1.000         1.000         1.000


           Covariance Coverage
              HOPEFUL
              ________
 HOPEFUL        1.000



UNIVARIATE SAMPLE STATISTICS


     UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS

```

```
          Variable/         Mean/      Skewness/   Minimum/ % with
               Percentiles
         Sample Size      Variance    Kurtosis    Maximum  Min/Max        20%/60%     40%/
              80%     Median

      SAD                  0.454       1.831       0.000   65.77%        0.000
           0.000       0.000
             2442.000      0.548       3.185       3.000    3.81%        0.000
                  1.000
      BLUES                0.337       2.417       0.000   78.05%        0.000
           0.000       0.000
             2442.000      0.545       5.259       3.000    4.34%        0.000
                  1.000
      DEPRESS              0.510       1.749       0.000   64.41%        0.000
           0.000       0.000
             2442.000      0.674       2.457       3.000    5.81%        0.000
                  1.000
      HAPPY                0.630       1.241       0.000   67.08%        0.000
           0.000       0.000
             2442.000      0.979       0.069       3.000    7.25%        0.000
                  2.000
      ENJOY                0.465       1.742       0.000   75.96%        0.000
           0.000       0.000
             2442.000      0.816       1.624       3.000    5.94%        0.000
                  1.000
      HOPEFUL              0.973       0.681       0.000   57.70%        0.000
           0.000       0.000
             2442.000      1.517      -1.263       3.000   19.57%        1.000
                  2.000


THE MODEL ESTIMATION TERMINATED NORMALLY
```

```



MODEL FIT INFORMATION

Number of Free Parameters                         18

Loglikelihood

          H0 Value                        -17278.438
          H1 Value                        -17036.545

Information Criteria

          Akaike (AIC)                     34592.876
          Bayesian (BIC)                   34697.286
          Sample-Size Adjusted BIC         34640.096
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              483.785
          Degrees of Freedom                       9
          P-Value                             0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                             0.147
          90 Percent C.I.                      0.136  0.158
          Probability RMSEA <= .05             0.000

CFI/TLI
```

```

          CFI                                0.883
          TLI                                0.805

Chi-Square Test of Model Fit for the Baseline Model

          Value                           4064.686
          Degrees of Freedom                    15
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.057



MODEL RESULTS

                                                   Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F1       BY
    SAD                1.000      0.000    999.000    999.000
    BLUES              0.963      0.035     27.751      0.000
    DEPRESS            1.193      0.040     30.075      0.000
    HAPPY              1.230      0.049     24.975      0.000
    ENJOY              1.075      0.045     23.950      0.000
    HOPEFUL            0.883      0.057     15.489      0.000

 Intercepts
    SAD                0.454      0.015     30.325      0.000
    BLUES              0.337      0.015     22.553      0.000
```

```
    DEPRESS              0.510      0.017     30.685      0.000
    HAPPY                0.630      0.020     31.451      0.000
    ENJOY                0.465      0.018     25.421      0.000
    HOPEFUL              0.973      0.025     39.018      0.000

 Variances
    F1                   0.260      0.015     17.461      0.000

 Residual Variances
    SAD                  0.288      0.011     27.263      0.000
    BLUES                0.304      0.011     27.677      0.000
    DEPRESS              0.304      0.013     23.883      0.000
    HAPPY                0.586      0.021     27.902      0.000
    ENJOY                0.516      0.018     28.700      0.000
    HOPEFUL              1.315      0.039     33.536      0.000


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix              0.141E-01
       (ratio of smallest to largest eigenvalue)


STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F1       BY
```

```
    SAD                0.689      0.014     50.000      0.000
    BLUES              0.665      0.015     45.558      0.000
    DEPRESS            0.741      0.013     56.771      0.000
    HAPPY              0.634      0.016     40.045      0.000
    ENJOY              0.606      0.016     36.810      0.000
    HOPEFUL            0.366      0.020     18.142      0.000

 Intercepts
    SAD                0.614      0.022     27.819      0.000
    BLUES              0.456      0.021     21.463      0.000
    DEPRESS            0.621      0.022     28.096      0.000
    HAPPY              0.636      0.022     28.680      0.000
    ENJOY              0.514      0.022     23.890      0.000
    HOPEFUL            0.790      0.023     34.068      0.000

 Variances
    F1                 1.000      0.000    999.000    999.000

 Residual Variances
    SAD                0.526      0.019     27.698      0.000
    BLUES              0.558      0.019     28.739      0.000
    DEPRESS            0.451      0.019     23.359      0.000
    HAPPY              0.598      0.020     29.838      0.000
    ENJOY              0.632      0.020     31.669      0.000
    HOPEFUL            0.866      0.015     58.825      0.000


STDY Standardization


                                                Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

```

```
 F1       BY
    SAD                   0.689      0.014     50.000      0.000
    BLUES                 0.665      0.015     45.558      0.000
    DEPRESS               0.741      0.013     56.771      0.000
    HAPPY                 0.634      0.016     40.045      0.000
    ENJOY                 0.606      0.016     36.810      0.000
    HOPEFUL               0.366      0.020     18.142      0.000

 Intercepts
    SAD                   0.614      0.022     27.819      0.000
    BLUES                 0.456      0.021     21.463      0.000
    DEPRESS               0.621      0.022     28.096      0.000
    HAPPY                 0.636      0.022     28.680      0.000
    ENJOY                 0.514      0.022     23.890      0.000
    HOPEFUL               0.790      0.023     34.068      0.000

 Variances
    F1                    1.000      0.000    999.000    999.000

 Residual Variances
    SAD                   0.526      0.019     27.698      0.000
    BLUES                 0.558      0.019     28.739      0.000
    DEPRESS               0.451      0.019     23.359      0.000
    HAPPY                 0.598      0.020     29.838      0.000
    ENJOY                 0.632      0.020     31.669      0.000
    HOPEFUL               0.866      0.015     58.825      0.000


STD Standardization

                                                  Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
```

```

 F1        BY
    SAD                  0.510       0.015      34.922       0.000
    BLUES                0.491       0.015      33.059       0.000
    DEPRESS              0.608       0.016      37.799       0.000
    HAPPY                0.627       0.021      30.522       0.000
    ENJOY                0.548       0.019      28.846       0.000
    HOPEFUL              0.450       0.027      16.672       0.000

 Intercepts
    SAD                  0.454       0.015      30.325       0.000
    BLUES                0.337       0.015      22.553       0.000
    DEPRESS              0.510       0.017      30.685       0.000
    HAPPY                0.630       0.020      31.451       0.000
    ENJOY                0.465       0.018      25.421       0.000
    HOPEFUL              0.973       0.025      39.018       0.000

 Variances
    F1                   1.000       0.000     999.000     999.000

 Residual Variances
    SAD                  0.288       0.011      27.263       0.000
    BLUES                0.304       0.011      27.677       0.000
    DEPRESS              0.304       0.013      23.883       0.000
    HAPPY                0.586       0.021      27.902       0.000
    ENJOY                0.516       0.018      28.700       0.000
    HOPEFUL              1.315       0.039      33.536       0.000


R-SQUARE

    Observed                                        Two-Tailed
```

```
    Variable         Estimate       S.E.  Est./S.E.    P-Value

    SAD                 0.474      0.019     25.000      0.000
    BLUES               0.442      0.019     22.779      0.000
    DEPRESS             0.549      0.019     28.385      0.000
    HAPPY               0.402      0.020     20.022      0.000
    ENJOY               0.368      0.020     18.405      0.000
    HOPEFUL             0.134      0.015      9.071      0.000


TECHNICAL 1 OUTPUT


     PARAMETER SPECIFICATION


           NU
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
                  1             2             3             4             5


           NU
              HOPEFUL
              ________
                  6


           LAMBDA
              F1
              ________
 SAD               0
```

```
 BLUES              7
 DEPRESS            8
 HAPPY              9
 ENJOY             10
 HOPEFUL           11


           THETA
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              --------      --------      --------      --------      --------
 SAD               12
 BLUES              0            13
 DEPRESS            0             0            14
 HAPPY              0             0             0            15
 ENJOY              0             0             0             0            16
 HOPEFUL            0             0             0             0             0


           THETA
              HOPEFUL
              --------
 HOPEFUL           17


           ALPHA
              F1
              --------
                  0


           BETA
              F1
```

```
                 --------
F1                      0


          PSI
             F1
             --------
F1                     18


    STARTING VALUES


          NU
             SAD           BLUES         DEPRESS       HAPPY         ENJOY
             --------      --------      --------      --------      --------
               0.454         0.337         0.510         0.630         0.465


          NU
             HOPEFUL
             --------
               0.973


          LAMBDA
             F1
             --------
SAD            1.000
BLUES          0.991
DEPRESS        1.218
HAPPY          1.176
```

```
ENJOY            1.031
HOPEFUL          0.847


          THETA
             SAD           BLUES         DEPRESS       HAPPY         ENJOY
             --------      --------      --------      --------      --------
SAD            0.274
BLUES          0.000         0.273
DEPRESS        0.000         0.000         0.337
HAPPY          0.000         0.000         0.000         0.490
ENJOY          0.000         0.000         0.000         0.000         0.408
HOPEFUL        0.000         0.000         0.000         0.000         0.000


