Mplus Short Courses
Topic 3

Growth Modeling With Latent Variables Using Mplus: Introductory And Intermediate Growth Models

Linda K. Muthén
Bengt Muthén

www.statmodel.com
05/17/2010

Table Of Contents

Mplus Background

- Inefficient dissemination of statistical methods:
  - Many good methods contributions from biostatistics, psychometrics, etc are underutilized in practice
- Fragmented presentation of methods:
  - Technical descriptions in many different journals
  - Many different pieces of limited software
- Mplus: Integration of methods in one framework
  - Easy to use: Simple, non-technical language, graphics
  - Powerful: General modeling capabilities

Mplus Background

- Mplus versions
  - V1: November 1998
  - V2: February 2001
  - V3: March 2004
  - V4: February 2006
  - V5: November 2007
  - V5.21: May 2009
  - V6: April, 2010

- Mplus team: Linda & Bengt Muthén, Thuy Nguyen, Tihomir Asparouhov, Michelle Conn, Jean Maninger

Statistical Analysis With Latent Variables A General Modeling Framework

Statistical Concepts Captured By Latent Variables

Continuous Latent Variables

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

Categorical Latent Variables

- Latent classes
- Clusters
- Finite mixtures
- Missing data

Statistical Analysis With Latent Variables A General Modeling Framework (Continued)

Models That Use Latent Variables

Continuous Latent Variables

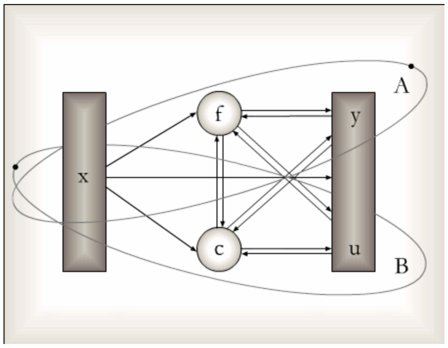
- Factor analysis models
- Structural equation models
- Growth curve models
- Multilevel models

Categorical Latent Variables

- Latent class models
- Mixture models
- Discrete-time survival models
- Missing data models

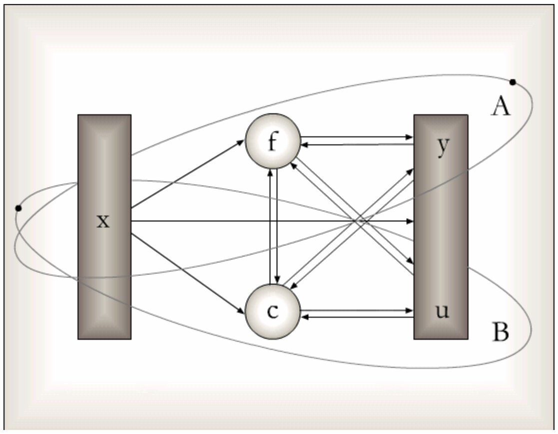
Mplus integrates the statistical concepts captured by latent variables into a general modeling framework that includes not only all of the models listed above but also combinations and extensions of these models.

General Latent Variable Modeling Framework

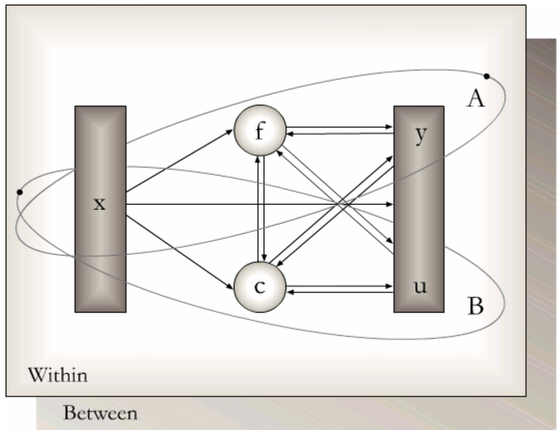


- Observed variables
  - x background variables (no model structure)
  - y continuous and censored outcome variables
  - u categorical (dichotomous, ordinal, nominal) and count outcome variables
- Latent variables
  - f continuous variables
    - – interactions among f's
  - c categorical variables
    - – multiple c's

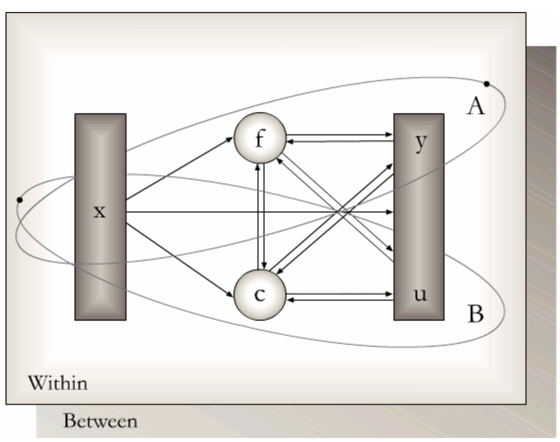
General Latent Variable Modeling Framework



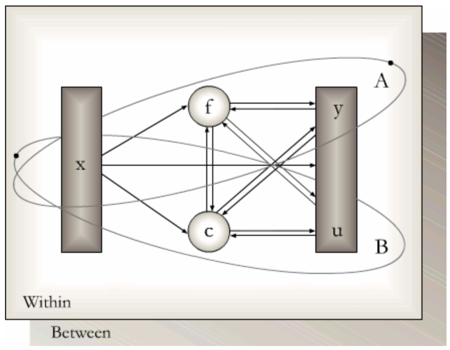
General Latent Variable Modeling Framework



General Latent Variable Modeling Framework



General Latent Variable Modeling Framework



- Observed variables
  - x background variables (no model structure)
  - y continuous and censored outcome variables
  - u categorical (dichotomous, ordinal, nominal) and count outcome variables
- Latent variables
  - f continuous variables
    - – interactions among f's
  - c categorical variables
    - – multiple c's

Mplus

Several programs in one

- Exploratory factor analysis
- Structural equation modeling
- Item response theory analysis
- Latent class analysis
- Latent transition analysis
- Survival analysis
- Growth modeling
- Multilevel analysis
- Complex survey data analysis
- Monte Carlo simulation

Fully integrated in the general latent variable framework

Overview Of Mplus Courses

- **Topic 1.** August 20, 2009, Johns Hopkins University: Introductory - advanced factor analysis and structural equation modeling with continuous outcomes
- **Topic 2.** August 21, 2009, Johns Hopkins University: Introductory - advanced regression analysis, IRT, factor analysis and structural equation modeling with categorical, censored, and count outcomes

- **Topic 3.** March 22, 2010, Johns Hopkins University: Introductory and intermediate growth modeling
- **Topic 4.** March 23, 2010, Johns Hopkins University: Advanced growth modeling, survival analysis, and missing data analysis

Overview Of Mplus Courses (Continued)

- **Topic 5.** August 16, 2010, Johns Hopkins University:

Categorical latent variable modeling with cross-sectional data

- **Topic 6.** August 17, 2010, Johns Hopkins University:

Categorical latent variable modeling with longitudinal data

- **Extra Topic.** August 18, 2010, Johns Hopkins University:

What's new in Mplus version 6?

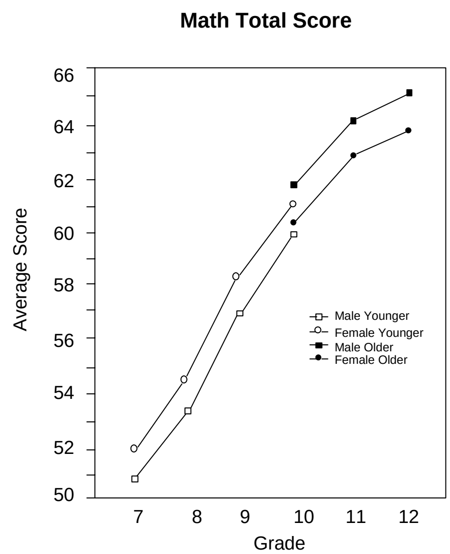
- **Topic 7.** March, 2011, Johns Hopkins University: Multilevel modeling of cross-sectional data
- **Topic 8.** March, 2011, Johns Hopkins University: Multilevel modeling of longitudinal data

Typical Examples Of Growth Modeling

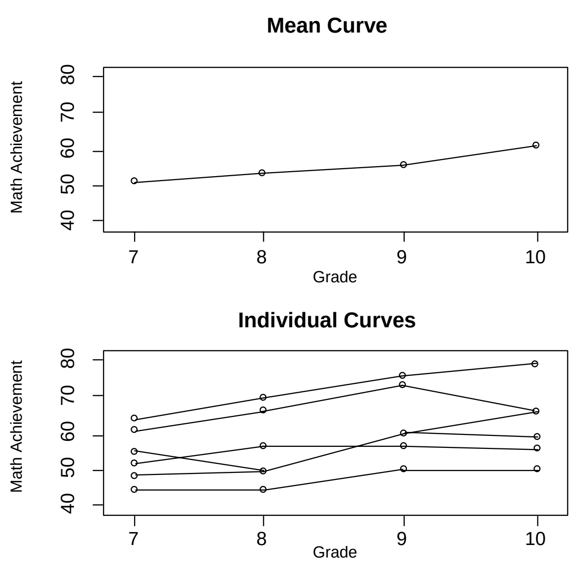
LSAY Data

Longitudinal Study of American Youth (LSAY)

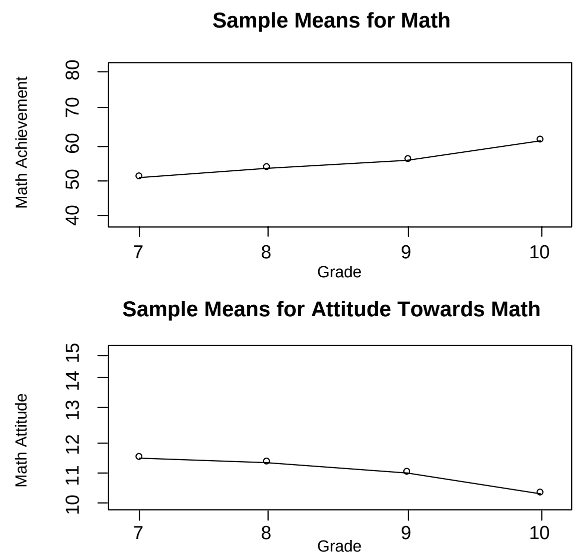
- Two cohorts measured each year beginning in 1987
  - Cohort 1 - Grades 10, 11, and 12
  - Cohort 2 - Grades 7, 8, 9, 10, 11, and 12
- Each cohort contains approximately 60 schools with approximately 60 students per school
- Variables - math and science achievement items, math and science attitude measures, and background variables from parents, teachers, and school principals
- Approximately 60 items per test with partial item overlap across grades - adaptive tests


Math Total Score
Average Score
66
64
62
60
58
56
54
52
50
7
8
9
10
11
12
Grade
Male Younger
Female Younger
Male Older
Female Older

17


Mean Curve
Math Achievement
80
70
60
50
40
7
8
9
10
Grade
Individual Curves
Math Achievement
80
70
60
50
40
7
8
9
10
Grade

18

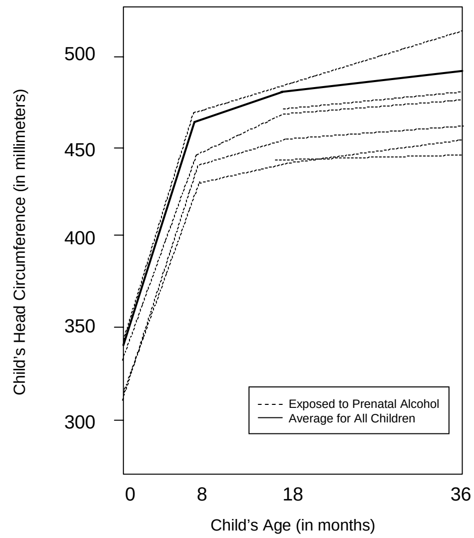


Maternal Health Project Data

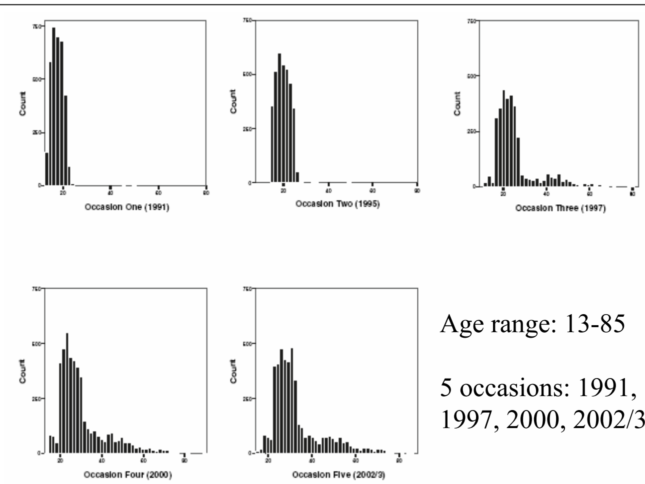
Maternal Health Project (MHP)

- Mothers who drank at least three drinks a week during their first trimester plus a random sample of mothers who used alcohol less often
- Mothers measured at fourth month and seventh month of pregnancy, at delivery, and at 8, 18, and 36 months postpartum
- Offspring measured at 0, 8, 18 and 36 months
- Variables for mothers - demographic, lifestyle, current environment, medical history, maternal psychological status, alcohol use, tobacco use, marijuana use, other illicit drug use
- Variables for offspring - head circumference, height, weight, gestational age, gender, and ethnicity

MHP: Offspring Head Circumference



Loneliness In Twins



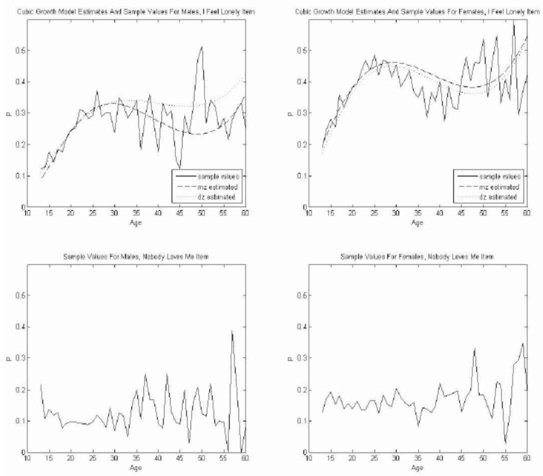
Age range: 13-85

5 occasions: 1991, 1995, 1997, 2000, 2002/3

Boomsma, D.I., Cacioppo, J.T., Muthen, B., Asparouhov, T., & Clark, S. (2007). Longitudinal Genetic Analysis for Loneliness in Dutch Twins. Twin Research and Human Genetics, 10, 267-273.

Loneliness In Twins

Males Females



I feel lonely

Nobody loves me

Basic Modeling Ideas

Longitudinal Data: Three Approaches

Three modeling approaches for the regression of outcome on time (n is sample size, T is number of timepoints):

- **Use all n x T data points to do a single regression analysis**: Gives an intercept and a slope estimate common to all individuals - does not account for individual differences or lack of independence of observations
- **Use each individual's T data points to do n regression analyses:** Gives an intercept and a slope estimate for each individual. Accounts for individual differences, but does not account for similarities among individuals
- **Use all n x T data points to do a single random effect regression analysis:** Gives an intercept and a slope estimate for each individual. Accounts for similarities among individuals by stipulating that all individuals' random effects come from a single, common population and models the non-independence of observations as show on the next page

Individual Development Over Time

(1) $y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}$

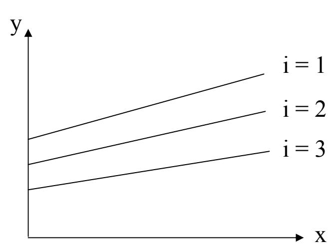
(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

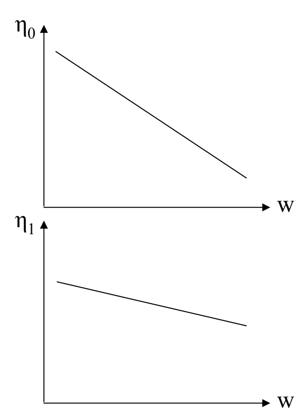
(2b) $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$

t = timepoint   i = individual   w = time-invariant covariate

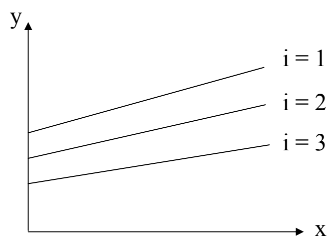
y = outcome   x = time score

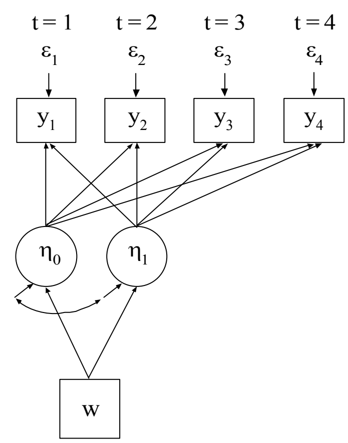
η_0 = intercept η_1 = slope





Individual Development Over Time





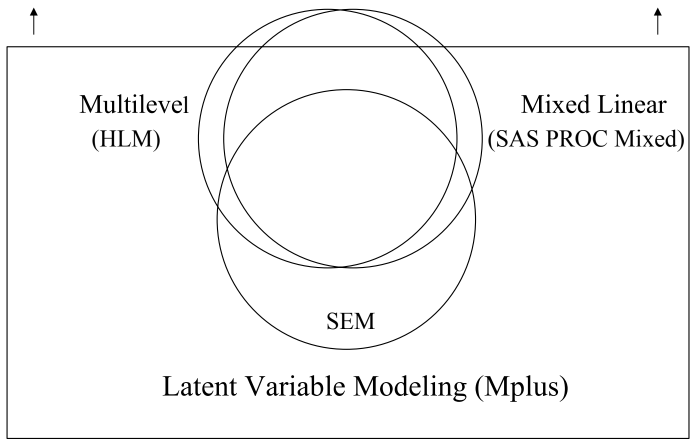
(1) $y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}$

(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

(2b) $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$

Growth Modeling Frameworks

Growth Modeling Frameworks/Software



Comparison Summary Of Multilevel, Mixed Linear, And SEM Growth Models

- Multilevel and mixed linear models are the same
- SEM differs from the multilevel and mixed linear models in two ways
  - Treatment of time scores
    - Time scores are data for multilevel and mixed linear models -- individuals can have different times of measurement
    - Time scores are parameters for SEM growth models -- time scores can be estimated
  - Treatment of time-varying covariates
    - Time-varying covariates have random effect coefficients for multilevel and mixed linear models -- coefficients vary over individuals
    - Time-varying covariates have fixed effect coefficients for SEM growth models -- coefficients vary over time

Random Effects: Multilevel And Mixed Linear Modeling

Individual i ($i = 1, 2, \ldots, n$) observed at time point t ($t = 1, 2, \ldots T$).

Multilevel model with two levels (e.g. Raudenbush & Bryk, 2002, HLM).

- Level 1: $y_{ti} = \eta_{0i} + \eta_{1i}\, x_{ti} + \kappa_i\, w_{ti} + \varepsilon_{ti}$ (39)
- Level 2: $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$ (40)
  $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$ (41)
  $\kappa_i = \alpha + \gamma\, w_i + \zeta_i$ (42)

Random Effects: Multilevel And Mixed Linear Modeling (Continued)

Mixed linear model:

$$y_{ti} = \textit{fixed part} + \textit{random part} \quad (43)$$
$$= \alpha_0 + \gamma_0\, w_i + (\alpha_1 + \gamma_1\, w_i)\, x_{ti} + (\alpha + \gamma\, w_i)\, w_{ti} \quad (44)$$
$$+ \zeta_{0i} + \zeta_{1i}\, x_{ti} + \zeta_i\, w_{ti} + \varepsilon_{ti}\,. \quad (45)$$

E.g. "*time* x w_i" refers to γ_1 (e.g. Rao, 1958; Laird & Ware, 1982; Jennrich & Sluchter, 1986; Lindstrom & Bates, 1988; BMDP5V; Goldstein, 2003, MLwiN; SAS PROC MIXED-Littell et al. 1996 and Singer, 1999).

Random Effects: SEM And Multilevel Modeling

SEM (Tucker, 1958; Meredith & Tisak, 1990; McArdle & Epstein 1987; SEM software):

Measurement part:

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \kappa_t\, w_{ti} + \varepsilon_{ti}\,. \qquad (46)$$

Compare with level 1 of multilevel:

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_{ti} + \kappa_i\, w_{ti} + \varepsilon_{ti}\,. \qquad (47)$$

Multilevel approach:

- x_{ti} as data: Flexible individually-varying times of observation
- Slopes for time-varying covariates vary over individuals

Random Effects: SEM And Multilevel Modeling (Continued)

SEM approach:

- x_t as parameters: Flexible growth function form
- Slopes for time-varying covariates vary over time points

Structural part (same as level 2, except for κ_t):

$$\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}\,, \qquad (48)$$

$$\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}\,, \qquad (49)$$

κ_t not involved (parameter).

Random Effects: Mixed Linear Modeling And SEM

Mixed linear model in matrix form:

$$\boldsymbol{y}_i = (y_{1i}, y_{2i}, \ldots, y_{Ti})' \qquad (51)$$

$$= \boldsymbol{X}_i \boldsymbol{\alpha} + \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{e}_i . \qquad (52)$$

Here, ***X, Z*** are design matrices with known values, $\boldsymbol{\alpha}$ contains fixed effects, and ***b*** contains random effects. Compare with (43) - (45).

Random Effects: Mixed Linear Modeling And SEM (Continued)

SEM in matrix form:

$$y_i = \nu + \Lambda\, \eta_i + K\, x_i + \varepsilon_i , \qquad (53)$$
$$\eta_i = \alpha + B\, \eta_i + \Gamma\, x_i + \zeta_i . \qquad (54)$$

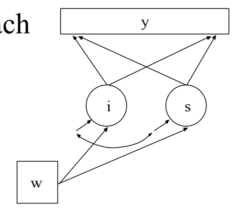
$$y_i = \textit{fixed part} + \textit{random part}$$
$$= \nu + \Lambda\, (I - B)^{-1}\, \alpha + \Lambda\, (I - B)^{-1}\, \Gamma\, x_i + K\, x_i$$
$$+ \Lambda\, (I - B)^{-1}\, \zeta_i + \varepsilon_i .$$

Assume $x_{ti} = x_t$, $\kappa_i = \kappa_t$ in (39). Then (39) is handled by (53) and (40) – (41) are handled by (54), putting x_t in Λ and w_{ti}, w_i in x_i.

Need for Λ_i, K_i, B_i, Γ_i.

Growth Modeling Approached In Two Ways: Data Arranged As Wide Versus Long

- Wide: Multivariate, Single-Level Approach

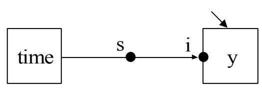


$y_{ti} = i_i + s_i \times \text{time}_{ti} + \varepsilon_{ti}$

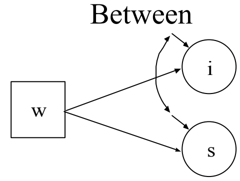
i_i regressed on w_i
s_i regressed on w_i

- Long: Univariate, 2-Level Approach (CLUSTER = id)

Within



Between



The intercept *i* is called *y* in Mplus

Pros And Cons Of Wide Versus Long

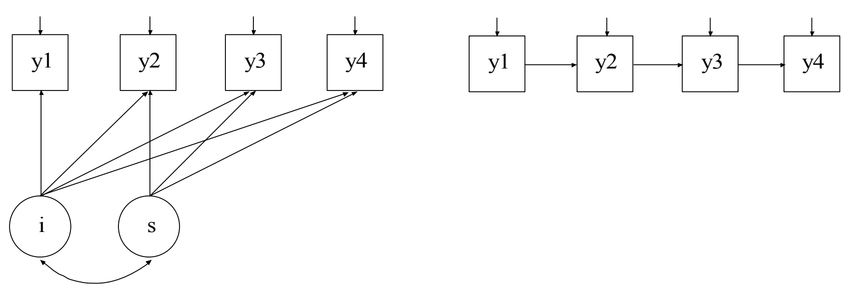
- Advantages of the wide approach:
  - Modeling flexibility
    - Unequal residual variances and covariances
    - Testing of measurement invariance with multiple indicator growth
    - Allowing partial measurement non-invariance
  - Missing data modeling
  - Reduction of the number of levels by one (or more)
- Advantages of the long approach
  - Many time points
  - Individually-varying times of observation with missingness

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- Multiple processes
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

Alternative Models For Longitudinal Data

Growth Curve Model | Auto-Regressive Model



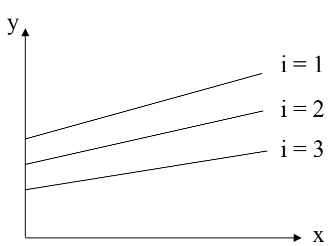
Hybrid Models

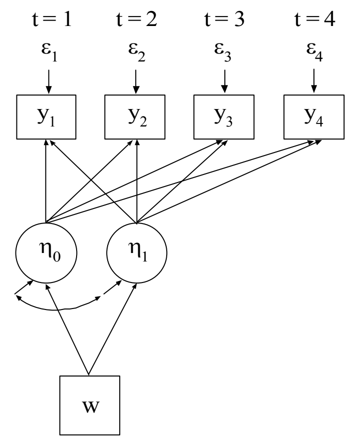
Curran & Bollen (2001)
McArdle & Hamagami (2001)
Bollen & Curran (2006)

The Latent Variable Growth Model In Practice

41

Individual Development Over Time





(1) $y_{ti} = \eta_{0i} + \eta_{1i} x_t + \varepsilon_{ti}$

(2a) $\eta_{0i} = \alpha_0 + \gamma_0 w_i + \zeta_{0i}$

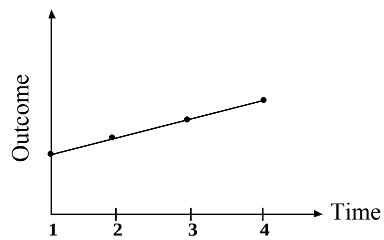
(2b) $\eta_{1i} = \alpha_1 + \gamma_1 w_i + \zeta_{1i}$

42

Specifying Time Scores For Linear Growth Models

Linear Growth Model

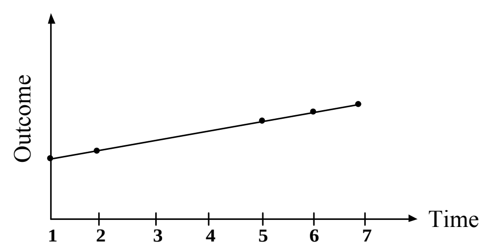
- Need two latent variables to describe a linear growth model: Intercept and slope



- Equidistant time scores for slope:

0	1	2	3	or
0	.1	.2	.3	

Specifying Time Scores For Linear Growth Models (Continued)



- Nonequidistant time scores for slope:

0	1	4	5	6	or
0	.1	.4	.5	.6	

Interpretation Of The Linear Growth Factors

Model:

$$y_{ti} = \eta_{0i} + \eta_{1i} x_t + \varepsilon_{ti} , \quad (17)$$

where in the example t = 1, 2, 3, 4 and x_t = 0, 1, 2, 3:

$$y_{1i} = \eta_{0i} + \eta_{1i}\, 0 + \varepsilon_{1i} , \quad (18)$$

$$\eta_{0i} = y_{1i} - \varepsilon_{1i}, \quad (19)$$

$$y_{2i} = \eta_{0i} + \eta_{1i}\, 1 + \varepsilon_{2i} , \quad (20)$$

$$y_{3i} = \eta_{0i} + \eta_{1i}\, 2 + \varepsilon_{3i} , \quad (21)$$

$$y_{4i} = \eta_{0i} + \eta_{1i}\, 3 + \varepsilon_{4i} . \quad (22)$$

Interpretation Of The Linear Growth Factors (Continued)

Interpretation of the intercept growth factor
η_{0i} (initial status, level):
Systematic part of the variation in the outcome variable at the time point where the time score is zero.

- Unit factor loadings

Interpretation of the slope growth factor
η_{1i} (growth rate, trend):
Systematic part of the increase in the outcome variable for a time score increase of one unit.

- Time scores determined by the growth curve shape

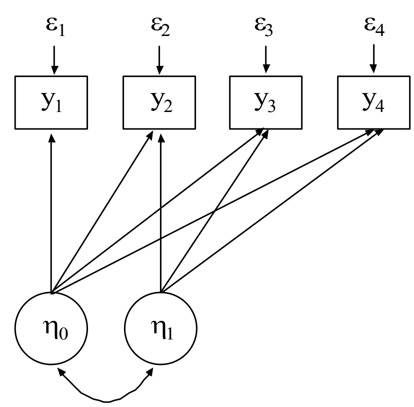
Interpreting Growth Model Parameters

- Intercept Growth Factor Parameters
  - Mean
    - Average of the outcome over individuals at the timepoint with the time score of zero;
    - When the first time score is zero, it is the intercept of the average growth curve, also called initial status
  - Variance
    - Variance of the outcome over individuals at the timepoint with the time score of zero, excluding the residual variance

Interpreting Growth Model Parameters (Continued)

- Linear Slope Growth Factor Parameters
  - Mean – average growth rate over individuals
  - Variance – variance of the growth rate over individuals
  - Covariance with Intercept – relationship between individual intercept and slope values
- Outcome Parameters
  - Intercepts – not estimated in the growth model – fixed at zero to represent measurement invariance
  - Residual Variances – time-specific and measurement error variation
  - Residual Covariances – relationships between time-specific and measurement error sources of variation across time

Latent Growth Model Parameters And Sources Of Model Misfit



Latent Growth Model Parameters For Four Time Points

Linear growth over four time points, no covariates.

Free parameters in the H_1 unrestricted model:

- 4 means and 10 variances-covariances

Free parameters in the H_0 growth model:

(9 parameters, 5 d.f.):

- Means of intercept and slope growth factors
- Variances of intercept and slope growth factors
- Covariance of intercept and slope growth factors
- Residual variances for outcomes

Fixed parameters in the H_0 growth model:

- Intercepts of outcomes at zero
- Loadings for intercept growth factor at one
- Loadings for slope growth factor at time scores
- Residual covariances for outcomes at zero

Latent Growth Model Sources Of Misfit

Sources of misfit:

- Time scores for slope growth factor
- Residual covariances for outcomes
- Outcome variable intercepts
- Loadings for intercept growth factor

Model modifications:

- Recommended
  - Time scores for slope growth factor
  - Residual covariances for outcomes
- Not recommended
  - Outcome variable intercepts
  - Loadings for intercept growth factor

Latent Growth Model Parameters For Three Time Points

Linear growth over three time points, no covariates.

Free parameters in the H_1 unrestricted model:

- 3 means and 6 variances-covariances

Free parameters in the H_0 growth model

(8 parameters, 1 d.f.)