          THETA
             HOPEFUL
             --------
HOPEFUL        0.759


          ALPHA
             F1
             --------
               0.000


          BETA
             F1
             --------
F1             0.000

```

```

           PSI
              F1
              --------
 F1             0.050


     Beginning Time:  14:07:22
        Ending Time:  14:07:22
       Elapsed Time:  00:00:00


Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
  Maximum number of dependent variables: 6
  Maximum number of independent variables: 2
  Maximum number of between variables: 2
  Maximum number of continuous latent variables in time series analysis: 2


MUTHEN & MUTHEN
3463 Stoner Ave.
Los Angeles, CA  90066

Tel: (310) 391-9971
Fax: (310) 391-8971
Web: www.StatModel.com
Support: Support@StatModel.com

Copyright (c) 1998-2022 Muthen & Muthen
```

ex0101.out

A Note Model Fit

- Model fit is based on how close the model-implied covariance matrix is to the observed covariance matrix
- Chi-Square should be low, P-value high
  - an *absolute fit* index
  - very sensitive to sample size
- CFI > 0.95
  - a *parsimony-adjusted relative fit* index
  - CFI ranges from 0 to 1
  - Bentler. Psych Bull, 1990; 107:238-46.
- RMSEA $< .05$
  - an *absolute fit* index
  - Hu & Bentler. Psych Methods, 1998; 4:424-53.

A note on Mplus Standardized Estimates

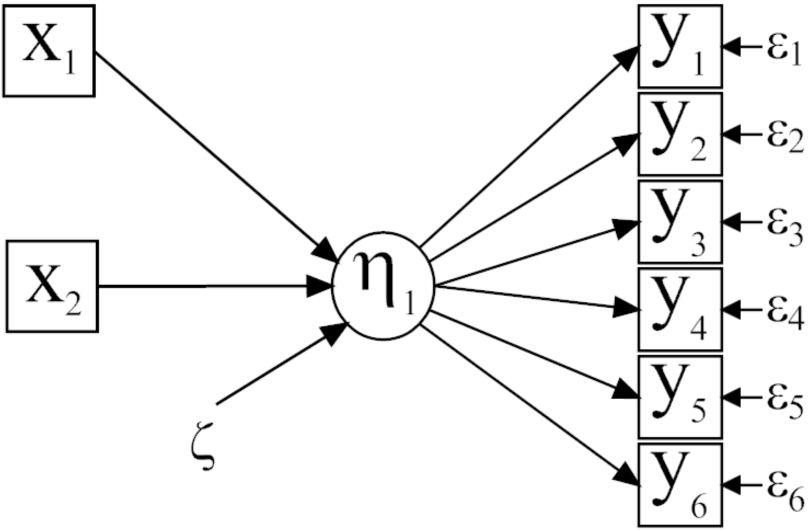
- Standardized with respect to...
  - **STDYX** ...the variances of the continuous latent variables, background, and outcome variables. Use this if the background variable is continuous (and it otherwise makes sense to do so).
  - **STDY** ...the variances of the continuous latent variables and the outcome variables. Use this one if the background variable is discrete.
  - **STD** ... the variances of the continuous latent variables.
- Standard errors and significance testing
  - Each standardization solution may come with standard errors, z-test, P-value (e.g., STDY not available for categorical dependent variables)
  - Significance tests don't always agree with unstandardized estimates
  - I recommend basing inferences on unstandardized estimates

Example 1.2

CFA with Covariates
CES-D Depression, Age, Sex
New Haven EPESE

CFA with Covariates Example 1.2 ex01-02.inp

MIMIC - Multiple Indicators Multiple Causes EPESE CES-D (six items)
Age, Gender


X_1
X_2
ζ
η_1
y_1
y_2
y_3
y_4
y_5
y_6
ε_1
ε_2
ε_3
ε_4
ε_5
ε_6

```
Mplus VERSION 8.8 DEMO (Mac)
MUTHEN & MUTHEN
07/11/2022   2:07 PM

INPUT INSTRUCTIONS

  TITLE:    EX0102 CFA WITH COVARIATES EPESE CESD ;

  DATA:     FILE = ex0101.dat ;

  VARIABLE: NAMES =  sad blues depress happy enjoy hopeful age male ;
            MISSING ARE ALL (-9999) ;

  DEFINE:   CENTER age male (grandmean) ;

  OUTPUT:   TECH1 ; standardized ;

  MODEL:    f by sad-hopeful ;
            f on age male ;






INPUT READING TERMINATED NORMALLY



EX0102 CFA WITH COVARIATES EPESE CESD ;

SUMMARY OF ANALYSIS
```

```

Number of groups                                                 1
Number of observations                                        2442

Number of dependent variables                                    6
Number of independent variables                                  2
Number of continuous latent variables                            1

Observed dependent variables

  Continuous
   SAD         BLUES       DEPRESS     HAPPY       ENJOY       HOPEFUL

Observed independent variables
   AGE         MALE

Continuous latent variables
   F

Variables with special functions

  Centering (GRANDMEAN)
   AGE         MALE


Estimator                                                       ML
Information matrix                                        OBSERVED
Maximum number of iterations                                  1000
Convergence criterion                                    0.500D-04
Maximum number of steepest descent iterations                   20
Maximum number of iterations for H1                           2000
Convergence criterion for H1                             0.100D-03
```

```

Input data file(s)
  ex0101.dat

Input data format  FREE


SUMMARY OF DATA

     Number of missing data patterns             1


COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value   0.100


     PROPORTION OF DATA PRESENT


           Covariance Coverage
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
 SAD            1.000
 BLUES          1.000         1.000
 DEPRESS        1.000         1.000         1.000
 HAPPY          1.000         1.000         1.000         1.000
 ENJOY          1.000         1.000         1.000         1.000         1.000
 HOPEFUL        1.000         1.000         1.000         1.000         1.000
 AGE            1.000         1.000         1.000         1.000         1.000
 MALE           1.000         1.000         1.000         1.000         1.000

```

```

           Covariance Coverage
              HOPEFUL       AGE           MALE
              ________      ________      ________
 HOPEFUL         1.000
 AGE             1.000         1.000
 MALE            1.000         1.000         1.000



UNIVARIATE SAMPLE STATISTICS


     UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS

         Variable/         Mean/     Skewness/   Minimum/ % with
                Percentiles
        Sample Size      Variance    Kurtosis    Maximum  Min/Max      20%/60%    40%/
             80%    Median

     SAD                   0.454       1.831       0.000   65.77%       0.000
          0.000       0.000
            2442.000       0.548       3.185       3.000    3.81%       0.000
                 1.000
     BLUES                 0.337       2.417       0.000   78.05%       0.000
          0.000       0.000
            2442.000       0.545       5.259       3.000    4.34%       0.000
                 1.000
     DEPRESS               0.510       1.749       0.000   64.41%       0.000
          0.000       0.000
            2442.000       0.674       2.457       3.000    5.81%       0.000
                 1.000
```

```
     HAPPY                 0.630        1.241        0.000   67.08%        0.000
          0.000       0.000
             2442.000      0.979        0.069        3.000    7.25%        0.000
               2.000
     ENJOY                 0.465        1.742        0.000   75.96%        0.000
          0.000       0.000
             2442.000      0.816        1.624        3.000    5.94%        0.000
               1.000
     HOPEFUL               0.973        0.681        0.000   57.70%        0.000
          0.000       0.000
             2442.000      1.517       -1.263        3.000   19.57%        1.000
               2.000
     AGE                   0.000        0.497       -7.316   28.50%       -7.316
         -2.316      -2.316
             2442.000     40.696       -0.862       12.684    8.39%        2.684
               7.684
     MALE                  0.000        0.307       -0.424   57.58%       -0.424
         -0.424      -0.424
             2442.000      0.244       -1.906        0.576   42.42%        0.576
               0.576


THE MODEL ESTIMATION TERMINATED NORMALLY



MODEL FIT INFORMATION

Number of Free Parameters                        20

Loglikelihood

```

```
          H0 Value                        -17255.177
          H1 Value                        -17006.684

Information Criteria

          Akaike (AIC)                     34550.354
          Bayesian (BIC)                   34666.365
          Sample-Size Adjusted BIC         34602.820
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              496.985
          Degrees of Freedom                      19
          P-Value                             0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                             0.101
          90 Percent C.I.                      0.094  0.109
          Probability RMSEA <= .05             0.000

CFI/TLI

          CFI                                  0.883
          TLI                                  0.834

Chi-Square Test of Model Fit for the Baseline Model

          Value                             4124.409
          Degrees of Freedom                      27
          P-Value                             0.0000
```

```
SRMR (Standardized Root Mean Square Residual)

          Value                                  0.048



MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F        BY
    SAD                1.000      0.000    999.000    999.000
    BLUES              0.963      0.035     27.784      0.000
    DEPRESS            1.190      0.040     30.103      0.000
    HAPPY              1.231      0.049     25.035      0.000
    ENJOY              1.075      0.045     24.007      0.000
    HOPEFUL            0.884      0.057     15.518      0.000