- Means of intercept and slope growth factors
- Variances of intercept and slope growth factors
- Covariance of intercept and slope growth factors
- Residual variances for outcomes

Fixed parameters in the H_0 growth model:

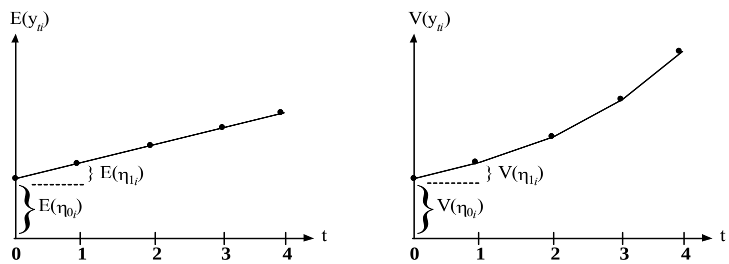
- Intercepts of outcomes at zero
- Loadings for intercept growth factor at one
- Loadings for slope growth factor at time scores
- Residual covariances for outcomes at zero

Growth Model Means And Variances

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,,$$
$$x_t = 0,\ 1,\ \ldots,\ T - 1.$$

Expectation (mean; E) and variance (V):

$$E\,(y_{ti}) = E\,(\eta_{0i}\,) + E\,(\eta_{1i})\, x_t\,,$$
$$V\,(y_{ti}) = V\,(\eta_{0i}\,) + V\,(\eta_{1i})\, x_t^2$$
$$+\ 2x_t\, Cov\,(\eta_{0i}\,,\ \eta_{1i}) + \mathrm{V}\,(\varepsilon_{ti})$$

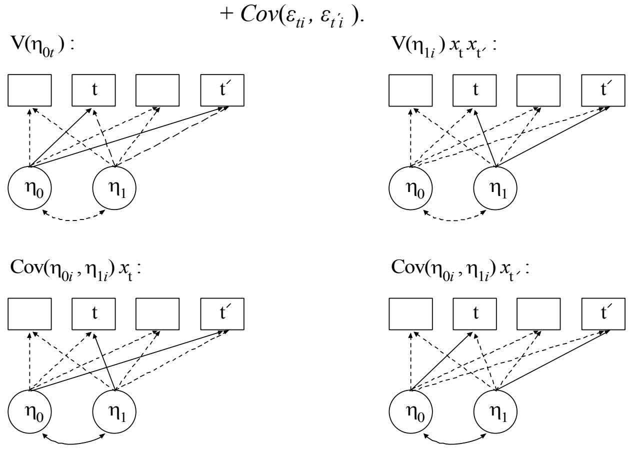


$V(\varepsilon_{ti})$ constant over t
$Cov(\eta_0\,,\ \eta_1) = 0$

Growth Model Covariances

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,,$$
$$x_t = 0,\ 1,\ \ldots,\ T - 1.$$

$$Cov(y_{ti}, y_{t'i}) = V(\eta_{0i}) + V(\eta_{1i})\, x_t\, x_{t'}$$
$$+\ Cov(\eta_{0i}\,,\ \eta_{1i})\ (x_t + x_{t'})$$
$$+\ Cov(\varepsilon_{ti}\,,\ \varepsilon_{t'i}\,).$$



Growth Model Estimation, Testing, And Model Modification

Growth Model Estimation, Testing, And Model Modification

- Estimation: Model parameters
  - Maximum-likelihood (ML) estimation under normality
  - ML and non-normality robust s.e.'s
  - Quasi-ML (MUML): clustered data (multilevel)
  - WLS: categorical outcomes
  - ML-EM: missing data, mixtures
- Model Testing
  - Likelihood-ratio chi-square testing; robust chi square
  - Root mean square of approximation (RMSEA): Close fit ($\leq .05$)
- Model Modification
  - Expected drop in chi-square, EPC
- Estimation: Individual growth factor values (factor scores)
  - Regression method – Bayes modal – Empirical Bayes
  - Factor determinacy

CFA Modeling Estimation And Testing

Estimators

In CFA, a covariance matrix and a mean vector are analyzed.

- ML – minimizes the differences between matrix summaries (determinant and trace) of observed and estimated variances/covariances
- Robust ML – same estimates as ML, standard errors and chi-square robust to non-normality of outcomes and non-independence of observations (MLM, MLR)

Chi-square test of model fit

Tests that the model does not fit significantly worse than a model where the variables correlate freely – p-values greater than or equal to .05 indicate good fit

H_0: Factor model
H_1: Free variance-covariance and mean model
If $p < .05$, H_0 is rejected
Note: We want large p

CFA Modeling Estimation And Testing (Continued)

Model fit indices (cutoff recommendations for good fit based on Yu, 2002 / Hu & Bentler, 1999; see also Marsh et al, 2004)

- CFI – chi-square comparisons of the target model to the baseline model – greater than or equal to .96/.95
- TLI – chi-square comparisons of the target model to the baseline model – greater than or equal to .95/.95
- RMSEA – function of chi-square, test of close fit – less than or equal to .05 (not good at n=100)/.06
- SRMR – average correlation residuals – less than or equal to .07 (not good with binary outcomes)/.08
- WRMR – average weighted residuals – less than or equal to 1.00 (also good with non-normal and categorical outcomes – not good with growth models with many timepoints or multiple group models)

Degrees Of Freedom For Chi-Square Testing Against An Unrestricted Model

The p value of the χ^2 test gives the probability of obtaining a χ^2 value this large or larger if the H_0 model is correct (we want high p values).

Degrees of Freedom:
(Number of parameters in H_1) – (number parameters in H_0)

Number of H_1 parameters with an unrestricted $\boldsymbol{\Sigma}$: $p\,(p + 1)/2$

Number of H_1 parameters with unrestricted $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$:
$p + p\,(p + 1)/2$

Chi-Square Difference Testing Of Nested Models

- When a model H_a imposes restrictions on parameters of model H_b, H_a is said to be nested within H_b
- To test if the nested model H_a fits significantly worse than H_b, a chi-square test can be obtained as the difference in the chi-square values for the two models (testing against an unrestricted model) using as degrees of freedom the difference in number of parameters for the two models
- The chi-square difference is the same as 2 times the difference in log likelihood values for the two models
- The chi-square theory does not hold if H_a has restricted any of the H_b parameters to be on the border of their admissible parameter space (e.g. variance = 0)

CFA Model Modification

Model modification indices are estimated for all parameters that are fixed or constrained to be equal.

- Modification Indices – expected drop in chi-square if the parameter is estimated
- Expected Parameter Change Indices – expected value of the parameter if it is estimated
- Standardized Expected Parameter Change Indices – standardized expected value of the parameter if it is estimated

Model Modifications

- Residual covariances
- Factor cross loadings

Alternative Growth Model Parameterizations

Parameterization 1 – for continuous outcomes

$$y_{ti} = \mathbf{0} + \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}, \qquad (32)$$

$$\eta_{0i} = \boldsymbol{\alpha}_0 + \zeta_{0i}, \qquad (33)$$

$$\eta_{1i} = \alpha_1 + \zeta_{1i}. \qquad (34)$$

Parameterization 2 – for categorical outcomes and multiple indicators

$$y_{ti} = \boldsymbol{v} + \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}, \qquad (35)$$

$$\eta_{0i} = \mathbf{0} + \zeta_{0i}, \qquad (36)$$

$$\eta_{1i} = \alpha_1 + \zeta_{1i}. \qquad (37)$$

Alternative Growth Model Parameterizations

Parameterization 1 – for continuous outcomes

- Outcome variable intercepts fixed at zero
- Growth factor means free to be estimated

```
MODEL:  i BY y1-y4@1;
        s BY y1@0 y2@1 y3@2 y4@3;
        [y1-y4@0 i s];
```

Parameterization 2 – for categorical outcomes and multiple indicators

- Outcome variable intercepts constrained to be equal
- Intercept growth factor mean fixed at zero

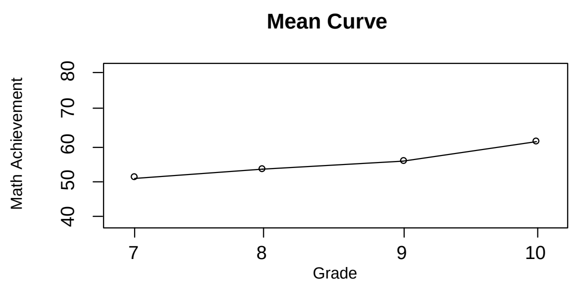
```
MODEL:  i BY y1-y4@1;
        s BY y1@0 y2@1 y3@2 y4@3;
        [y1-y4] (1);
        [i@0 s];
```

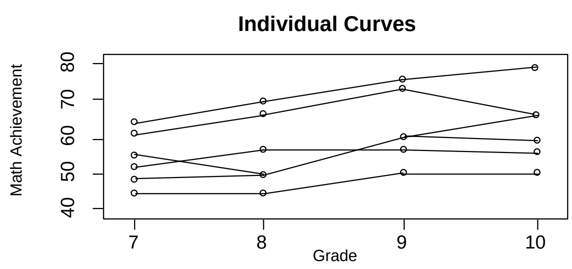

Simple Examples Of Growth Modeling

Steps In Growth Modeling

- Preliminary descriptive studies of the data: means, variances, correlations, univariate and bivariate distributions, outliers, etc.
- Determine the shape of the growth curve from theory and/or data
  - Individual plots
  - Mean plot
- Consider change in variance across time
- Fit model without covariates using fixed time scores
- Modify model as needed
- Add covariates

LSAY Math Achievement





Input For LSAY TYPE=BASIC Analysis

```
TITLE:      Basic run
DATA:       FILE = lsayfull_dropout.dat;
VARIABLE:   NAMES = lsayid schcode female mothed homeres math7
            math8 math9 math10 math11 math12
            mthcrs7 mthcrs8 mthcrs9 mthcrs10 mthcrs11 mthcrs12;
!lsayid =   Student id
!schcode =  7th grade school code
!mothed =   mother's education
!           (1=LT HS diploma, 2=HS diploma, 3=Some college,
!           4=4yr college degree, 5=advanced degree)
!homeres =  Home math and science resources
!mthcrs7-mthcrs12 = Highest math course taken during each grade
!           (0 = no course, 1 – low,basic, 2 = average, 3 = high,
!           4 = pre-algebra, 5 = algebra I, 6 = geometry,
!           7 = algebra II, 8 = pre-calc, 9 = calculus)
ANALYSIS:   TYPE = BASIC;
PLOT:       TYPE = PLOT3;
            SERIES = math7-math10(*);
```


Sample Statistics For LSAY Data

n = 3102

Means

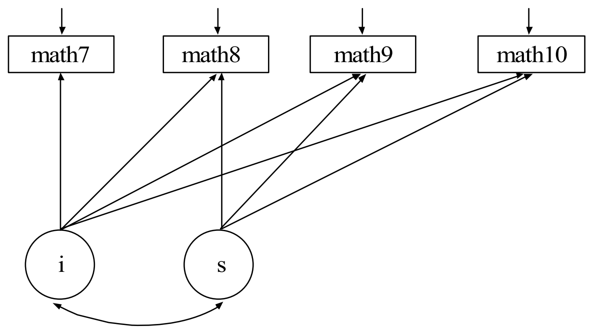
	MATH7	MATH8	MATH9	MATH10
	50.356	53.872	57.962	62.250

Covariances

	MATH7	MATH8	MATH9	MATH10
MATH7	103.868			
MATH8	93.096	121.294		
MATH9	104.328	121.439	161.394	
MATH10	110.003	125.355	157.656	189.096

Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	1.000			
MATH8	0.829	1.000		
MATH9	0.806	0.868	1.000	
MATH10	0.785	0.828	0.902	1.000



Input For LSAY Linear Growth Model Without Covariates

```
TITLE:      Growth 7 – 10, no covariates
DATA:       FILE = lsayfull_dropout.dat;
VARIABLE:   NAMES = lsayid schcode female mothed homeres
            math7 math8 math9 math10 math11 math12
            mthcrs7 mthcrs8 mthcrs9 mthcrs10 mthcrs11 mthcrs12;
            USEV = math7-math10;
            MISSING = ALL(9999);
MODEL:      i BY math7-math10@1;
            s BY math7@0 math8@1 math9@2 math10@3;
            [math7-math10@0];
            [i s];
OUTPUT:     SAMPSTAT STANDARDIZED RESIDUAL MODINDICES (3.84);
Alternative language:
MODEL:      i s | math7@0 math8@1 math9@2 math10@3;
```

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Model Results

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
I BY				
MATH7	1.000	0.000	999.000	999.000
MATH8	1.000	0.000	999.000	999.000
MATH9	1.000	0.000	999.000	999.000
MATH10	1.000	0.000	999.000	999.000
S BY				
MATH7	0.000	0.000	999.000	999.000
MATH8	1.000	0.000	999.000	999.000
MATH9	2.000	0.000	999.000	999.000
MATH10	3.000	0.000	999.000	999.000

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

	Estimates	S.E.	Est./S.E.	Two-tailed P-value
Means				
I	50.202	0.180	279.523	0.000
S	3.939	0.059	66.460	0.000
Intercepts				
MATH7	0.000	0.000	999.000	999.000
MATH8	0.000	0.000	999.000	999.000
MATH9	0.000	0.000	999.000	999.000
MATH10	0.000	0.000	999.000	999.000

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

	Estimates	S.E.	Est./S.E.	Two-tailed P-value
Residual Variances				
MATH7	17.430	1.002	17.400	0.000
MATH8	18.440	0.750	24.596	0.000
MATH9	16.184	0.757	20.561	0.000
MATH10	17.219	1.301	13.230	0.000
Variances				
I	86.159	2.606	33.067	0.000
S	4.792	0.295	16.262	0.000
I WITH				
S	8.031	0.654	12.276	0.000

R-Square

Observed Variable	R-Square
MATH7	0.832
MATH8	0.853
MATH9	0.895
MATH10	0.912

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Tests Of Model Fit

```
Chi-Square Test of Model Fit
          Value                          86.541
          Degrees of Freedom                  5
          P-Value                        0.0000

CFI/TLI
          CFI                            0.992
          TLI                            0.990

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                       0.073
          90 Percent C.I.                0.060    0.086
          Probability RMSEA <= .05       0.002

SRMR (Standardized Root Mean Square Residual)
          Value                          0.047
```

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Modification Indices

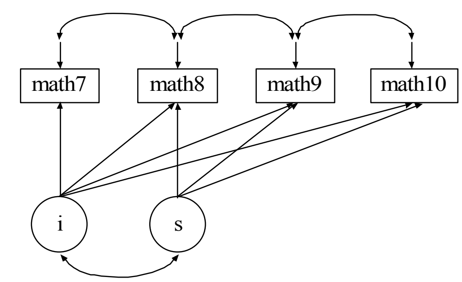
		M.I.	E.P.C.	Std.E.P.C.	StdYX E.P.C.
BY Statements					
I	BY MATH7	18.291	0.013	0.123	0.012
I	BY MATH8	15.115	-0.008	-0.073	-0.006
S	BY MATH7	22.251	0.178	0.389	0.038
S	BY MATH8	24.727	-0.120	-0.263	-0.023
WITH Statements					
MATH9	WITH MATH7	18.449	-2.930	-2.930	-0.174
MATH9	WITH MATH8	31.311	4.767	4.767	0.276
MATH10	WITH MATH7	30.282	5.742	5.742	0.331
MATH10	WITH MATH8	54.842	-6.353	-6.353	-0.357
MATH10	WITH MATH9	31.503	14.816	14.816	0.888

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Modification Indices

	M.I.	E.P.C.	Std.E.P.C.	StdYX E.P.C.
Means/Intercepts/Thresholds				
[MATH7]	18.011	0.671	0.671	0.066
[MATH8]	12.506	-0.362	-.362	-0.032

Linear Growth Model Without Covariates: Adding Correlated Residuals




```
MODEL:
        i s | math7@0 math8@1 math9@2 math10@3;
        math7-math9 PWITH math8-math10;
```


Output Excerpts LSAY Linear Growth Model Without Covariates: Adding Correlated Residuals

Tests Of Model Fit

```
Chi-Square Test of Model Fit
   Value                                  18.519
   Degrees of Freedom                          2
   P-Value                                0.0000

CFI/TLI
   CFI                                    0.998
   TLI                                    0.995

RMSEA (Root Mean Square Error Of Approximation)
   Estimate                               0.052
   90 Percent C.I.                        0.032   0.074
   Probability RMSEA <= .05               0.404

SRMR (Standardized Root Mean Square Residual)
   Value                                  0.011
```

Output Excerpts LSAY: Adding Correlated Residuals (Continued)

	Estimates	S.E.	Est./S.E.	Two-tailed P-value
S WITH				
I	6.133	1.379	4.447	0.000
MATH7 WITH				
MATH8	-5.078	2.146	-2.366	0.018
MATH8 WITH				
MATH9	4.917	0.916	5.365	0.000
MATH9 WITH				
MATH10	17.062	2.983	5.720	0.000
Means				
I	50.203	0.180	279.431	0.000
S	3.936	0.059	66.693	0.000

Output Excerpts LSAY: Adding Correlated Residuals (Continued)

	Estimates	S.E.	Est./S.E.	Two-tailed P-value
Variances				
I	92.038	4.167	22.085	0.000
S	3.043	0.789	3.858	0.000
Residual Variances				
MATH7	11.871	3.466	3.425	0.001
MATH8	14.027	1.980	7.085	0.000
MATH9	32.596	2.609	12.492	0.000
MATH10	33.857	4.815	7.032	0.000

Output Excerpts LSAY: Adding Correlated Residuals (Continued)

ESTIMATED MODEL AND RESIDUALS (OBSERVED – ESTIMATED)

Model Estimated Means/Intercepts/Thresholds

	MATH7	MATH8	MATH9	MATH10
1	50.203	54.140	58.076	62.012

Residuals for Means/Intercepts/Thresholds

	MATH7	MATH8	MATH9	MATH10
1	0.153	-0.267	-0.114	0.238

Standardized Residuals (z-scores) for Means/Intercepts/Thresholds

	MATH7	MATH8	MATH9	MATH10
1	4.198	-4.109	-1.256	5.199

Normalized Residuals for Means/Intercepts/Thresholds

	MATH7	MATH8	MATH9	MATH10
1	0.834	-1.317	-0.478	0.904

Output Excerpts LSAY: Adding Correlated Residuals (Continued)

Model Estimated Covariances/Correlations/Residual Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	103.910			
MATH8	93.093	121.375		
MATH9	104.304	121.441	161.339	
MATH10	110.437	125.700	158.025	190.083

Residuals for Covariances/Correlations/Residual Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	-0.041			
MATH8	0.002	-0.081		
MATH9	0.024	-0.002	0.055	
MATH10	-0.434	-0.345	-0.368	-0.987

Output Excerpts LSAY: Adding Correlated Residuals (Continued)

Standardized Residuals (z-scores) for Covariances/Correlations/Residual Corr

	MATH7	MATH8	MATH9	MATH10
MATH7	999.000			
MATH8	999.000	999.000		
MATH9	0.279	999.000	0.297	
MATH10	999.000	999.000	999.000	999.000

Normalized Residuals for Covariances/Correlations/Residual Correlations

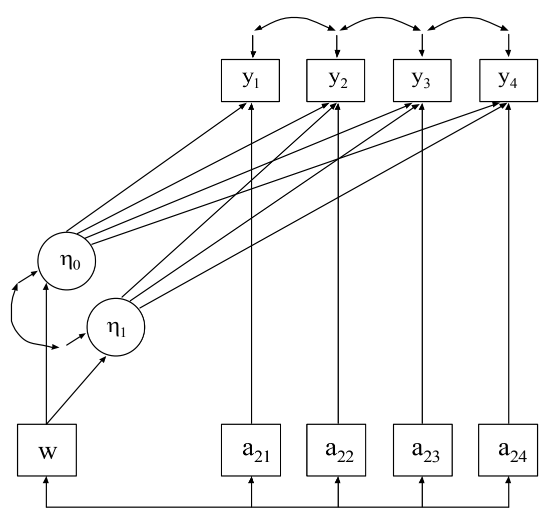
	MATH7	MATH8	MATH9	MATH10
MATH7	-0.016			
MATH8	0.001	-0.025		
MATH9	0.008	-0.001	0.012	
MATH10	-0.130	-0.092	-0.081	-0.185

Covariates In The Growth Model

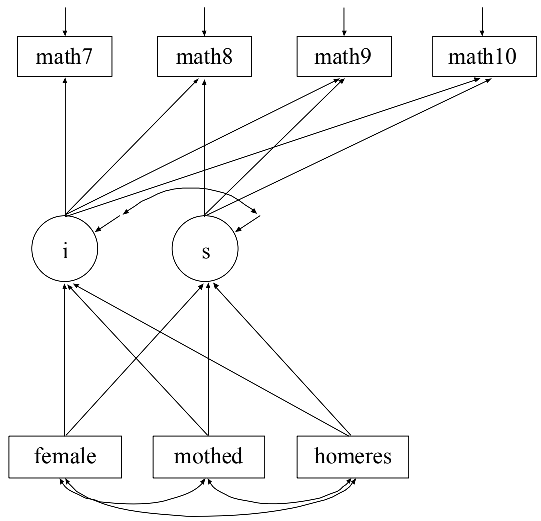
Covariates In The Growth Model

- Types of covariates
  - Time-invariant covariates—vary across individuals not time, explain the variation in the growth factors
  - Time-varying covariates—vary across individuals and time, explain the variation in the outcomes beyond the growth factors

Time-Invariant And Time-Varying Covariates



LSAY Growth Model With Time-Invariant Covariates



Input Excerpts For LSAY Linear Growth Model With Time-Invariant Covariates

```
TITLE:      Growth 7 – 10, no covariates
DATA:       FILE = lsayfull_dropout.dat;
VARIABLE:   NAMES = lsayid schcode female mothed homeres
            math7 math8 math9 math10 math11 math12
            mthcrs7 mthcrs8 mthcrs9 mthcrs10 mthcrs11 mthcrs12;
            MISSING = ALL (999);
            USEVAR = math7-math10 female mothed homeres;
ANALYSIS:   !ESTIMATOR = MLR;
MODEL:      i s | math7@0 math8@1 math9@2 math10@3;
            i s ON female mothed homeres;

Alternative language:
MODEL:      i BY math7-math10@1;
            s BY math7@0 math8@1 math9@2 math10@3;
            [math7-math10@0];
            [i s];
            i s ON female mothed homeres;
```

Output Excerpts LSAY Growth Model With Time-Invariant Covariates

n = 3116

Tests Of Model Fit for ML

```
Chi-Square Test of Model Fit
             Value                              33.611
             Degrees of Freedom                      8
             P-Value                             0.000
CFI/TLI
             CFI                                 0.998
             TLI                                 0.994
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                            0.032
             90 Percent C.I.                     0.021  0.044
             Probability RMSEA <= .05            0.996
SRMR (Standardized Root Mean Square Residual)
             Value                               0.010
```


Output Excerpts LSAY Growth Model With Time-Invariant Covariates (Continued)

Tests Of Model Fit for MLR

```
Chi-Square Test of Model Fit
             Value                              33.290*
             Degrees of Freedom                      8
             P-Value                            0.0001
             Scaling Correction Factor           1.010
                 for MLR
CFI/TLI
             CFI                                 0.997
             TLI                                 0.993
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                            0.015
             90 Percent C.I.                     0.021  0.043
             Probability RMSEA <= .05            0.996
SRMR (Standardized Root Mean Square Residual)
             Value                               0.010
```

Output Excerpts LSAY Growth Model With Time-Invariant Covariates (Continued)

Selected Estimates For ML

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I ON				
FEMALE	2.123	0.327	6.499	0.000
MOTHED	2.262	0.164	13.763	0.000
HOMERES	1.751	0.104	16.918	0.000
S ON				
FEMALE	-0.134	0.116	-1.153	0.249
MOTHED	0.223	0.059	3.771	0.000
HOMERES	0.273	0.037	7.308	0.000

Output Excerpts LSAY Growth Model With Time-Invariant Covariates (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
S WITH				
I	4.131	1.244	3.320	0.001
Residual Variances				
I	71.888	3.630	19.804	0.000
S	3.313	0.724	4.579	0.000
Intercepts				
I	38.434	0.497	77.391	0.000
S	2.636	0.181	14.561	0.000

Output Excerpts LSAY Growth Model With Time-Invariant Covariates (Continued)

R-Square

Observed Variable	R-Square
MATH7	0.876
MATH8	0.863
MATH9	0.817
MATH10	0.854

Latent Variable	R-Square
I	.204
S	.091

Output Excerpts LSAY Growth Model With Time-Invariant Covariates (Continued)

TECHNICAL 4 OUTPUT

ESTIMATES DERIVED FROM THE MODEL

ESTIMATED MEANS FOR THE LATENT VARIABLES

I	S	FEMALE	MOTHED	HOMERES
50.219	3.944	0.478	2.347	3.118

Output Excerpts LSAY Growth Model With Time-Invariant Covariates (Continued)

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	I	S	FEMALE	MOTHED	HOMERES
I	90.264				
S	6.411	3.647			
FEMALE	0.350	-0.058	0.250		
MOTHED	3.226	0.373	-0.024	1.088	
HOMERES	5.901	0.891	-0.071	0.467	2.853

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	I	S	FEMALE	MOTHED	HOMERES
I	1.000				
S	0.353	1.000			
FEMALE	0.074	-0.061	1.000		
MOTHED	0.326	0.187	-0.047	1.000	
HOMERES	0.368	0.276	-0.084	0.265	1.000

Model Estimated Average And Individual Growth Curves With Covariates

Model:

$$y_{ti} = \eta_{0i} + \eta_{1i} x_t + \varepsilon_{ti} , \quad (23)$$

$$\eta_{0i} = \alpha_0 + \gamma_0 w_i + \zeta_{0i} , \quad (24)$$

$$\eta_{1i} = \alpha_1 + \gamma_1 w_i + \zeta_{1i} , \quad (25)$$

Estimated growth factor means:

$$\hat{E}(\eta_{0i}) = \hat{\alpha}_0 + \hat{\gamma}_0 \overline{w} , \quad (26)$$

$$\hat{E}(\eta_{1i}) = \hat{\alpha}_1 + \hat{\gamma}_1 \overline{w} . \quad (27)$$

Estimated outcome means:

$$\hat{E}(y_{ti}) = \hat{E}(\eta_{0i}) + \hat{E}(\eta_{1i}) x_t . \quad (28)$$

Estimated outcomes for individual *i* :

$$\hat{y}_{ti} = \hat{\eta}_{0i} + \hat{\eta}_{1i} x_t \quad (29)$$

where $\hat{\eta}_{0i}$ and $\hat{\eta}_{1i}$ are estimated factor scores. $\hat{y}_{ti}$ can be used for prediction purposes.

Model Estimated Means With Covariates

Model estimated means are available using the TECH4 and RESIDUAL options of the OUTPUT command.

Estimated Intercept Mean = Estimated Intercept +
Estimated Slope (Female)*Sample Mean (Female) +
Estimated Slope (Mothed)*Sample Mean (Mothed) +
Estimated Slope (Homeres)*Sample Mean (Homeres)

38.43 + 2.12*0.48 + 2.26*2.35 + 1.75*3.12 = 50.22

Estimated Slope Mean = Estimated Intercept +
Estimated Slope (Female)*Sample Mean (Female) +
Estimated Slope (Mothed)*Sample Mean (Mothed) +
Estimated Slope (Homeres)*Sample Mean (Homeres)

2.64 – 0.13*0.48 + 0.22*2.35 + 0.27*3.11 = 3.94

Model Estimated Means With Covariates (Continued)

Estimated Outcome Mean at Timepoint t =

Estimated Intercept Mean +
Estimated Slope Mean * (Time Score at Timepoint t)

Estimated Outcome Mean at Timepoint 1 =
50.22 + 3.94 * (0) = **50.22**

Estimated Outcome Mean at Timepoint 2 =
50.22 + 3.94 * (1.00) = **54.16**

Estimated Outcome Mean at Timepoint 3 =
50.22 + 3.94 * (2.00) = **58.11**

Estimated Outcome Mean at Timepoint 4 =
50.22 + 3.94 * (3.00) = **62.05**

Centering

Centering

- Centering determines the interpretation of the intercept growth factor
- The centering point is the timepoint at which the time score is zero
- A model can be estimated for different centering points depending on which interpretation is of interest
- Models with different centering points give the same model fit because they are reparameterizations of the model
- Changing the centering point in a linear growth model with four timepoints

	1	2	3	4	
Timepoints	1	2	3	4	Centering at
Time scores	0	1	2	3	Timepoint 1
	-1	0	1	2	Timepoint 2
	-2	-1	0	1	Timepoint 3
	-3	-2	-1	0	Timepoint 4

Input Excerpts For LSAY Growth Model With Covariates Centered At Grade 10

```
MODEL:      i s | math7@-3 math8@-2 math9@-1 math10@0;
            i s ON female mothed homeres;
            math7-math9 PWITH math8-math10;
OUTPUT:     TECH1 RESIDUAL STANDARDIZED MODINDICES TECH4;
```

Alternative language:

```
MODEL:      i BY math7-math10@1;
            s BY math7@-3 math8@-2 math9@-1 math10@0;
            math7-math9 PWITH math8-math10;
            [math7-math10@0];
            [i s];
            i s ON female mothed homeres;
```


Output Excerpts LSAY Growth Model With Covariates Centered At Grade 10

n = 3116

Tests of Model Fit

```
CHI-SQUARE TEST OF MODEL FIT

          Value                        33.611
          Degrees of Freedom                8
          P-Value                       0.000

RMSEA (ROOT MEAN SQUARE ERROR OF APPROXIMATION)

          Estimate                      .032
          90 Percent C.I.               .021    .044
          Probability RMSEA <= .05      .996
```

Output Excerpts LSAY Growth Model With Covariates Centered At Grade 10 (Continued)

SELECTED ESTIMATES

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I ON				
FEMALE	1.723	0.473	3.643	0.000
MOTHED	2.930	0.239	12.249	0.000
HOMERES	2.569	0.151	17.002	0.000
S ON				
FEMALE	-0.133	0.116	-1.153	0.249
MOTHED	0.223	0.059	3.771	0.000
HOMERES	0.273	0.037	7.308	0.000

Further Readings On Introductory Growth Modeling

Bijleveld, C. C. J. H., & van der Kamp, T. (1998). Longitudinal data analysis: Designs, models, and methods. Newbury Park: Sage.

Bollen, K.A. & Curran, P.J. (2006). Latent curve models. A structural equation perspective. New York: Wiley.

Duncan, T., Duncan S. & Strycker, L. (2006). An introduction to latent variable growth curve modeling. Second edition. Lawrence Erlbaum: New York.

Muthén, B. & Khoo, S.T. (1998). Longitudinal studies of achievement growth using latent variable modeling. Learning and Individual Differences, Special issue: latent growth curve analysis, 10, 73-101. (#80)

Muthén, B. & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol, 61, 290-300. (#83)

Further Readings On Introductory Growth Modeling (Continued)

Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods. Second edition. Newbury Park, CA: Sage Publications.

Singer, J.D. & Willett, J.B. (2003). Applied longitudinal data analysis. Modeling change and event occurrence. New York, NY: Oxford University Press.

Snijders, T. & Bosker, R. (1999). Multilevel analysis. An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.