 F        ON
    AGE                0.006      0.002      3.234      0.001
    MALE              -0.134      0.023     -5.778      0.000

 Intercepts
    SAD                0.454      0.015     30.494      0.000
    BLUES              0.337      0.015     22.670      0.000
    DEPRESS            0.510      0.017     30.881      0.000
    HAPPY              0.630      0.020     31.599      0.000
    ENJOY              0.465      0.018     25.530      0.000
    HOPEFUL            0.973      0.025     39.078      0.000
```

```
 Residual Variances
    SAD                0.288      0.011     27.278      0.000
    BLUES              0.304      0.011     27.713      0.000
    DEPRESS            0.306      0.013     24.034      0.000
    HAPPY              0.585      0.021     27.942      0.000
    ENJOY              0.516      0.018     28.740      0.000
    HOPEFUL            1.314      0.039     33.541      0.000
    F                  0.254      0.015     17.456      0.000


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix              0.568E-03
       (ratio of smallest to largest eigenvalue)


STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                   Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F        BY
    SAD                0.689      0.014     50.114      0.000
    BLUES              0.665      0.015     45.608      0.000
    DEPRESS            0.739      0.013     56.729      0.000
    HAPPY              0.634      0.016     40.219      0.000
    ENJOY              0.607      0.016     36.956      0.000
    HOPEFUL            0.366      0.020     18.183      0.000

```

```
 F          ON
    AGE                  0.072      0.022      3.249      0.001
    MALE                -0.130      0.022     -5.880      0.000

 Intercepts
    SAD                  0.614      0.022     27.949      0.000
    BLUES                0.456      0.021     21.564      0.000
    DEPRESS              0.621      0.022     28.247      0.000
    HAPPY                0.636      0.022     28.793      0.000
    ENJOY                0.514      0.021     23.980      0.000
    HOPEFUL              0.790      0.023     34.108      0.000

 Residual Variances
    SAD                  0.525      0.019     27.705      0.000
    BLUES                0.558      0.019     28.782      0.000
    DEPRESS              0.453      0.019     23.521      0.000
    HAPPY                0.598      0.020     29.868      0.000
    ENJOY                0.632      0.020     31.705      0.000
    HOPEFUL              0.866      0.015     58.794      0.000
    F                    0.977      0.007    145.590      0.000


STDY Standardization

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F          BY
    SAD                  0.689      0.014     50.114      0.000
    BLUES                0.665      0.015     45.608      0.000
    DEPRESS              0.739      0.013     56.729      0.000
    HAPPY                0.634      0.016     40.219      0.000
```

```
    ENJOY               0.607      0.016     36.956      0.000
    HOPEFUL             0.366      0.020     18.183      0.000

 F          ON
    AGE                 0.011      0.003      3.252      0.001
    MALE               -0.263      0.045     -5.900      0.000

 Intercepts
    SAD                 0.614      0.022     27.949      0.000
    BLUES               0.456      0.021     21.564      0.000
    DEPRESS             0.621      0.022     28.247      0.000
    HAPPY               0.636      0.022     28.793      0.000
    ENJOY               0.514      0.021     23.980      0.000
    HOPEFUL             0.790      0.023     34.108      0.000

 Residual Variances
    SAD                 0.525      0.019     27.705      0.000
    BLUES               0.558      0.019     28.782      0.000
    DEPRESS             0.453      0.019     23.521      0.000
    HAPPY               0.598      0.020     29.868      0.000
    ENJOY               0.632      0.020     31.705      0.000
    HOPEFUL             0.866      0.015     58.794      0.000
    F                   0.977      0.007    145.590      0.000


STD Standardization

                                                        Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F          BY
    SAD                 0.510      0.015     34.965      0.000
```

```
    BLUES              0.491      0.015     33.077      0.000
    DEPRESS            0.607      0.016     37.766      0.000
    HAPPY              0.628      0.021     30.604      0.000
    ENJOY              0.548      0.019     28.920      0.000
    HOPEFUL            0.451      0.027     16.703      0.000

 F        ON
    AGE                0.011      0.003      3.252      0.001
    MALE              -0.263      0.045     -5.900      0.000

 Intercepts
    SAD                0.454      0.015     30.494      0.000
    BLUES              0.337      0.015     22.670      0.000
    DEPRESS            0.510      0.017     30.881      0.000
    HAPPY              0.630      0.020     31.599      0.000
    ENJOY              0.465      0.018     25.530      0.000
    HOPEFUL            0.973      0.025     39.078      0.000

 Residual Variances
    SAD                0.288      0.011     27.278      0.000
    BLUES              0.304      0.011     27.713      0.000
    DEPRESS            0.306      0.013     24.034      0.000
    HAPPY              0.585      0.021     27.942      0.000
    ENJOY              0.516      0.018     28.740      0.000
    HOPEFUL            1.314      0.039     33.541      0.000
    F                  0.977      0.007    145.590      0.000


R-SQUARE

    Observed                                        Two-Tailed
    Variable        Estimate       S.E.  Est./S.E.    P-Value
```

```

    SAD                0.475      0.019     25.057      0.000
    BLUES              0.442      0.019     22.804      0.000
    DEPRESS            0.547      0.019     28.364      0.000
    HAPPY              0.402      0.020     20.110      0.000
    ENJOY              0.368      0.020     18.478      0.000
    HOPEFUL            0.134      0.015      9.091      0.000

     Latent                                         Two-Tailed
    Variable        Estimate       S.E.  Est./S.E.    P-Value

    F                  0.023      0.007      3.466      0.001


TECHNICAL 1 OUTPUT


     PARAMETER SPECIFICATION


           NU
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
                  1             2             3             4             5


           NU
              HOPEFUL       AGE           MALE
              ________      ________      ________
                  6             0             0


```

```
             LAMBDA
                F              AGE            MALE
                --------       --------       --------
 SAD                   0              0              0
 BLUES                 7              0              0
 DEPRESS               8              0              0
 HAPPY                 9              0              0
 ENJOY                10              0              0
 HOPEFUL              11              0              0
 AGE                   0              0              0
 MALE                  0              0              0


             THETA
                SAD            BLUES          DEPRESS        HAPPY          ENJOY
                --------       --------       --------       --------       --------
 SAD                  12
 BLUES                 0             13
 DEPRESS               0              0             14
 HAPPY                 0              0              0             15
 ENJOY                 0              0              0              0             16
 HOPEFUL               0              0              0              0              0
 AGE                   0              0              0              0              0
 MALE                  0              0              0              0              0


             THETA
                HOPEFUL        AGE            MALE
                --------       --------       --------
 HOPEFUL              17
 AGE                   0              0
 MALE                  0              0              0
```

```


          ALPHA
             F              AGE            MALE
             ________       ________       ________
                 0              0              0


          BETA
             F              AGE            MALE
             ________       ________       ________
 F                   0             18             19
 AGE                 0              0              0
 MALE                0              0              0


          PSI
             F              AGE            MALE
             ________       ________       ________
 F                  20
 AGE                 0              0
 MALE                0              0              0


     STARTING VALUES


          NU
             SAD            BLUES          DEPRESS        HAPPY          ENJOY
             ________       ________       ________       ________       ________
               0.454          0.337          0.510          0.630          0.465

```

```

           NU
              HOPEFUL       AGE           MALE
              --------      --------      --------
                 0.973         0.000         0.000


           LAMBDA
              F             AGE           MALE
              --------      --------      --------
 SAD             1.000         0.000         0.000
 BLUES           0.991         0.000         0.000
 DEPRESS         1.218         0.000         0.000
 HAPPY           1.176         0.000         0.000
 ENJOY           1.031         0.000         0.000
 HOPEFUL         0.847         0.000         0.000
 AGE             0.000         1.000         0.000
 MALE            0.000         0.000         1.000


           THETA
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              --------      --------      --------      --------      --------
 SAD             0.274
 BLUES           0.000         0.273
 DEPRESS         0.000         0.000         0.337
 HAPPY           0.000         0.000         0.000         0.490
 ENJOY           0.000         0.000         0.000         0.000         0.408
 HOPEFUL         0.000         0.000         0.000         0.000         0.000
 AGE             0.000         0.000         0.000         0.000         0.000
 MALE            0.000         0.000         0.000         0.000         0.000

```

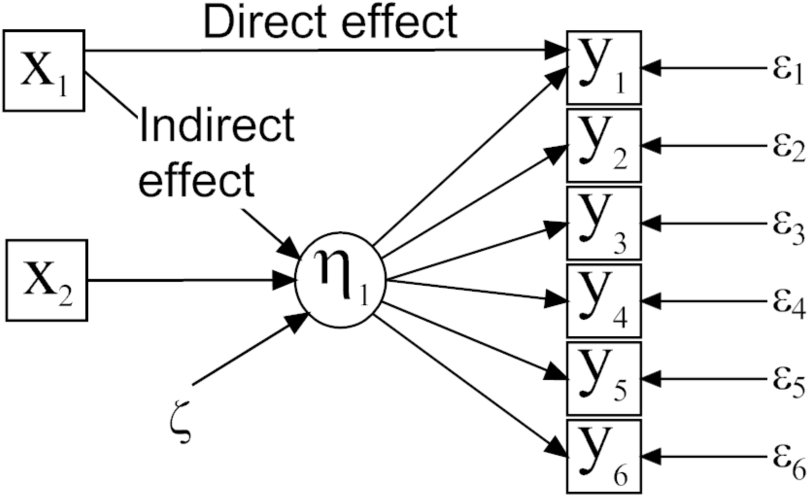
```

           THETA
              HOPEFUL       AGE           MALE
              ________      ________      ________
 HOPEFUL        0.759
 AGE            0.000         0.000
 MALE           0.000         0.000         0.000


           ALPHA
              F             AGE           MALE
              ________      ________      ________
                0.000         0.000         0.000


           BETA
              F             AGE           MALE
              ________      ________      ________
 F              0.000         0.000         0.000
 AGE            0.000         0.000         0.000
 MALE           0.000         0.000         0.000


           PSI
              F             AGE           MALE
              ________      ________      ________
 F              0.050
 AGE            0.000        40.696
 MALE           0.000        -0.190         0.244


     Beginning Time:  14:07:22
```

```
        Ending Time:  14:07:22
       Elapsed Time:  00:00:00


Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
  Maximum number of dependent variables: 6
  Maximum number of independent variables: 2
  Maximum number of between variables: 2
  Maximum number of continuous latent variables in time series analysis: 2


MUTHEN & MUTHEN
3463 Stoner Ave.
Los Angeles, CA  90066

Tel: (310) 391-9971
Fax: (310) 391-8971
Web: www.StatModel.com
Support: Support@StatModel.com

Copyright (c) 1998-2022 Muthen & Muthen
```

ex0102.out

Example 1.3

CFA with Covariates and Direct Effects
CES-D Depression, Age, Sex
New Haven EPESE



Direct effect
Indirect effect
x1
x2
η1
ζ
y1
y2
y3
y4
y5
y6
ε1
ε2
ε3
ε4
ε5
ε6

```
Mplus VERSION 8.8 DEMO (Mac)
MUTHEN & MUTHEN
07/11/2022   2:07 PM

INPUT INSTRUCTIONS

  TITLE:     EX0103 CFA WITH COVARIATES AND DIRECT EFFECT EPESE CESD ;

  DATA:      FILE = ex0101.dat ;

  VARIABLE: NAMES =  sad blues depress happy enjoy hopeful age male ;
             MISSING ARE ALL (-9999) ;

  OUTPUT:    TECH1 ;

  MODEL:     f by sad-hopeful ;
             f on age male ;
             sad-hopeful ON age-male@0 ;
             hopeful ON age ;





INPUT READING TERMINATED NORMALLY



EX0103 CFA WITH COVARIATES AND DIRECT EFFECT EPESE CESD ;

SUMMARY OF ANALYSIS

```

```
Number of groups                                                 1
Number of observations                                        2442

Number of dependent variables                                    6
Number of independent variables                                  2
Number of continuous latent variables                            1

Observed dependent variables

  Continuous
   SAD         BLUES       DEPRESS     HAPPY       ENJOY       HOPEFUL

Observed independent variables
   AGE         MALE

Continuous latent variables
   F


Estimator                                                       ML
Information matrix                                        OBSERVED
Maximum number of iterations                                  1000
Convergence criterion                                    0.500D-04
Maximum number of steepest descent iterations                   20
Maximum number of iterations for H1                           2000
Convergence criterion for H1                             0.100D-03

Input data file(s)
  ex0101.dat

Input data format  FREE

```

```

SUMMARY OF DATA

     Number of missing data patterns             1


COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value   0.100


     PROPORTION OF DATA PRESENT


           Covariance Coverage
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
 SAD            1.000
 BLUES          1.000         1.000
 DEPRESS        1.000         1.000         1.000
 HAPPY          1.000         1.000         1.000         1.000
 ENJOY          1.000         1.000         1.000         1.000         1.000
 HOPEFUL        1.000         1.000         1.000         1.000         1.000
 AGE            1.000         1.000         1.000         1.000         1.000
 MALE           1.000         1.000         1.000         1.000         1.000


           Covariance Coverage
              HOPEFUL       AGE           MALE
              ________      ________      ________
 HOPEFUL        1.000
 AGE            1.000         1.000
```

```
MALE             1.000          1.000          1.000



UNIVARIATE SAMPLE STATISTICS


     UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS

         Variable/          Mean/      Skewness/   Minimum/ % with
              Percentiles
        Sample Size      Variance    Kurtosis    Maximum  Min/Max      20%/60%    40%/
             80%    Median

     SAD                  0.454       1.831       0.000   65.77%       0.000
          0.000      0.000
            2442.000      0.548       3.185       3.000    3.81%       0.000
                1.000
     BLUES                0.337       2.417       0.000   78.05%       0.000
          0.000      0.000
            2442.000      0.545       5.259       3.000    4.34%       0.000
                1.000
     DEPRESS              0.510       1.749       0.000   64.41%       0.000
          0.000      0.000
            2442.000      0.674       2.457       3.000    5.81%       0.000
                1.000
     HAPPY                0.630       1.241       0.000   67.08%       0.000
          0.000      0.000
            2442.000      0.979       0.069       3.000    7.25%       0.000
                2.000
     ENJOY                0.465       1.742       0.000   75.96%       0.000
          0.000      0.000
```

```
             2442.000        0.816         1.624         3.000     5.94%        0.000
                  1.000
        HOPEFUL              0.973         0.681         0.000    57.70%        0.000
             0.000       0.000
             2442.000        1.517        -1.263         3.000    19.57%        1.000
                  2.000
        AGE                 74.316         0.497        67.000    28.50%       67.000
             72.000      72.000
             2442.000       40.696        -0.862        87.000     8.39%       77.000
                  82.000
        MALE                 0.424         0.307         0.000    57.58%        0.000
             0.000       0.000
             2442.000        0.244        -1.906         1.000    42.42%        1.000
                  1.000


THE MODEL ESTIMATION TERMINATED NORMALLY



MODEL FIT INFORMATION

Number of Free Parameters                        21

Loglikelihood

          H0 Value                      -17251.875
          H1 Value                      -17006.684

Information Criteria

          Akaike (AIC)                   34545.751
```

```
          Bayesian (BIC)                  34667.563
          Sample-Size Adjusted BIC        34600.841
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                             490.382
          Degrees of Freedom                     18
          P-Value                            0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                            0.104
          90 Percent C.I.                     0.096  0.112
          Probability RMSEA <= .05            0.000

CFI/TLI

          CFI                                 0.885
          TLI                                 0.827

Chi-Square Test of Model Fit for the Baseline Model

          Value                            4124.409
          Degrees of Freedom                     27
          P-Value                            0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                               0.048


```

```

MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F        BY
    SAD                1.000      0.000    999.000    999.000
    BLUES              0.963      0.035     27.792      0.000
    DEPRESS            1.191      0.040     30.113      0.000
    HAPPY              1.230      0.049     25.034      0.000
    ENJOY              1.074      0.045     24.002      0.000
    HOPEFUL            0.872      0.057     15.319      0.000

 F        ON
    AGE                0.005      0.002      3.007      0.003
    MALE              -0.134      0.023     -5.779      0.000

 SAD      ON
    AGE                0.000      0.000    999.000    999.000
    MALE               0.000      0.000    999.000    999.000

 BLUES    ON
    AGE                0.000      0.000    999.000    999.000
    MALE               0.000      0.000    999.000    999.000

 DEPRESS  ON
    AGE                0.000      0.000    999.000    999.000
    MALE               0.000      0.000    999.000    999.000

 HAPPY    ON
    AGE                0.000      0.000    999.000    999.000
```

```
    MALE               0.000      0.000    999.000    999.000

 ENJOY    ON
    AGE                0.000      0.000    999.000    999.000
    MALE               0.000      0.000    999.000    999.000

 HOPEFUL  ON
    AGE                0.010      0.004      2.572      0.010
    MALE               0.000      0.000    999.000    999.000

 Intercepts
    SAD                0.110      0.135      0.818      0.413
    BLUES              0.006      0.130      0.047      0.963
    DEPRESS            0.101      0.160      0.627      0.531
    HAPPY              0.207      0.166      1.245      0.213
    ENJOY              0.096      0.145      0.657      0.511
    HOPEFUL           -0.037      0.291     -0.126      0.900

 Residual Variances
    SAD                0.287      0.011     27.267      0.000
    BLUES              0.304      0.011     27.708      0.000
    DEPRESS            0.305      0.013     24.009      0.000
    HAPPY              0.586      0.021     27.941      0.000
    ENJOY              0.516      0.018     28.747      0.000
    HOPEFUL            1.312      0.039     33.574      0.000
    F                  0.254      0.015     17.467      0.000


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix              0.178E-06
       (ratio of smallest to largest eigenvalue)
```