Non-Linear Growth

Six Ways To Model Non-Linear Growth

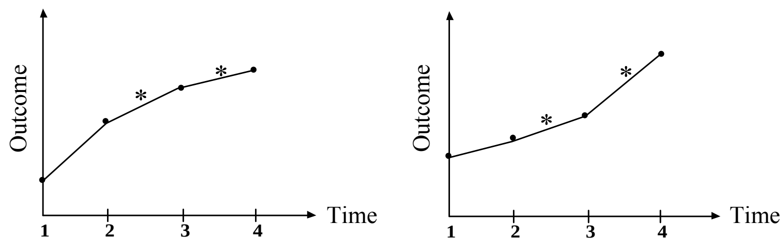
- Estimated time scores
- Quadratic (cubic) growth model
- Fixed non-linear time scores
- Piecewise growth modeling
- Time-varying covariates
- Non-linearity of random effects

Growth Model With Free Time Scores

Specifying Time Scores For Non-Linear Growth Models With Estimated Time Scores

Non-linear growth models with estimated time scores

- Need two latent variables to describe a non-linear growth model: Intercept and slope



Time scores: 0 1 Estimated Estimated

Input Excerpts For LSAY Linear Growth Model With Free Time Scores Without Covariates

```
MODEL:     i s | math7@0 math8@1 math9@2 math10@3 math11@4
           math12*5;

Alternative language:

MODEL:     i BY math7-math12@1;
           s BY math7@0 math8@1 math9@2 math10@3 math11@4
           math12*5;
           [math7-math12@0];
           [i s];
```

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates

n = 3102

Tests Of Model Fit

```
Chi-Square Test of Model Fit
             Value                                121.095
             Degrees of Freedom                        10
             P-Value                               0.0000
CFI/TLI
             CFI                                    0.992
             TLI                                    0.989
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                               0.060
             90 Percent C.I.                        0.051   0.070
             Probability RMSEA <= .05               0.041
SRMR (Standardized Root Mean Square Residual)
             Value                                  0.034
```


Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I |				
MATH7	1.000	0.000	999.000	999.000
MATH8	1.000	0.000	999.000	999.000
MATH9	1.000	0.000	999.000	999.000
MATH10	1.000	0.000	999.000	999.000
MATH11	1.000	0.000	999.000	999.000
MATH12	1.000	0.000	999.000	999.000
S |				
MATH7	0.000	0.000	999.000	999.000
MATH8	1.000	0.000	999.000	999.000
MATH9	2.000	0.000	999.000	999.000
MATH10	3.000	0.000	999.000	999.000

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
MATH11	4.000	0.000	999.000	999.000
MATH12	4.095	0.042	97.236	0.000
S WITH				
I	4.986	0.741	6.725	0.000
Variances				
I	91.374	3.046	29.994	0.000
S	4.001	0.276	14.666	0.000
Means				
I	50.323	0.180	279.612	0.000
S	3.752	0.049	76.472	0.000

Growth Model With Free Time Scores

- Identification of the model – for a model with two growth factors, at least one time score must be fixed to a non-zero value (usually one) in addition to the time score that is fixed at zero (centering point)
- Interpretation—cannot interpret the mean of the slope growth factor as a constant rate of change over all timepoints, but as the rate of change for a time score change of one.
- Approach—fix the time score following the centering point at one

Interpretation Of Slope Growth Factor Mean For Non-Linear Models

- The slope growth factor mean is the expected change in the outcome variable for a one unit change in the time score
- In non-linear growth models, the time scores should be chosen so that a one unit change occurs between timepoints of substantive interest.
  - An example of 4 timepoints representing grades 7, 8, 9, and 10
    - Time scores of 0 1 * * – slope factor mean refers to expected change between grades 7 and 8
    - Time scores of 0 * * 1 – slope factor mean refers to expected change between grades 7 and 10

Specifying Time Scores For Quadratic Growth Models

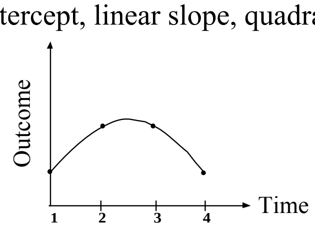
Quadratic growth model

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \eta_{2i}\; x_t^2 + \varepsilon_{ti} \qquad \text{or}$$

$$y_{ti} = \eta_{0i} + \eta_{1i} \cdot (x_t - c) + \eta_{2i} \cdot (x_t - c)^2$$

where c is a centering constant, e.g. $\bar{x}$

- Need three latent variables to describe a quadratic growth model: Intercept, linear slope, quadratic slope



- Linear slope time scores: 0 1 2 3 or 0 .1 .2 .3
- Quadratic slope time scores: 0 1 4 9 or 0 .01 .04 .09

Mplus Specification Of Several Growth Factors

- Quadratic:

```
i s q | y1@0 y2@1 y3@2 y4@3;
```

or alternatively

```
i BY y1-y4@1;
s BY y1@0 y2@1 y3@2 y4@3;
q BY y1@0 y2@1 y3@4 y4@9;
```

- Cubic

```
i s q c | y1@0 y2@1 y3@2 y4@3;
```

- Intercept only

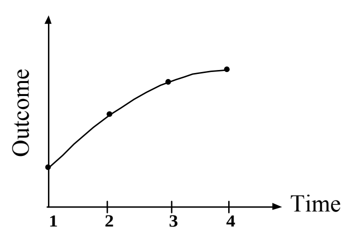
```
i | y1-y4;
```


Specifying Time Scores For Non-Linear Growth Models With Fixed Time Scores

Non-Linear Growth Models with Fixed Time scores

- Need two latent variables to describe a non-linear growth model: Intercept and slope

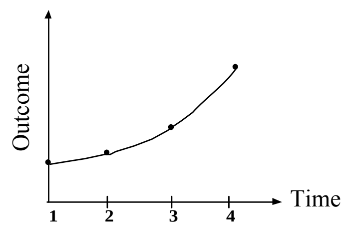
Growth model with a logarithmic growth curve--ln(t)



Time scores: 0 0.69 1.10 1.39

Specifying Time Scores For Non-Linear Growth Models With Fixed Time Scores (Continued)

Growth model with an exponential growth curve–
exp(t-1) - 1

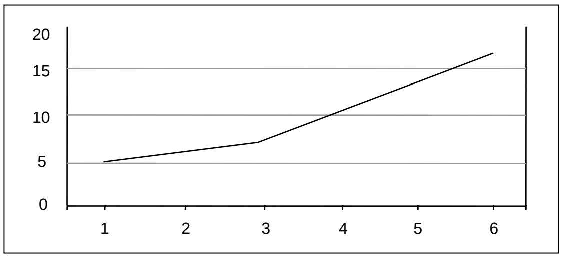


Time scores: 0 1.72 6.39 19.09

Piecewise Growth Modeling

Piecewise Growth Modeling

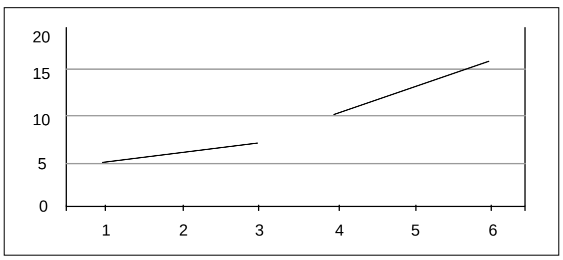
- Can be used to represent different phases of development
- Can be used to capture non-linear growth
- Each piece has its own growth factor(s)
- Each piece can have its own coefficients for covariates



One intercept growth factor, two slope growth factors

s1:	0	1	2	2	2	2	Time scores piece 1
s2:	0	0	0	1	2	3	Time scores piece 2

Piecewise Growth Modeling (Continued)

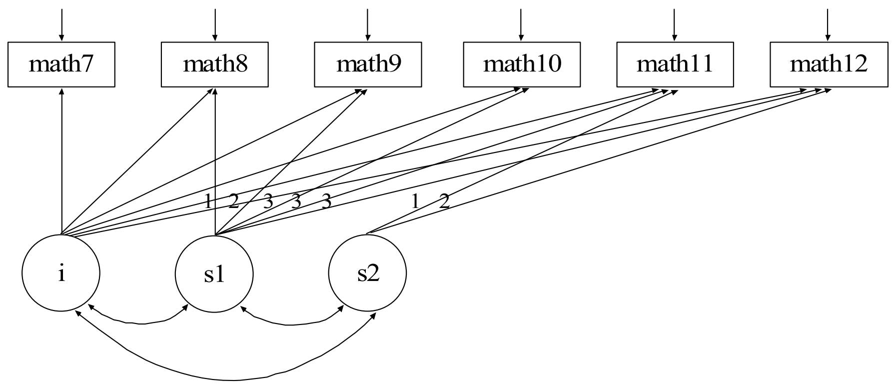


Two intercept growth factors, two slope growth factors

0	1	2				Time scores piece 1
			0	1	2	Time scores piece 2

Sequential model

LSAY Piecewise Linear Growth Modeling: Grades 7-10 and 10-12



Input For LSAY Piecewise Growth Model With Covariates

```
MODEL:     i s1 | math7@0 math8@1 math9@2 math10@3 math11@3
           math12@3;
           i s2 | math7@0 math8@0 math9@0 math10@0 math11@1
           math12@2;
           i s1 s2 ON female mothed homeres;
```

```
Alternative language:
MODEL:     i BY math7-math12@1;
           s1 BY math7@0 math8@1 math9@2 math10@3 math11@3
           math12@3;
           s2 BY math7@0 math8@0 math9@0 math10@0 math11@1
           math12@2;
           [math7-math12@0];
           [i s1 s2];
           i s1 s2 ON female mothed homeres;
```

Output Excerpts LSAY Piecewise Growth Model With Covariates

n = 3116

Tests of Model Fit

```
CHI-SQUARE TEST OF MODEL FIT

          Value                        229.22
          Degrees of Freedom               21
          P-Value                      0.0000

RMSEA (ROOT MEAN SQUARE ERROR OF APPROXIMATION)

          Estimate                      0.056
          90 Percent C.I.               0.050  0.063
          Probability RMSEA <= .05      0.051
```


Output Excerpts LSAY Piecewise Growth Model With Covariates (Continued)

SELECTED ESTIMATES

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I ON				
FEMALE	2.126	0.327	6.496	0.000
MOTHED	2.282	0.165	13.867	0.000
HOMERES	1.757	0.104	16.953	0.000
S1 ON				
FEMALE	-0.121	0.114	-1.065	0.287
MOTHED	0.216	0.058	3.703	0.000
HOMERES	0.269	0.037	7.325	0.000
S2 ON				
FEMALE	-0.178	0.191	-0.935	0.350
MOTHED	0.071	0.099	0.719	0.472
HOMERES	0.047	0.061	0.758	0.449

Intermediate Growth Models

Growth Model With Individually-Varying Times Of Observation And Random Slopes For Time-Varying Covariates

Growth Modeling In Multilevel Terms

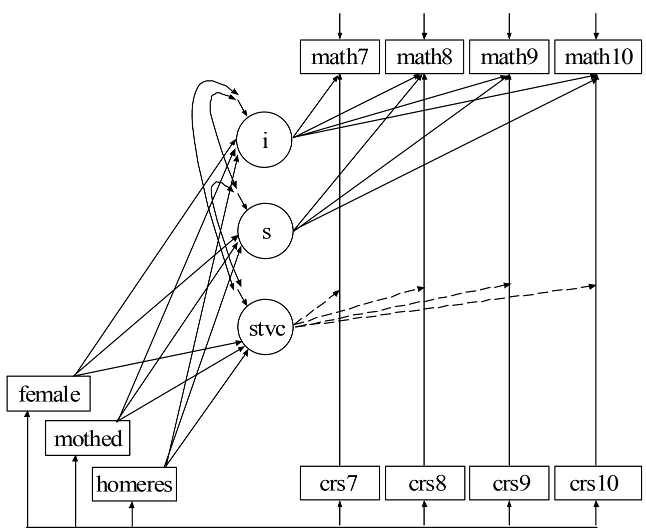
Time point *t*, individual *i* (two-level modeling, no clustering):

y_{ti} : repeated measures of the outcome, e.g. math achievement
a_{1ti} : time-related variable; e.g. grade 7-10
a_{2ti} : time-varying covariate, e.g. math course taking
x_i : time-invariant covariate, e.g. grade 7 expectations

Two-level analysis with individually-varying times of observation and random slopes for time-varying covariates:

$$Level\ 1: y_{ti} = \pi_{0i} + \pi_{1i}\, a_{1ti} + \pi_{2ti}\, a_{2ti} + e_{ti}, \tag{55}$$

$$Level\ 2: \begin{cases} \pi_{0i} = \beta_{00} + \beta_{01}\, x_i + r_{0i}, \\ \pi_{1i} = \beta_{10} + \beta_{11}\, x_i + r_{1i}, \\ \pi_{2i} = \beta_{20} + \beta_{21}\, x_i + r_{2i}. \end{cases} \tag{56}$$



Input For Growth Model With Individually Varying Times Of Observation

```
TITLE:      Growth model with individually varying times of
            observation and random slopes
DATA:       FILE IS lsaynew.dat;
            FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:   NAMES ARE math7 math8 math9 math10 crs7 crs8 crs9
            crs10 female mothed homeres a7-a10;

            ! crs7-crs10 = highest math course taken during each
            ! grade (0=no course, 1=low, basic, 2=average, 3=high.
            ! 4=pre-algebra, 5=algebra I, 6=geometry,
            ! 7=algebra II, 8=pre-calc, 9=calculus)

            MISSING ARE ALL (9999);
            CENTER = GRANDMEAN (crs7-crs10 mothed homeres);
            TSCORES = a7-a10;
```


Input For Growth Model With Individually Varying Times Of Observation (Continued)

```
DEFINE:     math7 = math7/10;
            math8 = math8/10;
            math9 = math9/10;
            math10 = math10/10;
ANALYSIS:   TYPE = RANDOM MISSING;
            ESTIMATOR = ML;
            MCONVERGENCE = .001;
MODEL:      i s | math7-math10 AT a7-a10;
            stvc | math7 ON crs7;
            stvc | math8 ON crs8;
            stvc | math9 ON crs9;
            stvc | math10 ON crs10;
            i ON female mothed homeres;
            s ON female mothed homeres;
            stvc ON female mothed homeres;
            i WITH s;
            stvc WITH i;
            stvc WITH s;
OUTPUT:     TECH8;
```

Output Excerpts For Growth Model With Individually Varying Times Of Observation And Random Slopes For Time-Varying Covariates

n = 2271

Tests of Model Fit

```
Loglikelihood

          H0 Value                     -8199.311

Information Criteria

          Number of Free Parameters          22
          Akaike (AIC)                 16442.623
          Bayesian (BIC)               16568.638
          Sample-Size Adjusted BIC     16498.740
             (n* = (n + 2) / 24)
```


Output Excerpts For Growth Model With Individually Varying Times Of Observation And Random Slopes For Time-Varying Covariates (Continued)

Model Results	Estimates	S.E.	Est./S.E.
I ON			
FEMALE	0.187	0.036	5.247
MOTHED	0.187	0.018	10.231
HOMERES	0.159	0.011	14.194
S ON			
FEMALE	-0.025	0.012	-2.017
MOTHED	0.015	0.006	2.429
HOMERES	0.019	0.004	4.835
STVC ON			
FEMALE	-0.008	0.013	-0.590
MOTHED	0.003	0.007	0.429
HOMERES	0.009	0.004	2.167
I WITH			
S	0.038	0.006	6.445
STVC WITH			
I	0.011	0.005	2.087
S	0.004	0.002	2.033

Output Excerpts For Growth Model With Individually Varying Times Of Observation And Random Slopes For Time-Varying Covariates (Continued)

Intercepts			
MATH7	0.000	0.000	0.000
MATH8	0.000	0.000	0.000
MATH9	0.000	0.000	0.000
MATH10	0.000	0.000	0.000
I	4.992	0.025	198.456
S	0.417	0.009	47.275
STVC	0.113	0.010	11.416
Residual Variances			
MATH7	0.185	0.011	16.464
MATH8	0.178	0.008	22.232
MATH9	0.156	0.008	18.497
MATH10	0.169	0.014	12.500
I	0.570	0.023	25.087
S	0.036	0.003	12.064
STVC	0.012	0.002	5.055

Why No Chi-Square With Random Slopes For Random Variables?

Consider as an example individually-varying times of observation a_{1ti}:

$$y_{ti} = \pi_{0i} + \pi_{1i}\, a_{1ti} + e_{ti}$$

$$V(y_{ti} \mid a_{1ti}) = V(\pi_{0i}) + V(\pi_{1i})\, a_{1ti}^2 + 2\, a_{1ti}\, Cov(\pi_{0i}, \pi_{1i}) + V(e_{ti})$$

The variance of y changes as a function of a_{1ti} values.

Not a constant Σ to test the model fit for.

Maximum-Likelihood Alternatives

Note that [y, x] = [y | x] * [x], where the marginal distribution [x] is unrestricted.

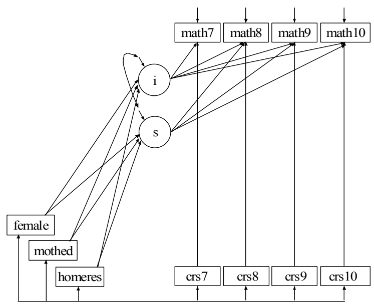
Normal theory ML for

- [y, x]: Gives the same results as [y | x] when there is no missing data (Joreskog & Goldberger, 1975). Typically used in SEM
  - With missing data on x, the normality assumption for x is an additional assumption not used with [y | x]
- [y | x]: Makes normality assumptions for residuals, not for x. Typically used outside SEM
  - Used with Type = Random, Type = Mixture, and with categorical, censored, and count outcomes
  - Deletes individuals with missing on any x
- [y, x] versus [y | x] gives different sample sizes and the likelihood and BIC values are not on a comparable scale

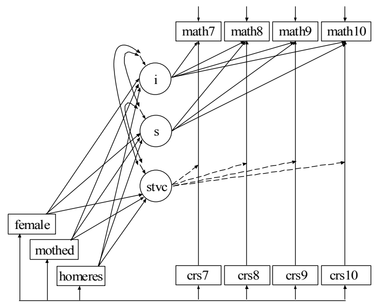
Alternative Models With Time-Varying Covariates

Alternative Models With Time-Varying Covariates

Model M1



Model M2



Input Excerpts Model M1

```
ANALYSIS:   TYPE = RANDOM; ! gives loglikelihood in [y | x] metric

MODEL:      i s | math7@0 math8@1 math9@2 math10@3;
            i s ON female mothed homeres;
            math7 ON mthcrs7;
            math8 ON mthcrs8;
            math9 ON mthcrs9;
            math10 ON mthcrs10;
```

Output Excerpts Model M1

```
TESTS OF MODEL FIT

Chi-Square Test of Model Fit
          Value                              1143.173*
          Degrees of Freedom                        23
          P-Value                                0.000
          Scaling Correlation Factor             1.058
             for MLR
*  The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and
WLSMV cannot be used for chi-square difference tests.  MLM,
MLR and WLSM chi-square difference testing is described in
the Mplus Technical Appendices at www.statmodel.com.  See
chi-square difference testing in the index of the Mplus
User's Guide.
```


Output Excerpts Model M1 (Continued)

```
Chi-Square Test of Model Fit for the Baseline Model
          Value                              8680.167
          Degrees of Freedom                       34
          P-Value                               0.000
CFI/TLI
          CFI                                   0.870
          TLI                                   0.808
Loglikelihood
          H0 Value                         -26869.760
          H0 Scaling Correlation Factor         1.159
            for MLR
          H1 Value                         -26264.830
          H1 Scaling Correlation Factor         1.104
            for MLR
```

Output Excerpts Model M1 (Continued)

```
Information Criteria
         Number of Free Parameters                            19
         Akaike (AIC)                                  53777.520
         Bayesian (BIC)                                53886.351
         Sample-Size Adjusted BIC                      53825.985

                      (n* = (n = 2) / 24)

RMSEA (Root Mean Square Error of Approximation)
         Estimate                                          0.146

SRMR (Standardized Root Mean Square Residual)
         Value                                             0.165
```


Output Excerpts Model M1 (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I ON				
FEMALE	1.877	0.357	5.261	0.000
MOTHED	1.926	0.203	9.497	0.000
HOMERES	1.608	0.113	14.181	0.000
S ON				
FEMALE	-0.236	0.125	-1.893	0.058
MOTHED	0.167	0.066	2.545	0.011
HOMERES	0.193	0.042	4.556	0.000
MATH7 ON				
MTHCRS7	1.042	0.157	6.644	0.000
MATH8 ON				
MTHCRS8	0.898	0.102	8.794	0.000

Output Excerpts Model M1 (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
MATH9 ON				
MTHCRS9	0.929	0.087	10.638	0.000
MATH10 ON				
MTHCRS10	0.911	0.102	8.966	0.000
S WITH				
I	4.200	0.687	6.113	0.000
Intercepts				
MATH7	0.000	0.000	999.000	999.000
MATH8	0.000	0.000	999.000	999.000
MATH9	0.000	0.000	999.000	999.000
MATH10	0.000	0.000	999.000	999.000
I	50.063	0.263	190.158	0.000
S	4.202	0.096	43.621	0.000

Output Excerpts Model M1 (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Residual Variances				
MATH7	18.640	1.341	13.895	0.000
MATH8	18.554	1.002	18.518	0.000
MATH9	16.672	1.010	16.501	0.000
MATH10	17.795	1.671	10.651	0.000
I	58.919	2.393	24.622	0.000
S	3.800	0.359	10.581	0.000

Output Excerpts Model M1 (Continued)

```
MODIFICATION INDICES
Minimum M.I. value for printing the modification index     10.000

                                  M.I.       E.P.C.
ON/BY Statements
MATH 7    ON I         /
I         BY MATH7              15.393       0.014
MATH7     ON S         /
S         BY MATH7              16.813       0.172
MATH8     ON I         /
I         BY MATH8              11.067      -0.008
MATH8     ON S         /
S         BY MATH8              15.769      -0.107
S         ON I         /
I         BY S                 999.000       0.000
```


Output Excerpts Model M1 (Continued)

```
                                  M.I.       E.P.C.
ON Statements
I         ON MATH7              60.201      0.718
I         ON MATH8              58.550      0.464
I         ON MATH9             116.447      0.600
I         ON MATH10            118.956      0.786
I         ON MTHCRS7           582.970      5.844
I         ON MTHCRS8           373.181      3.119
I         ON MTHCRS9           475.187      2.540
I         ON MTHCRS10          379.535      2.012
S         ON MATH7              55.444      0.298
S         ON MATH9             118.064      0.322
S         ON MATH10             24.355      0.221
S         ON MTHCRS7           203.710      1.334
S         ON MTHCRS8            86.109      0.543
```

Output Excerpts Model M1 (Continued)

		M.I.	E.P.C.
S	ON MTHCRS9	90.560	0.453
S	ON MTHCRS10	118.478	0.559
MATH7	ON MATH7	15.393	0.014
MATH7	ON MATH8	17.359	0.013
MATH7	ON MATH9	14.805	0.011
MATH7	ON MATH10	18.991	0.012
MATH7	ON MTHCRS8	48.865	0.873
MATH7	ON MTHCRS9	63.490	0.676
MATH7	ON MTHCRS10	22.160	0.337
MATH8	ON MATH8	11.438	-0.007
MATH8	ON MATH10	13.204	-0.007
MATH8	ON MTHCRS7	82.739	1.467
MATH8	ON MTHCRS9	12.743	0.321

Output Excerpts Model M1 (Continued)

		M.I.	E.P.C.
MATH9	ON MTHCRS7	26.183	0.776
MATH9	ON MTHCRS8	16.027	0.494
MATH9	ON MTHCRS10	69.480	0.781
MATH10	ON MTHCRS8	19.665	0.629
MATH10	ON MTHCRS9	48.678	0.911

Alternative Models
With Time-Varying Covariates

Model	Loglikelihood	# of parameters	BIC
M1	-26,870	19	53,886
M2	-26,846	22	53,861
M3	-26,463	26	53,127

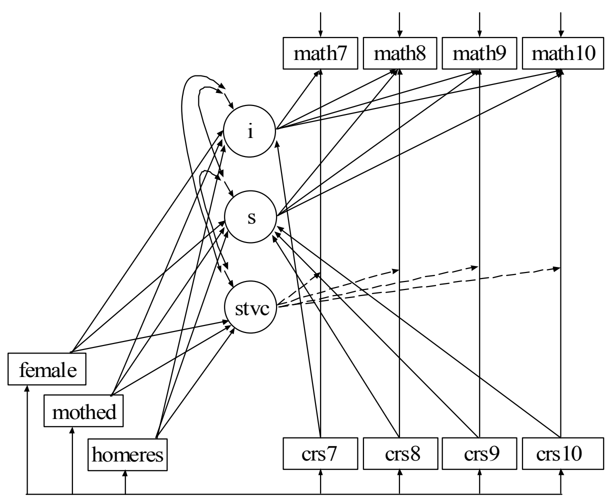
n = 2271 (using [y|x] approach)

M1: Fixed slopes for TVCs, varying across grade

M2: Random slope for TVCs, same across grade

M3: M2 + i and s regressed on TVCs (see model diagram)

Time-Varying Covariates: Model M3



Input Excerpt Model M3

```
ANALYSIS:     TYPE = RANDOM;

MODEL:        i s | math7@0 math8@1 math9@2 math10@3;
              stvc | math7 ON mthcrs7;
              stvc | math8 ON mthcrs8;
              stvc | math9 ON mthcrs9;
              stvc | math10 ON mthcrs10;
              stvc WITH i s;
              i s stvc ON female mothed homeres;
              i ON mthcrs7;
              s ON mthcrs8-mthcrs10;
```


Output Excerpts Time-Varying Covariates: Model M3

	Estimate	**S.E.**	**Est./S.E.**	**Two-Tailed P-Value**
I ON				
FEMALE	1.444	0.325	4.449	0.000
MOTHED	1.259	0.184	6.860	0.000
HOMERES	1.144	0.104	11.041	0.000
MTHCRS7	5.095	0.188	27.040	0.000
S ON				
FEMALE	-0.395	0.123	-3.215	0.001
MOTHED	-0.018	0.064	-0.283	0.777
HOMERES	0.052	0.042	1.249	0.212
MTHCRS8	0.099	0.061	1.627	0.104
MTHCRS9	0.254	0.061	4.188	0.000
MTHCRS10	0.341	0.053	6.471	0.000

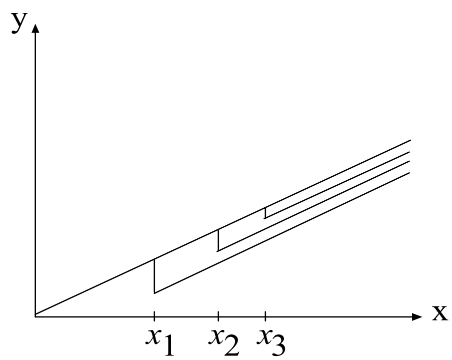
Output Excerpts: Model M3 (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
STVC ON				
FEMALE	-0.083	0.123	-0.677	0.499
MOTHED	0.009	0.066	0.129	0.898
HOMERES	0.070	0.041	1.710	0.087
STVC WITH				
I	-0.078	0.453	-0.173	0.863
S	0.015	0.185	0.083	0.934
S WITH				
I	0.480	0.630	0.762	0.446
Intercepts				
MATH7	0.000	0.000	999.000	999.000
MATH8	0.000	0.000	999.000	999.000

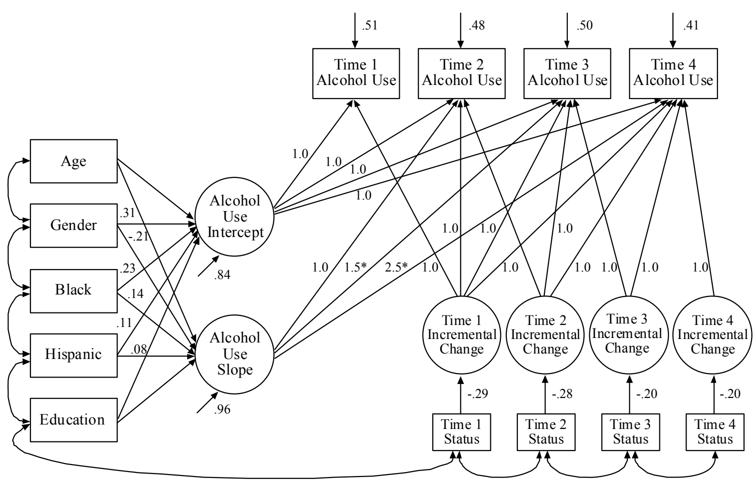
Output Excerpts: Model M3 (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
MATH9	0.000	0.000	999.000	999.000
MATH10	0.000	0.000	999.000	999.000
I	50.244	0.240	209.085	0.000
S	4.257	0.094	45.071	0.000
STVC	0.231	0.106	2.188	0.029
Residual Variances				
MATH7	18.968	1.304	14.541	0.000
MATH8	17.061	0.931	18.322	0.000
MATH9	15.624	0.936	16.690	0.000
MATH10	16.550	1.494	11.074	0.000
I	44.980	1.891	23.792	0.000
S	3.423	0.338	10.118	0.000
STVC	0.615	0.255	2.410	0.016

Time-Varying Covariates Representing Status Change



Marital Status Change And Alcohol Use (Curran, Muthen, & Harford, 1998)



Input Excerpts Marital Status Change And Alcohol Use

```
MODEL:      i s | alcuse1@0 alcuse2@1 alcuse3@2 alcuse4@3;
            i s ON age gender black hispanic education;

            f1 BY alcuse1-alcuse4@1;
            f2 BY alcuse2-alcuse4@1;
            f3 BY alcuse3-alcuse4@1;
            f4 BY alcuse4@1;
            f1-f4@0;

            f1 ON status1;
            f2 ON status2;
            f3 ON status3;
            f4 ON status4;
```


Computational Issues For Growth Models

- Decreasing variances of the observed variables over time may make the modeling more difficult
- Scale of observed variables – keep on a similar scale
- Convergence – often related to starting values or the type of model being estimated
  - Program stops because maximum number of iterations has been reached
    - If no negative residual variances, either increase the number of iterations or use the preliminary parameter estimates as starting values
    - If there are large negative residual variances, try better starting values
  - Program stops before the maximum number of iterations has been reached
    - Check if variables are on a similar scale
    - Try new starting values
- Starting values – the most important parameters to give starting values to are residual variances and the intercept growth factor mean
- Convergence for models using the | symbol
  - Non-convergence may be caused by zero random slope variances which indicates that the slopes should be fixed rather than random

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- **Regressions among random effects**
- Multiple processes
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

Regressions Among Random Effects

Regressions Among Random Effects

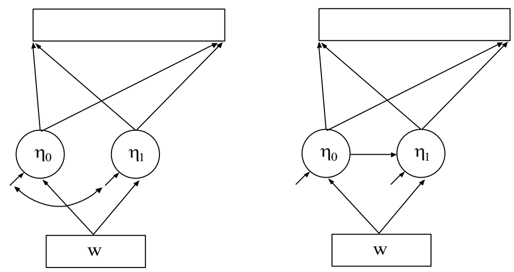
Standard multilevel model (where $x_t = 0, 1, \ldots, T$):

$$\textit{Level 1}: y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}, \quad (1)$$
$$\textit{Level 2a}: \eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}, \quad (2)$$
$$\textit{Level 2b}: \eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}. \quad (3)$$

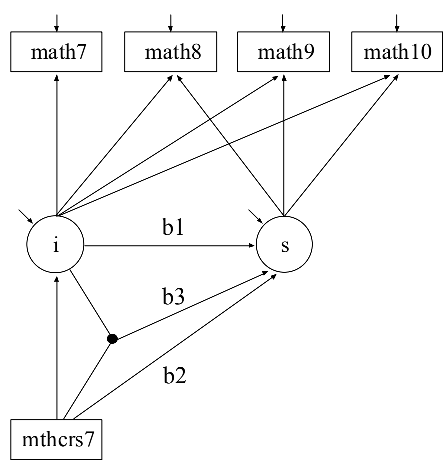
A useful type of model extension is to replace (3) by the regression equation

$$\eta_{1i} = \alpha + \beta\, \eta_{0i} + \gamma\, w_i + \zeta_i. \quad (4)$$

Example: Blood Pressure (Bloomqvist, 1977)