```


TECHNICAL 1 OUTPUT


     PARAMETER SPECIFICATION


           NU
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
                  0             0             0             0             0


           NU
              HOPEFUL       AGE           MALE
              ________      ________      ________
                  0             0             0


           LAMBDA
              F             SAD           BLUES         DEPRESS       HAPPY
              ________      ________      ________      ________      ________
 SAD                0             0             0             0             0
 BLUES              0             0             0             0             0
 DEPRESS            0             0             0             0             0
 HAPPY              0             0             0             0             0
 ENJOY              0             0             0             0             0
 HOPEFUL            0             0             0             0             0
 AGE                0             0             0             0             0
 MALE               0             0             0             0             0

```

```

          LAMBDA
             ENJOY         HOPEFUL       AGE           MALE
             --------      --------      --------      --------
 SAD                0             0             0             0
 BLUES              0             0             0             0
 DEPRESS            0             0             0             0
 HAPPY              0             0             0             0
 ENJOY              0             0             0             0
 HOPEFUL            0             0             0             0
 AGE                0             0             0             0
 MALE               0             0             0             0


          THETA
             SAD           BLUES         DEPRESS       HAPPY         ENJOY
             --------      --------      --------      --------      --------
 SAD                0
 BLUES              0             0
 DEPRESS            0             0             0
 HAPPY              0             0             0             0
 ENJOY              0             0             0             0             0
 HOPEFUL            0             0             0             0             0
 AGE                0             0             0             0             0
 MALE               0             0             0             0             0


          THETA
             HOPEFUL       AGE           MALE
             --------      --------      --------
 HOPEFUL            0
 AGE                0             0
```

```
 MALE                0             0             0


          ALPHA
             F            SAD           BLUES         DEPRESS       HAPPY
             --------     --------      --------      --------      --------
                 0            1             2             3             4


          ALPHA
             ENJOY        HOPEFUL       AGE           MALE
             --------     --------      --------      --------
                 5            6             0             0


          BETA
             F            SAD           BLUES         DEPRESS       HAPPY
             --------     --------      --------      --------      --------
 F                   0             0             0             0             0
 SAD                 0             0             0             0             0
 BLUES               9             0             0             0             0
 DEPRESS            10             0             0             0             0
 HAPPY              11             0             0             0             0
 ENJOY              12             0             0             0             0
 HOPEFUL            13             0             0             0             0
 AGE                 0             0             0             0             0
 MALE                0             0             0             0             0


          BETA
             ENJOY        HOPEFUL       AGE           MALE
             --------     --------      --------      --------
```

```
F                0              0              7              8
SAD              0              0              0              0
BLUES            0              0              0              0
DEPRESS          0              0              0              0
HAPPY            0              0              0              0
ENJOY            0              0              0              0
HOPEFUL          0              0             14              0
AGE              0              0              0              0
MALE             0              0              0              0


          PSI
             F              SAD            BLUES          DEPRESS        HAPPY
             --------       --------       --------       --------       --------
F               15
SAD              0             16
BLUES            0              0             17
DEPRESS          0              0              0             18
HAPPY            0              0              0              0             19
ENJOY            0              0              0              0              0
HOPEFUL          0              0              0              0              0
AGE              0              0              0              0              0
MALE             0              0              0              0              0


          PSI
             ENJOY          HOPEFUL        AGE            MALE
             --------       --------       --------       --------
ENJOY           20
HOPEFUL          0             21
AGE              0              0              0
MALE             0              0              0              0
```

```


     STARTING VALUES


           NU
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              ________      ________      ________      ________      ________
                0.000         0.000         0.000         0.000         0.000


           NU
              HOPEFUL       AGE           MALE
              ________      ________      ________
                0.000         0.000         0.000


           LAMBDA
              F             SAD           BLUES         DEPRESS       HAPPY
              ________      ________      ________      ________      ________
 SAD            0.000         1.000         0.000         0.000         0.000
 BLUES          0.000         0.000         1.000         0.000         0.000
 DEPRESS        0.000         0.000         0.000         1.000         0.000
 HAPPY          0.000         0.000         0.000         0.000         1.000
 ENJOY          0.000         0.000         0.000         0.000         0.000
 HOPEFUL        0.000         0.000         0.000         0.000         0.000
 AGE            0.000         0.000         0.000         0.000         0.000
 MALE           0.000         0.000         0.000         0.000         0.000


           LAMBDA
              ENJOY         HOPEFUL       AGE           MALE
```

```
               --------      --------      --------      --------
 SAD              0.000         0.000         0.000         0.000
 BLUES            0.000         0.000         0.000         0.000
 DEPRESS          0.000         0.000         0.000         0.000
 HAPPY            0.000         0.000         0.000         0.000
 ENJOY            1.000         0.000         0.000         0.000
 HOPEFUL          0.000         1.000         0.000         0.000
 AGE              0.000         0.000         1.000         0.000
 MALE             0.000         0.000         0.000         1.000


           THETA
              SAD           BLUES         DEPRESS       HAPPY         ENJOY
              --------      --------      --------      --------      --------
 SAD              0.000
 BLUES            0.000         0.000
 DEPRESS          0.000         0.000         0.000
 HAPPY            0.000         0.000         0.000         0.000
 ENJOY            0.000         0.000         0.000         0.000         0.000
 HOPEFUL          0.000         0.000         0.000         0.000         0.000
 AGE              0.000         0.000         0.000         0.000         0.000
 MALE             0.000         0.000         0.000         0.000         0.000


           THETA
              HOPEFUL       AGE           MALE
              --------      --------      --------
 HOPEFUL          0.000
 AGE              0.000         0.000
 MALE             0.000         0.000         0.000
```

```
           ALPHA
              F             SAD           BLUES         DEPRESS       HAPPY
              ________      ________      ________      ________      ________
                0.000         0.454         0.337         0.510         0.630


           ALPHA
              ENJOY         HOPEFUL       AGE           MALE
              ________      ________      ________      ________
                0.465         0.973        74.316         0.424


           BETA
              F             SAD           BLUES         DEPRESS       HAPPY
              ________      ________      ________      ________      ________
 F              0.000         0.000         0.000         0.000         0.000
 SAD            1.000         0.000         0.000         0.000         0.000
 BLUES          0.991         0.000         0.000         0.000         0.000
 DEPRESS        1.218         0.000         0.000         0.000         0.000
 HAPPY          1.176         0.000         0.000         0.000         0.000
 ENJOY          1.031         0.000         0.000         0.000         0.000
 HOPEFUL        0.847         0.000         0.000         0.000         0.000
 AGE            0.000         0.000         0.000         0.000         0.000
 MALE           0.000         0.000         0.000         0.000         0.000


           BETA
              ENJOY         HOPEFUL       AGE           MALE
              ________      ________      ________      ________
 F              0.000         0.000         0.000         0.000
 SAD            0.000         0.000         0.000         0.000
 BLUES          0.000         0.000         0.000         0.000
```

```
DEPRESS        0.000         0.000         0.000         0.000
HAPPY          0.000         0.000         0.000         0.000
ENJOY          0.000         0.000         0.000         0.000
HOPEFUL        0.000         0.000         0.000         0.000
AGE            0.000         0.000         0.000         0.000
MALE           0.000         0.000         0.000         0.000


          PSI
             F             SAD           BLUES         DEPRESS       HAPPY
             ________      ________      ________      ________      ________
F              0.050
SAD            0.000         0.274
BLUES          0.000         0.000         0.273
DEPRESS        0.000         0.000         0.000         0.337
HAPPY          0.000         0.000         0.000         0.000         0.490
ENJOY          0.000         0.000         0.000         0.000         0.000
HOPEFUL        0.000         0.000         0.000         0.000         0.000
AGE            0.000         0.000         0.000         0.000         0.000
MALE           0.000         0.000         0.000         0.000         0.000


          PSI
             ENJOY         HOPEFUL       AGE           MALE
             ________      ________      ________      ________
ENJOY          0.408
HOPEFUL        0.000         0.759
AGE            0.000         0.000        40.696
MALE           0.000         0.000        -0.190         0.244


    Beginning Time:  14:07:23
```

Example 1.4

Growth modeling
Alcohol Use from Singer and Willet (2003)

Model a: Intercept only model

https://stats.idre.ucla.edu/other/examples/alda/

https://stats.idre.ucla.edu/other/examples/alda/

UCLA Institute for Digital Research & Education

Search this website

HOME SOFTWARE RESOURCES SERVICES ABOUT US

TEXTBOOK EXAMPLES APPLIED LONGITUDINAL DATA ANALYSIS: MODELING CHANGE AND EVENT OCCURRENCEBY JUDITH D. SINGER AND JOHN B. WILLETT

This is one of the books available for loan from Academic Technology Services (see Statistics Books for Loan for other such books, and details about borrowing). We encourage you to obtain *Applied Longitudinal Data Analysis*, written by Judith D. Singer and John B. Willett, published by the Oxford University Press, to gain a deeper conceptual understanding of the analysis illustrated (see Where to buy books for tips on different places you can buy this book). We are very grateful to the authors for granting us permission to create these pages and to distribute the data files via our web pages. (In addition to the data formats below, you can also download the data files as comma separated text files).

Note: The files for MLwiN here were created for older version of MLwiN (version 2.0 and older). The newer version of MLwiN is compatible with these files. But the newer version of the same files created at MLwiN site will only work for newer version of MLwiN (version 2.10). Please visit MLwiN site to download the files and for other MLwiN related information.

Note: The files for HLM consists of two types of files, the data files in SPSS format and the corresponding .mdmt (template) file for creating the binary HLM files. (Updated in January, 2012).

https://stats.idre.ucla.edu/r/examples/alda/applied-longitudinal-data-analysis-chapter-15-r-textbook-examples/

	M*plus*	SAS	Stata	R	SPSS	Chapter Title
Download Data	Download	Download	Download	Download	Download	Table of Contents
Chapter 1						A framework for investigating change over time
Chapter 2		Chap 2	Chap 2	Chap 2	Chap 2	Exploring longitudinal data on change
Chapter 3		Chap 3	Chap 3	Chap 3	Chap 3	Introducing the multilevel model for change

Chapter 4. Doing Data Analysis with the Multilevel Model for Change. Singer and Willet Applied Longitudinal Data Analysis. Oxford. (2003)

- Adolescent Alcohol Use Example (Curran et al. 2002).
- 82 Adolescents, age 14, seen 3 times +1 year
- 4 item Questionnaire Alcohol use past year
  - (1) drank beer or wine, (2) drank liquor, (3) had 5+ drinks in a row, (4) got drunk
  - rated 0-7 (0=not at all, 7=every day)
  - alcuse = (Sy)0.5
- Covariates/predictors
  - coa : child of alcoholic (0/1)
  - peer : alcohol use (alcuse) among peers [0-2.53]

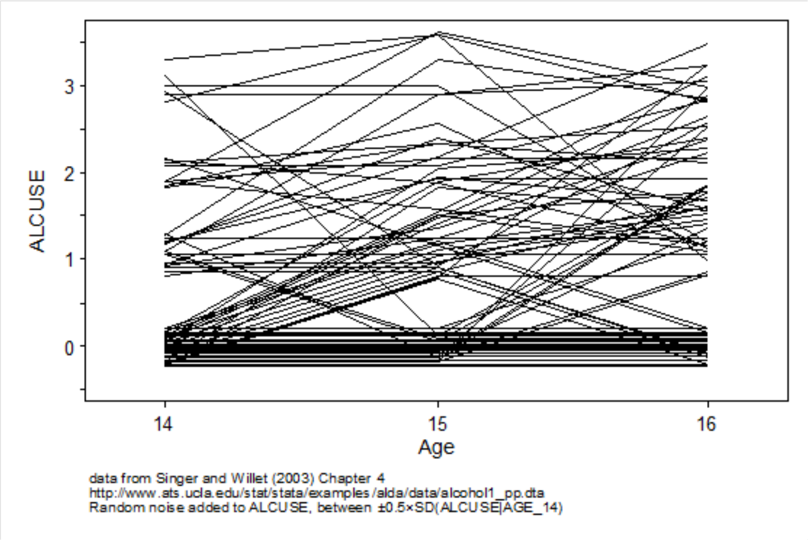

ALCUSE
0
1
2
3
14
15
16
Age
data from Singer and Willet (2003) Chapter 4
http://www.ats.ucla.edu/stat/stata/examples/alda/data/alcohol1_pp.dta
Random noise added to ALCUSE, between ±0.5×SD(ALCUSE|AGE_14)

Table 4.1: Alcohol use data (alcohol1_pp): Fitted multilevel models for change

Model A

Covariance Parameter Estimates

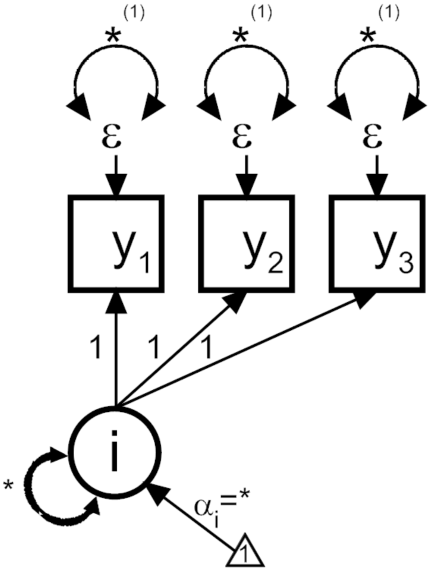
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	ID	0.5639	0.1191	4.73	<.0001
Residual		0.5617	0.06203	9.06	<.0001

Fit Statistics

-2 Log Likelihood	670.2
AIC (smaller is better)	676.2
AICC (smaller is better)	676.3
BIC (smaller is better)	683.4

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.9220	0.09571	81	9.63	<.0001


(1)
*
(1)
*
(1)
*
ε
ε
ε
y1
y2
y3
1
1
1
i
*
αi=*
1

```
Mplus VERSION 8.8 DEMO (Mac)
MUTHEN & MUTHEN
07/11/2022   2:07 PM

INPUT INSTRUCTIONS

  TITLE:     EX0104 Singer and Willet (2003) Chapter 4 Model A ;

  DATA:      FILE = swch4.dat ;

  VARIABLE: NAMES =  alcuse1 alcuse2 alcuse3 ;

  ANALYSIS: ESTIMATOR = mlr ;

  MODEL:     i by alcuse1-alcuse3@1 ;
             [alcuse1-alcuse3*]  ; ! force mean to intercept
             ![i*] ;                 ! intercept
             alcuse1-alcuse3 (1) ; ! homoscedasticity





INPUT READING TERMINATED NORMALLY



EX0104 Singer and Willet (2003) Chapter 4 Model A ;

SUMMARY OF ANALYSIS

Number of groups                                                 1
```

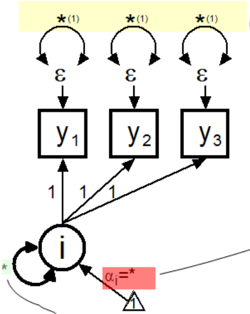
```
Number of observations                                          82

Number of dependent variables                                    3
Number of independent variables                                  0
Number of continuous latent variables                            1

Observed dependent variables

  Continuous
   ALCUSE1      ALCUSE2      ALCUSE3

Continuous latent variables
   I


Estimator                                                      MLR
Information matrix                                        OBSERVED
Maximum number of iterations                                  1000
Convergence criterion                                    0.500D-04
Maximum number of steepest descent iterations                   20

Input data file(s)
  swch4.dat

Input data format  FREE



UNIVARIATE SAMPLE STATISTICS


     UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS
```

```

          Variable/         Mean/     Skewness/   Minimum/ % with
               Percentiles
         Sample Size      Variance    Kurtosis    Maximum  Min/Max      20%/60%    40%/
              80%    Median

      ALCUSE1               0.630       1.320       0.000   62.20%       0.000
          0.000      0.000
               82.000       0.871       0.650       3.317    1.22%       0.000
                   1.414
      ALCUSE2               0.964       0.831       0.000   43.90%       0.000
          0.000      1.000
               82.000       1.084      -0.187       3.606    2.44%       1.000
                   2.000
      ALCUSE3               1.172       0.385       0.000   40.24%       0.000
          0.000      1.000
               82.000       1.273      -1.196       3.464    1.22%       1.414
                   2.236


THE MODEL ESTIMATION TERMINATED NORMALLY



MODEL FIT INFORMATION

Number of Free Parameters                        5

Loglikelihood

          H0 Value                        -323.402
          H0 Scaling Correction Factor      1.0942
```

```
              for MLR
           H1 Value                         -316.186
           H1 Scaling Correction Factor       1.2269
              for MLR

Information Criteria

           Akaike (AIC)                      656.804
           Bayesian (BIC)                    668.837
           Sample-Size Adjusted BIC          653.067
             (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

           Value                              10.361*
           Degrees of Freedom                      4
           P-Value                            0.0348
           Scaling Correction Factor          1.3928
             for MLR

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
    for chi-square difference testing in the regular way.  MLM, MLR and WLSM
    chi-square difference testing is described on the Mplus website.  MLMV, WLSMV,
    and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

           Estimate                            0.139
           90 Percent C.I.                     0.034  0.247
           Probability RMSEA <= .05            0.070

CFI/TLI
```

```

          CFI                               0.848
          TLI                               0.886

Chi-Square Test of Model Fit for the Baseline Model

          Value                            44.816
          Degrees of Freedom                    3
          P-Value                          0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                             0.117



MODEL RESULTS

                                                      Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 I        BY
    ALCUSE1            1.000      0.000    999.000    999.000
    ALCUSE2            1.000      0.000    999.000    999.000
    ALCUSE3            1.000      0.000    999.000    999.000

 Intercepts
    ALCUSE1            0.630      0.103      6.118      0.000
    ALCUSE2            0.964      0.115      8.383      0.000
    ALCUSE3            1.172      0.125      9.403      0.000