Growth Model With An Interaction



Input For A Growth Model With An Interaction Between A Latent And An Observed Variable

```
TITLE:     growth model with an interaction between a latent and an
           observed variable
DATA:      FILE IS lsay.dat;
VARIABLE:  NAMES ARE math7 math8 math9 math10 mthcrs7;
           MISSING ARE ALL (9999);
           CENTERING = GRANDMEAN (mthcrs7);
DEFINE:    math7 = math7/10;
           math8 = math8/10;
           math9 = math9/10;
           math10 = math10/10;
ANALYSIS:  TYPE=RANDOM MISSING;
MODEL:     i s | math7@0 math8@1 math9@2 math10@3;
           [math7-math10] (1);    !growth language defaults
           [i@0 s];               !overridden

           inter | i XWITH mthcrs7;
           s ON i mthcrs7 inter;
           i ON mthcrs7;
OUTPUT:    SAMPSTAT STANDARDIZED TECH1 TECH8;
```


Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable

Tests Of Model Fit

```
Loglikelihood

              H0 Value                          -10068.944
Information Criteria

              Number of Free Parameters                 12
              Akaike (AIC)                       20161.887
              Bayesian (BIC)                     20234.365
              Sample-Size Adjusted BIC           20196.236
                 (n* = (n + 2) / 24)
```

Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable (Continued)

Model Results

```
                    Estimates      S.E.  Est./S.E.
 I        |
    MATH7             1.000      0.000     0.000
    MATH8             1.000      0.000     0.000
    MATH9             1.000      0.000     0.000
    MATH10            1.000      0.000     0.000
 S        |
    MATH7             0.000      0.000     0.000
    MATH8             1.000      0.000     0.000
    MATH9             2.000      0.000     0.000
    MATH10            3.000      0.000     0.000
```


Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable (Continued)

```
                    Estimates      S.E.    Est./S.E.
 S        ON
    I                 0.087        0.012      7.023
    INTER            -0.047        0.006     -7.301
 S        ON
    MTHCRS7           0.045        0.013      3.555
 I        ON
    MTHCRS7           0.632        0.016     40.412
```

Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable (Continued)

	Estimates	S.E.	Est./S.E.
Intercepts			
MATH7	5.019	0.015	341.587
MATH8	5.019	0.015	341.587
MATH9	5.019	0.015	341.587
MATH10	5.019	0.015	341.587
I	0.000	0.000	0.000
S	0.417	0.007	57.749
Residual Variances			
MATH7	0.184	0.011	16.117
MATH8	0.178	0.009	20.109
MATH9	0.164	0.009	18.369
MATH10	0.173	0.015	11.509
I	0.528	0.018	28.935
S	0.037	0.004	10.027

Interpreting The Effect Of The Interaction Between Initial Status Of Growth In Math Achievement And Course Taking In Grade 6

- Model equation for slope s
  s = a + b1*i + b2*mthcrs7 + b3*i*mthcrs7 + e

  or, using a moderator function (Klein & Moosbrugger, 2000) where i moderates the influence of mthcrs7 on s
  s = a + b1*i + (b2 + b3*i)*mthcrs7 + e

- Estimated model

  Unstandardized
  s = 0.417 + 0.087*i + (0.045 – 0.047*i)*mthcrs7

  Standardized with respect to i and mthcrs7
  s = 0.42 + 0.08 * i + (0.04-0.04*i)*mthcrs7

Interpreting The Effect Of The Interaction Between Initial Status Of Growth In Math Achievement And Course Taking In Grade 6 (Continued)

- Interpretation of the standardized solution

  At the mean of i, which is zero, the slope increases 0.04 for 1 SD increase in mthcrs7

  At 1 SD below the mean of i, which is zero, the slope increases 0.08 for 1 SD increase in mthcrs7

  At 1 SD above the mean of i, which is zero, the slope does not increase as a function of mthcrs7

Growth Modeling With Parallel Processes

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- **Multiple processes**
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

Multiple Processes

- Parallel processes
- Sequential processes

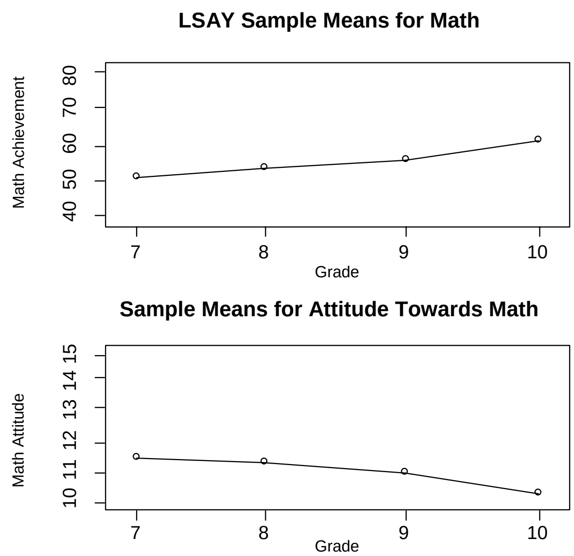
Growth Modeling With Parallel Processes

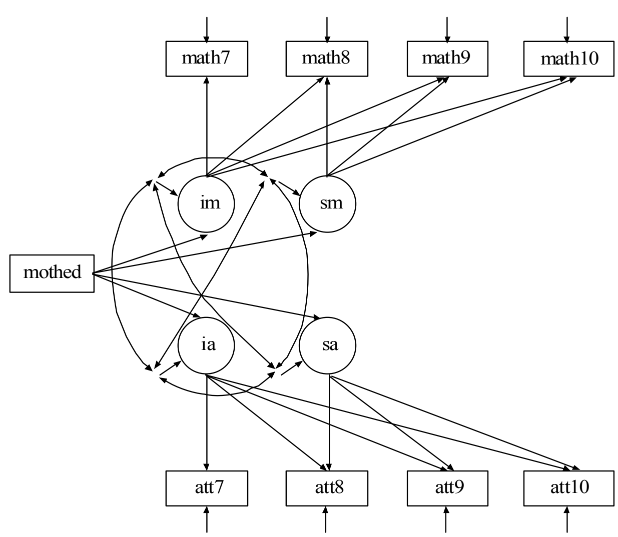
- Estimate a growth model for each process separately
  - Determine the shape of the growth curve
  - Fit model without covariates
  - Modify the model
- Joint analysis of both processes
- Add covariates

LSAY Data

The data come from the Longitudinal Study of American Youth (LSAY). Two cohorts were measured at four time points beginning in 1987. Cohort 1 was measured in Grades 10, 11, and 12. Cohort 2 was measured in Grades 7, 8, 9, and 10. Each cohort contains approximately 60 schools with approximately 60 students per school. The variables measured include math and science achievement items, math and science attitude measures, and background information from parents, teachers, and school principals. There are approximately 60 items per test with partial item overlap across grades—adaptive tests.

Data for the analysis include the younger females. The variables include math achievement and math attitudes from Grades 7, 8, 9, and 10 and mother's education.


LSAY Sample Means for Math
Math Achievement
40 50 60 70 80
7 8 9 10
Grade
Sample Means for Attitude Towards Math
Math Attitude
10 11 12 13 14 15
7 8 9 10
Grade


math7
math8
math9
math10
im
sm
mothed
ia
sa
att7
att8
att9
att10

Correlations Between Processes

- Through covariates
- Through growth factors (growth factor residuals)
- Through outcome residuals

Input For LSAY Parallel Process Growth Model

```
TITLE:     LSAY For Younger Females With Listwise Deletion
           Parallel Process Growth Model-Math Achievement and
           Math Attitudes

DATA:      FILE IS lsay.dat;
           FORMAT IS 3f8 f8.4 8f8.2 3f8 2f8.2;

VARIABLE:  NAMES ARE cohort id school weight math7 math8 math9
           math10 att7 att8 att9 att10 gender mothed homeres
           ses3 sesq3;
           USEOBS = (gender EQ 1 AND cohort EQ 2);
           MISSING = ALL (999);
           USEVAR = math7-math10 att7-att10 mothed;
```

Input For LSAY Parallel Process Growth Model (Continued)

```
MODEL:      im sm | math7@0 math8@1 math9 math10;
            ia sa | att7@0 att8@1 att9@2 att10@3;
            im-sa ON mothed;

OUTPUT:     MODINDICES STANDARDIZED;
```

Alternative language:

```
            im BY math7-math10@1;
            sm BY math7@0 math8@1 math9 math10;

            ia BY att7-att10@1;
            sa BY att7@0 att8@1 att9@2 att10@3;

            [math7-math10@0 att7-att10@0];
            [im sm ia sa];

            im-sa ON mothed;
```


Output Excerpts LSAY Parallel Process Growth Model

n = 910

Tests of Model Fit

```
Chi-Square Test of Model Fit

          Value                         43.161
          Degrees of Freedom                24
          P-Value                        .0095

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                        .030
          90 Percent C.I.                 .015   .044
          Probability RMSEA <= .05        .992
```

Output Excerpts LSAY Parallel Process Growth Model (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
IM ON					
MOTHED	**2.462**	**.280**	**8.798**	**.311**	**.303**
SM ON					
MOTHED	**.145**	**.066**	**2.195**	**.132**	**.129**
IA ON					
MOTHED	**.053**	**.086**	**.614**	**.025**	**.024**
SA ON					
MOTHED	**.012**	**.035**	**.346**	**.017**	**.017**

Output Excerpts LSAY Parallel Process Growth Model (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
SM WITH					
IM	3.032	.580	5.224	.350	.350
IA WITH					
IM	4.733	.702	6.738	.282	.282
SM	.544	.164	3.312	.235	.235
SA WITH					
IM	-.276	.279	-.987	-.049	-.049
SM	.130	.066	1.976	.168	.168
IA	-.567	.115	-4.913	-.378	-.378

Categorical Outcomes: Logistic And Probit Regression

Categorical Outcomes: Logit And Probit Regression

Probability varies as a function of x variables (here x_1, x_2)

$$P(u = 1 \mid x_1, x_2) = F[\beta_0 + \beta_1 x_1 + \beta_2 x_2], \qquad (22)$$

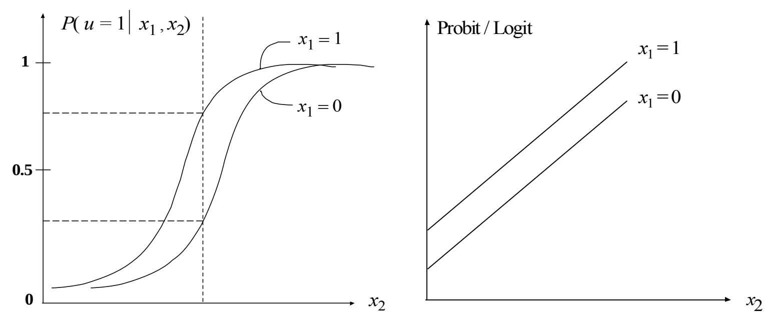
$P(u = 0 \mid x_1, x_2) = 1 - P[u = 1 \mid x_1, x_2]$, where $F[z]$ is either the standard normal ($\Phi[z]$) or logistic ($1/[1 + e^{-z}]$) distribution function.

Example: Lung cancer and smoking among coal miners

u	lung cancer ($u = 1$) or not ($u = 0$)
x_1	smoker ($x_1 = 1$), non-smoker ($x_1 = 0$)
x_2	years spent in coal mine

Categorical Outcomes: Logit And Probit Regression

$$P(u = 1 \mid x_1, x_2) = F\,[\beta_0 + \beta_1 x_1 + \beta_2 x_2\,], \qquad (22)$$



Interpreting Logit And Probit Coefficients

- Sign and significance
- Odds and odds ratios
- Probabilities

Logistic Regression And Log Odds

$Odds\ (u = 1 \mid x) = P(u = 1 \mid x) / P(u = 0 \mid x)$
$= P(u = 1 \mid x) / (1 - P(u = 1 \mid x)).$

The logistic function

$$P(u = 1 \mid x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

gives a log odds linear in x,

$logit = log\ [odds\ (u = 1 \mid x)] = log\ [P(u = 1 \mid x) / (1 - P(u = 1 \mid x))]$

$$= \log \left[\frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \ / \ (1 - \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}) \right]$$

$$= \log \left[\frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} * \frac{1 + e^{-(\beta_0 + \beta_1 x)}}{e^{-(\beta_0 + \beta_1 x)}} \right]$$

$$= \log \left[e^{(\beta_0 + \beta_1 x)} \right] = \beta_0 + \beta_1 x$$

Logistic Regression And Log Odds (Continued)

- $logit = log\ odds = \beta_0 + \beta_1 x$
- When x changes one unit, the *logit* (*log odds*) changes β_1 units
- When x changes one unit, the *odds* changes e^{β_1} units

Further Readings On Categorical Variable Analysis

Agresti, A. (2002). Categorical data analysis. Second edition. New York: John Wiley & Sons.

Agresti, A. (1996). An introduction to categorical data analysis. New York: Wiley.

Hosmer, D. W. & Lemeshow, S. (2000). Applied logistic regression. Second edition. New York: John Wiley & Sons.

Long, S. (1997). Regression models for categorical and limited dependent variables. Thousand Oaks: Sage.