 Variances
```

```
    I                  0.589      0.103      5.714      0.000

 Residual Variances
    ALCUSE1            0.487      0.071      6.845      0.000
    ALCUSE2            0.487      0.071      6.845      0.000
    ALCUSE3            0.487      0.071      6.845      0.000


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix              0.102E+00
       (ratio of smallest to largest eigenvalue)


     Beginning Time:  14:07:23
        Ending Time:  14:07:23
       Elapsed Time:  00:00:00


Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
  Maximum number of dependent variables: 6
  Maximum number of independent variables: 2
  Maximum number of between variables: 2
  Maximum number of continuous latent variables in time series analysis: 2


MUTHEN & MUTHEN
3463 Stoner Ave.
Los Angeles, CA  90066

Tel: (310) 391-9971
Fax: (310) 391-8971
```



MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I BY				
ALCUSE1	1.000	0.000	999.000	999.000
ALCUSE2	1.000	0.000	999.000	999.000
ALCUSE3	1.000	0.000	999.000	999.000
Means				
I	0.922	0.096	9.633	0.000
Intercepts				
ALCUSE1	0.000	0.000	999.000	999.000
ALCUSE2	0.000	0.000	999.000	999.000
ALCUSE3	0.000	0.000	999.000	999.000
Variances				
I	0.564	0.106	5.314	0.000
Residual Variances				
ALCUSE1	0.562	0.083	6.788	0.000
ALCUSE2	0.562	0.083	6.788	0.000
ALCUSE3	0.562	0.083	6.788	0.000

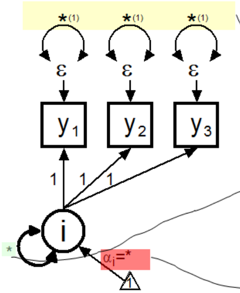


Table 4.1: Alcohol use data (alcohol1_pp): Fitted multilevel models for change

Model A

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	ID	0.5639	0.1191	4.73	<.0001
Residual		0.5617	0.06203	9.06	<.0001

Fit Statistics

-2 Log Likelihood	670.2
AIC (smaller is better)	676.2
AICC (smaller is better)	676.3
BIC (smaller is better)	683.4

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.9220	0.09571	81	9.63	<.0001

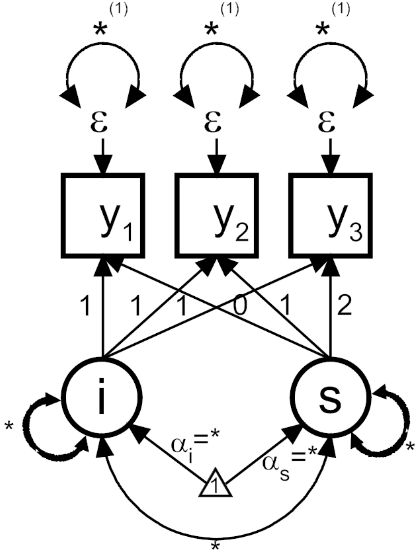
Note that the point estimates are the same. Standard errors of covariance parameters are slightly higher in PROC MIXED. It's a difference of N-1 (PROC Mixed) vs N in the denominator.

$$
\begin{aligned}
\text{SAS Proc Mixed} &\approx \text{M}\textit{plus} \\
\text{intercept} &\approx \text{Means(I)} \\
\text{Cov parm intercept} &\approx \text{Variance(I)} \\
\text{Cov parm residual} &\approx \text{Residual Variances(y)}
\end{aligned}
$$

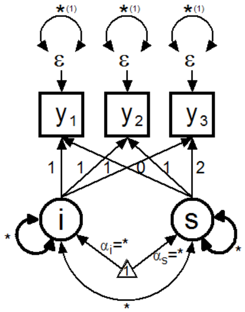
Example 1.5

Growth modeling
Alcohol Use from Singer and Willet (2003)

Model b: Intercept and Slope model

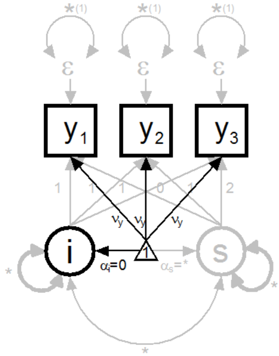

(1)
*
(1)
*
(1)
*
ε
ε
ε
y1
y2
y3
1
1
1
0
1
2
i
s
*
αi=*
αs=*
1
*
*

Model B (Table 4.1)



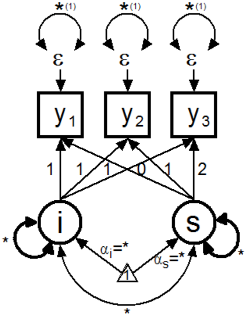

```
* SAS Example from http://www.ats.ucla.edu/stat/examples/alda.htm
*---------------------------------------------------------------------------
* SAS
proc mixed data="c:\alda\alcohol1_pp" method=ml noclprint noinfo covtest;
  title2 "Model B";
  class id;
  model alcuse = age_14/solution notest;
  random intercept age_14/type=un sub=id;

*---------------------------------------------------------------------------

* Mplus (short hand)
DATA:      FILE = C:\work\Shows\SHORTC~1\2009\data\swch4.dat ;
VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
ANALYSIS: ESTIMATOR = mlr ;
MODEL:     i s | alcuse1@0 alcuse2@1 alcuse3@2 ;
             alcuse1-alcuse3 (1) ;

*---------------------------------------------------------------------------

* Mplus (long hand)
DATA:      FILE = C:\work\Shows\SHORTC~1\2009\data\swch4.dat ;
VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
ANALYSIS: ESTIMATOR = mlr ;
MODEL:     i by alcuse1-alcuse3@1 ;
           s by alcuse1@0 alcuse2@1 alcuse3@2 ;
           [alcuse1-alcuse3@0] ;
           [i* s*] ;
           alcuse1-alcuse3 (1) ;
```

Model B (Table 4.1) Alternative Mplus Specification



note: $\nu_y = \nu_{y(1)} = \nu_{y(2)} = \nu_{y(3)}$

```
* Mplus  (alternate specification - long hand)
DATA:      FILE = swch4.dat ;

VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;

ANALYSIS: ESTIMATOR = mlr ;

MODEL:     alcuse1-alcuse3 (1) ;
           i by alcuse1-alcuse3@1 ;
           s by alcuse1@0 alcuse2@1 alcuse3@2 ;
           [alcuse1-alcuse3*0] (11) ;
           [i@0 s*] ;
```



$$V(\mathrm{y}_{ti}) = V(\mathrm{i})_i \boxed{+ V(\mathrm{s})_i t^2} + 2COV(\mathrm{i,s}) + V(\epsilon_{ti})$$

random slope contributes to increasing variance over time even if ε_t=ε

```
Mplus VERSION 8.8 DEMO (Mac)
MUTHEN & MUTHEN
07/11/2022   2:07 PM

INPUT INSTRUCTIONS

  TITLE:     EX0105 Singer and Willet (2003) Chapter 4 Model B ;

  DATA:      FILE = swch4.dat ;

  VARIABLE: NAMES =  alcuse1 alcuse2 alcuse3 ;
            missing are all (-9999) ;

  ANALYSIS: ESTIMATOR = mlr ;

  OUTPUT:    SAMPSTAT ; stdyx ;

  MODEL:
            ![alcuse1@0 alcuse2@0 alcuse3@0] ;
            !i by alcuse1@1 alcuse2@1 alcuse3@1 ;
            ! s by alcuse1@0 alcuse2@1 alcuse3@2 ;
            ![i*]; [s*];
            i s | alcuse1@0 alcuse2@1 alcuse3@2 ;
            alcuse1-alcuse3 (1) ; ! homoscedasticity





INPUT READING TERMINATED NORMALLY


```

```

EX0105 Singer and Willet (2003) Chapter 4 Model B ;

SUMMARY OF ANALYSIS

Number of groups                                                 1
Number of observations                                          82

Number of dependent variables                                    3
Number of independent variables                                  0
Number of continuous latent variables                            2

Observed dependent variables

  Continuous
   ALCUSE1     ALCUSE2     ALCUSE3

Continuous latent variables
   I           S


Estimator                                                      MLR
Information matrix                                        OBSERVED
Maximum number of iterations                                  1000
Convergence criterion                                    0.500D-04
Maximum number of steepest descent iterations                   20
Maximum number of iterations for H1                           2000
Convergence criterion for H1                             0.100D-03

Input data file(s)
  swch4.dat

```

```
Input data format  FREE


SUMMARY OF DATA

     Number of missing data patterns             1


COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value   0.100


     PROPORTION OF DATA PRESENT


           Covariance Coverage
              ALCUSE1       ALCUSE2       ALCUSE3
              ________      ________      ________
 ALCUSE1        1.000
 ALCUSE2        1.000         1.000
 ALCUSE3        1.000         1.000         1.000


SAMPLE STATISTICS


     ESTIMATED SAMPLE STATISTICS


           Means
              ALCUSE1       ALCUSE2       ALCUSE3
```

```
                  --------      --------      --------
                    0.630         0.964         1.172


               Covariances
                  ALCUSE1       ALCUSE2       ALCUSE3
                  --------      --------      --------
 ALCUSE1            0.871
 ALCUSE2            0.620         1.084
 ALCUSE3            0.432         0.714         1.273


               Correlations
                  ALCUSE1       ALCUSE2       ALCUSE3
                  --------      --------      --------
 ALCUSE1            1.000
 ALCUSE2            0.638         1.000
 ALCUSE3            0.411         0.608         1.000


     MAXIMUM LOG-LIKELIHOOD VALUE FOR THE UNRESTRICTED (H1) MODEL IS -316.186


UNIVARIATE SAMPLE STATISTICS


     UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS

         Variable/         Mean/     Skewness/   Minimum/ % with
                Percentiles
        Sample Size      Variance    Kurtosis    Maximum  Min/Max      20%/60%    40%/
               80%    Median
```

```
     ALCUSE1                0.630        1.320          0.000   62.20%         0.000
         0.000      0.000
               82.000       0.871        0.650          3.317    1.22%         0.000
                    1.414
     ALCUSE2                0.964        0.831          0.000   43.90%         0.000
         0.000      1.000
               82.000       1.084       -0.187          3.606    2.44%         1.000
                    2.000
     ALCUSE3                1.172        0.385          0.000   40.24%         0.000
         0.000      1.000
               82.000       1.273       -1.196          3.464    1.22%         1.414
                    2.236


THE MODEL ESTIMATION TERMINATED NORMALLY



MODEL FIT INFORMATION

Number of Free Parameters                        6

Loglikelihood

          H0 Value                         -318.306
          H0 Scaling Correction Factor       1.1192
            for MLR
          H1 Value                         -316.186
          H1 Scaling Correction Factor       1.2269
            for MLR

```

```
Information Criteria

          Akaike (AIC)                     648.611
          Bayesian (BIC)                   663.051
          Sample-Size Adjusted BIC         644.128
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              2.939*
          Degrees of Freedom                     3
          P-Value                           0.4011
          Scaling Correction Factor         1.4422
            for MLR

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
    for chi-square difference testing in the regular way.  MLM, MLR and WLSM
    chi-square difference testing is described on the Mplus website.  MLMV, WLSMV,
    and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.000
          90 Percent C.I.                    0.000  0.185
          Probability RMSEA <= .05           0.490

CFI/TLI

          CFI                                1.000
          TLI                                1.000

Chi-Square Test of Model Fit for the Baseline Model
```

```

          Value                              44.816
          Degrees of Freedom                      3
          P-Value                            0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                               0.055



MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 I        |
    ALCUSE1            1.000      0.000    999.000    999.000
    ALCUSE2            1.000      0.000    999.000    999.000
    ALCUSE3            1.000      0.000    999.000    999.000

 S        |
    ALCUSE1            0.000      0.000    999.000    999.000
    ALCUSE2            1.000      0.000    999.000    999.000
    ALCUSE3            2.000      0.000    999.000    999.000

 S        WITH
    I                 -0.068      0.074     -0.920      0.357

 Means
    I                  0.651      0.105      6.198      0.000
    S                  0.271      0.062      4.334      0.000
```

```

 Intercepts
    ALCUSE1            0.000      0.000    999.000    999.000
    ALCUSE2            0.000      0.000    999.000    999.000
    ALCUSE3            0.000      0.000    999.000    999.000

 Variances
    I                  0.624      0.162      3.861      0.000
    S                  0.151      0.056      2.721      0.007

 Residual Variances
    ALCUSE1            0.337      0.066      5.105      0.000
    ALCUSE2            0.337      0.066      5.105      0.000
    ALCUSE3            0.337      0.066      5.105      0.000


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix              0.422E-01
       (ratio of smallest to largest eigenvalue)


STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                    Two-Tailed
                   Estimate       S.E.  Est./S.E.    P-Value

 I        |
    ALCUSE1            0.806      0.046     17.355      0.000
```

```
    ALCUSE2              0.800        0.083        9.686        0.000
    ALCUSE3              0.695        0.092        7.552        0.000

 S          |
    ALCUSE1              0.000        0.000      999.000      999.000
    ALCUSE2              0.394        0.080        4.938        0.000
    ALCUSE3              0.684        0.120        5.679        0.000

 S          WITH
    I                   -0.223        0.193       -1.154        0.249

 Means
    I                    0.824        0.097        8.465        0.000
    S                    0.696        0.179        3.891        0.000

 Intercepts
    ALCUSE1              0.000        0.000      999.000      999.000
    ALCUSE2              0.000        0.000      999.000      999.000
    ALCUSE3              0.000        0.000      999.000      999.000

 Variances
    I                    1.000        0.000      999.000      999.000
    S                    1.000        0.000      999.000      999.000

 Residual Variances
    ALCUSE1              0.351        0.075        4.688        0.000
    ALCUSE2              0.346        0.062        5.588        0.000
    ALCUSE3              0.261        0.058        4.486        0.000


R-SQUARE

```

```
     Observed                                                 Two-Tailed
     Variable          Estimate        S.E.  Est./S.E.    P-Value

     ALCUSE1              0.649      0.075      8.678      0.000
     ALCUSE2              0.654      0.062     10.581      0.000
     ALCUSE3              0.739      0.058     12.706      0.000


      Beginning Time:  14:07:23
         Ending Time:  14:07:23
        Elapsed Time:  00:00:00


Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
  Maximum number of dependent variables: 6
  Maximum number of independent variables: 2
  Maximum number of between variables: 2
  Maximum number of continuous latent variables in time series analysis: 2


MUTHEN & MUTHEN
3463 Stoner Ave.
Los Angeles, CA  90066

Tel: (310) 391-9971
Fax: (310) 391-8971
Web: www.StatModel.com
Support: Support@StatModel.com

Copyright (c) 1998-2022 Muthen & Muthen
```

ex0105.out

Table 4.1: Alcohol use data (alcohol1_pp): Fitted multilevel models for change

Model B

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	ID	0.6244	0.1481	4.22	<.0001
UN(2,1)	ID	-0.06844	0.07008	-0.98	0.3288
UN(2,2)	ID	0.1512	0.05647	2.68	0.0037
Residual		0.3373	0.05268	6.40	<.0001

Fit Statistics

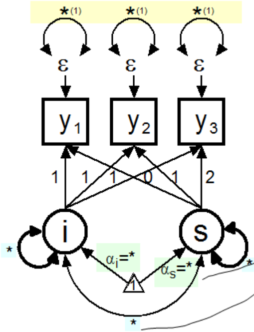
-2 Log Likelihood	636.6
AIC (smaller is better)	648.6
AICC (smaller is better)	649.0
BIC (smaller is better)	663.1

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	79.70	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.6513	0.1051	81	6.20	<.0001
AGE_14	0.2707	0.06245	81	4.33	<.0001



MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
S WITH				
I	-0.068	0.074	-0.920	0.357
Means				
I	0.651	0.105	6.198	0.000
S	0.271	0.062	4.334	0.000
Intercepts				
ALCUSE1	0.000	0.000	999.000	999.000
ALCUSE2	0.000	0.000	999.000	999.000
ALCUSE3	0.000	0.000	999.000	999.000
Variances				
I	0.624	0.162	3.861	0.000
S	0.151	0.056	2.721	0.007
Residual Variances				
ALCUSE1	0.337	0.066	5.105	0.000
ALCUSE2	0.337	0.066	5.105	0.000
ALCUSE3	0.337	0.066	5.105	0.000

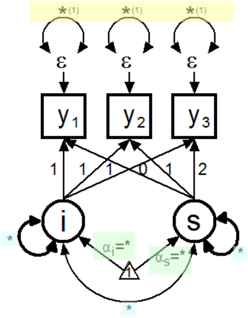


Table 4.1: Alcohol use data (alcohol1_pp): Fitted multilevel models for change
Model B

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	ID	0.6244	0.1481	4.22	<.0001
UN(2,1)	ID	-0.06844	0.07008	-0.98	0.3288
UN(2,2)	ID	0.1512	0.05647	2.68	0.0037
Residual		0.3373	0.05268	6.40	<.0001

Fit Statistics

-2 Log Likelihood	636.6
AIC (smaller is better)	648.6
AICC (smaller is better)	649.0
BIC (smaller is better)	663.1

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	79.70	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.6513	0.1051	81	6.20	<.0001
AGE_14	0.2707	0.06245	81	4.33	<.0001

$$
\begin{aligned}
\text{SAS Proc Mixed} &\approx \text{M}\textit{plus} \\
\text{Intercept} &\approx \text{Means(I)} \\
\text{AGE\_14} &\approx \text{Means(S)} \\
\text{Cov parm UN(1,1)} &\approx \text{Variance(I)} \\
\text{Cov parm UN(2,1)} &\approx \text{S WITH I} \\
\text{Cov parm UN(2,2)} &\approx \text{Variance(S)} \\
\text{Cov parm residual} &\approx \text{Residual Variances(y)}
\end{aligned}
$$

LAB SESSION Alcohol Use LGC Models

Questions?

END OF DAY ONE