Growth Models With Categorical Outcomes

Growth Model With Categorical Outcomes

- Individual differences in development of probabilities over time
- Logistic model considers growth in terms of log odds (logits), e.g.

$$(1) \qquad \log\left[\frac{P(u_{ti}=1 \mid \eta_{0i},\eta_{1i},\eta_{2i},x_{ti})}{P(u_{ti}=0 \mid \eta_{0i},\eta_{1i},\eta_{2i},x_{ti})}\right] = \eta_{0i} + \eta_{1i}\cdot(x_{ti}-c) + \eta_{2i}\cdot(x_{ti}-c)^2$$

for a binary outcome using a quadratic model with centering at time c. The growth factors η_{0i}, η_{1i}, and η_{2i} are assumed multivariate normal given covariates,

$$(2a) \quad \eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$$

$$(2b) \quad \eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$$

$$(2c) \quad \eta_{2i} = \alpha_2 + \gamma_2\, w_i + \zeta_{2i}$$

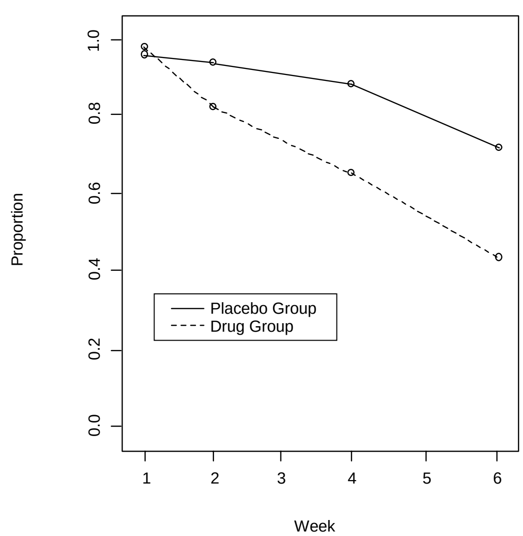
Growth Models With Categorical Outcomes

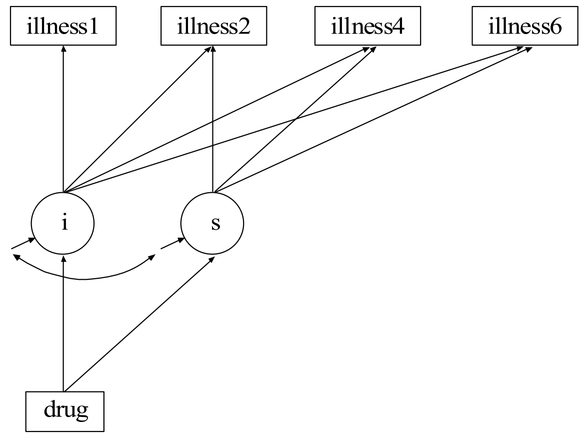
- Measurement invariance of the outcome over time is represented by the equality of thresholds over time (rather than intercepts)
- Thresholds not set to zero but held equal across timepoints—intercept factor mean value fixed at zero (parameterization 2)
- Differences in variances of the outcome over time are represented by allowing scale parameters to vary over time (WLS)

The NIMH Schizophrenia Collaborative Study

- The Data—The NIMH Schizophrenia Collaborative Study (Schizophrenia Data)
  - A group of 64 patients using a placebo and 249 patients on a drug for schizophrenia measured at baseline and at weeks one through six
  - Variables—severity of illness, background variables, and treatment variable
- Data for the analysis—severity of illness at weeks one, two, four, and six and treatment

Schizophrenia Data: Sample Proportions





Input For Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group

```
TITLE:      Growth model on schizophrenia data

DATA:       FILE = SCHIZ.DAT; FORMAT = 5F1;

VARIABLE:   NAMES = illness1 illness2 illness4 illness6 drug;
            CATEGORICAL = illness1-illness6;
            USEV = illness1-illness6;
            USEOBS = drug EQ 1;

ANALYSIS:   ESTIMATOR = ML;

MODEL:      i s | illness1@0 illness2@1 illness4@3 illness6@5;

OUTPUT:     TECH1 TECH10;
```

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group

```
TEST OF MODEL FIT

Loglikelihood
          H0 Value                             -405.068

Information Criteria
          Number of Free Parameters                   5
          Akaike (AIC)                          820.136
          Bayesian (BIC)                        837.724
          Sample-Size Adjusted BIC              821.873
            (n* = (n + 2) / 24)
```


Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

```
Chi-Square Test of Model Fit for the Binary and Ordered
Categorical (Ordinal) Outcomes

          Pearson Chi-Square

          Value                                  49.923
          Degrees of Freedom                         10
          P-Value                                0.0000

          Likelihood Ratio Chi-Square

          Value                                  42.960
          Degrees of Freedom                         10
          P-Value                                0.0000
```

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

```
TECHNICAL 10 OUTPUT

     MODEL FIT INFORMATION FOR THE LATENT CLASS INDICATOR MODEL PART

     RESPONSE PATTERNS

      No.   Pattern   No.   Pattern    No.   Pattern   No.   Pattern
      1     1001      2     1000       3     0000      4     1111
      5     1110      6     1010       7     1100      8     0011
      9     1011      10    0111       11    1101
```


Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

Response Pattern	Frequency Observed	Frequency Estimated	Standardized Residual (z-score)	Chi-Square Pearson	Contribution Loglikelihood
1	4.00	0.59	4.43	19.55	15.26
2	**24.00**	**12.89**	**3.18**	**9.57**	**29.83**
3	2.00	4.89	-1.32	1.71	-3.57
4	97.00	99.00	-0.26	0.04	-3.95
5	69.00	59.53	1.41	1.51	20.38
6	2.00	2.78	-0.47	0.22	-1.32
7	43.00	54.42	-1.75	2.40	-20.26
8	1.00	0.47	0.78	0.60	1.51
9	5.00	1.90	2.26	5.07	9.69
10	1.00	1.84	-0.62	0.38	-1.22
11	1.00	5.47	-1.93	3.65	-3.40

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

BIVARIATE MODEL FIT INFORMATION
Estimated Probabilities

Variable	Variable	H1	H0	Standardized Residual (z-score)
ILLNESS1	ILLNESS2			
Category 1	Category 1	0.012	0.025	-1.295
Category 1	Category 2	0.004	0.025	-2.127
Category 2	Category 1	0.141	0.073	**4.103**
Category 2	Category 2	0.843	0.877	-1.623
Bivariate Pearson Chi-Square				**21.979**
Bivariate Log-Likelihood Chi-Square				**21.430**

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

Variable	Variable	H1	H0	Standardized Residual (z-score)
ILLNESS1	ILLNESS4			
Category 1	Category 1	0.008	0.034	-2.271
Category 1	Category 2	0.008	0.016	-0.977
Category 2	Category 1	0.289	0.295	-0.191
Category 2	Category 2	0.695	0.655	1.307
Bivariate Pearson Chi-Square				6.533
Bivariate Log-Likelihood Chi-Square				8.977

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

Variable	Variable	H1	H0	Standardized Residual (z-score)
ILLNESS1	ILLNESS6			
Category 1	Category 1	0.008	0.038	-2.459
Category 1	Category 2	0.008	0.012	-0.599
Category 2	Category 1	0.554	0.521	1.063
Category 2	Category 2	0.430	0.430	0.006
Bivariate Pearson Chi-Square				6.713
Bivariate Log-Likelihood Chi-Square				9.529

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

Variable	Variable	H1	H0	Standardized Residual (z-score)
ILLNESS2	ILLNESS4			
Category 1	Category 1	0.120	0.075	2.722
Category 1	Category 2	0.032	0.023	0.996
Category 2	Category 1	0.177	0.254	-2.796
Category 2	Category 2	0.671	0.648	0.735
Bivariate Pearson Chi-Square				13.848
Bivariate Log-Likelihood Chi-Square				13.361

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

Variable	Variable	H1	H0	Standardized Residual (z-score)
ILLNESS2	ILLNESS6			
Category 1	Category 1	0.112	0.085	1.578
Category 1	Category 2	0.040	0.013	3.745
Category 2	Category 1	0.450	0.474	-0.754
Category 2	Category 2	0.398	0.429	-0.988
Bivariate Pearson Chi-Square				16.976
Bivariate Log-Likelihood Chi-Square				11.831

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes: Treatment Group (Continued)

Variable	Variable	H1	H0	Standardized Residual (z-score)
ILLNESS4	ILLNESS6			
Category 1	Category 1	0.277	0.302	-0.841
Category 1	Category 2	0.20	0.027	-0.697
Category 2	Category 1	0.285	0.257	1.027
Category 2	Category 2	0.418	0.414	0.103
Bivariate Pearson Chi-Square				1.757
Bivariate Log-Likelihood Chi-Square				1.791
Overall Bivariate Pearson Chi-Square				67.806
Overall Bivariate Log-Likelihood Chi-Square				66.920

Input For Schizophrenia Data Growth Model For Binary Outcomes With A Treatment Variable

```
TITLE:      Schizophrenia Data
            Growth Model for Binary Outcomes
            With a Treatment Variable and Scaling Factors
DATA:       FILE IS schiz.dat; FORMAT IS 5F1;
VARIABLE:   NAMES ARE illness1 illness2 illness4 illness6
            drug;   ! 0=placebo (n=64)  1=drug (n=249)
            CATEGORICAL ARE illness1-illness6;
ANALYSIS:   ESTIMATOR = ML;
            !ESTIMATOR = WLSMV;
MODEL:      i s | illness1@0 illness2@1 illness4@3
                  illness6@5;
            i s ON drug;
Alternative language:
MODEL:      i BY illness1-illness6@1;
            s BY illness1@0 illness2@1 illness4@3 illness6@5;
            [illness1$1 illness2$1 illness4$1 illness6$1] (1);
            [s];
            i s ON drug;
            !{illness1@1 illness2-illness6};
```


Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes With A Treatment Variable

n = 313

Tests Of Model Fit

```
Loglikelihood
          HO Value                     -486.337
Information Criteria
          Number of Free Parameters          7
          Akaike (AIC)                  986.674
          Bayesian (BIC)               1012.898
          Sample-Size Adjusted BIC      990.696
             (n* = (n + 2) / 24)
```

Output Excerpts Schizophrenia Data Growth Model For Binary Outcomes With A Treatment Variable (Continued)

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
I ON					
DRUG	-0.429	0.825	-0.521	-0.156	-0.063
S ON					
DRUG	-0.651	0.259	-2.512	-0.684	-0.276
I WITH					
S	-0.925	0.621	-1.489	-0.353	-0.353
Intercepts					
I	0.000	0.000	0.000	0.000	0.000
S	-0.555	0.255	-2.182	-0.583	-0.583
Thresholds					
ILLNESS1$1	-5.706	1.047	-5.451		
ILLNESS2$1	-5.706	1.047	-5.451		
ILLNESS4$1	-5.706	1.047	-5.451		
ILLNESS5$1	-5.706	1.047	-5.451		
Residual Variances					
I	7.543	3.213	2.348	0.996	0.996
S	0.838	0.343	2.440	0.924	0.924

Further Readings On Growth Analysis With Categorical Outcomes

Fitzmaurice, G.M., Laird, N.M. & Ware, J.H. (2004). Applied longitudinal analysis. New York: Wiley.

Gibbons, R.D. & Hedeker, D. (1997). Random effects probit and logistic regression models for three-level data. Biometrics, 53, 1527-1537.

Hedeker, D. & Gibbons, R.D. (1994). A random-effects ordinal regression model for multilevel analysis. Biometrics, 50, 933-944.

Muthén, B. (1996). Growth modeling with binary responses. In A. V. Eye, & C. Clogg (Eds.), Categorical variables in developmental research: methods of analysis (pp. 37-54). San Diego, CA: Academic Press. (#64)

Muthén, B. & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Mplus Web Note #4 (www.statmodel.com).

References

(To request a Muthén paper, please email bmuthen@ucla.edu.)

Analysis With Longitudinal Data

Introductory

Bollen, K.A. & Curran, P.J. (2006). Latent curve models. A structural equation perspective. New York: Wiley.

Collins, L.M. & Sayer, A. (Eds.) (2001). New methods for the analysis of change. Washington, D.C.: American Psychological Association.

Curran, P.J. & Bollen, K.A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In Collins, L.M. & Sayer, A.G. (Eds.) New methods for the analysis of change (pp. 105-136). Washington, DC: American Psychological Association.

Duncan, T., Duncan S. & Strycker, L. (2006). An introduction to latent variable growth curve modeling. Second edition. Lawrence Erlbaum: New York.

Goldstein, H. (2003). Multilevel statistical models. Third edition. London: Edward Arnold.

References (Continued)

Jennrich, R.I. & Schluchter, M.D. (1986). Unbalanced repeated-measures models with structured covariance matrices. Biometrics, 42, 805-820.

Laird, N.M., & Ware, J.H. (1982). Random-effects models for longitudinal data. Biometrics, 38, 963-974.

Lindstrom, M.J. & Bates, D.M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. Journal of the American Statistical Association, 83, 1014-1022.

Littell, R., Milliken, G.A., Stroup, W.W., & Wolfinger, R.D. (1996). SAS system for mixed models. Cary NC: SAS Institute.

McArdle, J.J. & Epstein, D. (1987). Latent growth curves within developmental structural equation models. Child Development, 58, 110-133.

McArdle, J.J. & Hamagami, F. (2001). Latent differences score structural models for linear dynamic analyses with incomplete longitudinal data. In Collins, L.M. & Sayer, A. G. (Eds.), New methods for the analysis of change (pp. 137-175). Washington, D.C.: American Psychological Association.

Meredith, W. & Tisak, J. (1990). Latent curve analysis Psychometrika, 55, 107-122.

References (Continued)

Muthén, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), Best methods for the analysis of change. Recent advances, unanswered questions, future directions (pp. 1-17). Washington D.C.: American Psychological Association.

Muthén, B. (2000). Methodological issues in random coefficient growth modeling using a latent variable framework: Applications to the development of heavy drinking. In Multivariate applications in substance use research, J. Rose, L. Chassin, C. Presson & J. Sherman (eds.), Hillsdale, N.J.: Erlbaum, pp. 113-140.

Muthén, B. & Khoo, S.T. (1998). Longitudinal studies of achievement growth using latent variable modeling. Learning and Individual Differences. Special issue: latent growth curve analysis, 10, 73-101.

Muthén, B. & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol, 61, 290-300.

Muthén, B. & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics, 55, 463-469.

References (Continued)

Rao, C.R. (1958). Some statistical models for comparison of growth curves. Biometrics, 14, 1-17.

Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods. Second edition. Newbury Park, CA: Sage Publications.

Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics, 23, 323-355.

Singer, J.D. & Willett, J.B. (2003). Applied longitudinal data analysis. Modeling change and event occurrence. New York, NY: Oxford University Press.

Snijders, T. & Bosker, R. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.

Tucker, L.R. (1958). Determination of parameters of a functional relation by factor analysis. Psychometrika, 23, 19-23.

References (Continued)

Advanced

Albert, P.S. & Shih, J.H. (2003). Modeling tumor growth with random onset. Biometrics, 59, 897-906.

Brown, C.H. & Liao, J. (1999). Principles for designing randomized preventive trials in mental health: An emerging development epidemiologic perspective. American Journal of Community Psychology, special issue on prevention science, 27, 673-709.

Brown, C.H., Indurkhya, A. & Kellam, S.K. (2000). Power calculations for data missing by design: applications to a follow-up study of lead exposure and attention. Journal of the American Statistical Association, 95, 383-395.

Collins, L.M. & Sayer, A. (Eds.), New methods for the analysis of change. Washington, D.C.: American Psychological Association.

Curran, P.J., Muthén, B., & Harford, T.C. (1998). The influence of changes in marital status on developmental trajectories of alcohol use in young adults. Journal of Studies on Alcohol, 59, 647-658.

Duan, N., Manning, W.G., Morris, C.N. & Newhouse, J.P. (1983). A comparison of alternative models for the demand for medical care. Journal of Business and Economic Statistics, 1, 115-126.

References (Continued)

Ferrer, E. & McArdle, J.J. (2004). Alternative structural models for multivariate longitudinal data analysis. Structural Equation Modeling, 10, 493-524.

Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (2008). Longitudinal Data Analysis. Chapman & Hall/CRC Press.

Klein, A. & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. Psychometrika, 65, 457-474.

Khoo, S.T. & Muthén, B. (2000). Longitudinal data on families: Growth modeling alternatives. In Multivariate applications in substance use research, J. Rose, L. Chassin, C. Presson & J. Sherman (eds), Hillsdale, N.J.: Erlbaum, pp. 43-78.

Miyazaki, Y. & Raudenbush, S.W. (2000). A test for linkage of multiple cohorts from an accelerated longitudinal design. Psychological Methods, 5, 44-63.

Moerbeek, M., Breukelen, G.J.P. & Berger, M.P.F. (2000). Design issues for experiments in multilevel populations. Journal of Educational and Behavioral Statistics, 25, 271-284.

References (Continued)

Muthén, B. (1996). Growth modeling with binary responses. In A.V. Eye & C. Clogg (eds), Categorical variables in developmental research: methods of analysis (pp. 37-54). San Diego, CA: Academic Press.

Muthén, B. (1997). Latent variable modeling with longitudinal and multilevel data. In A. Raftery (ed), Sociological Methodology (pp. 453-380). Boston: Blackwell Publishers.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), Handbook of quantitative methodology for the social sciences (pp. 345-368). Newbury Park, CA: Sage Publications.

Muthén, B. & Curran, P. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. Psychological Methods, 2, 371-402.

Muthén, B. & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol, 61, 290-300.

References (Continued)

Muthén, L.K. and Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation Modeling, 4, 599-620.

Olsen, M.K. & Schafer, J.L. (2001). A two-part random effects model for semicontinuous longitudinal data. Journal of the American Statistical Association, 96, 730-745.

Raudenbush, S.W. (1997). Statistical analysis and optimal design for cluster randomized trials. Psychological Methods, 2, 173-185.

Raudenbush, S.W. & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. Psychological Methods, 5, 199-213.

Roeder, K., Lynch, K.G., & Nagin, D.S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. Journal of the American Statistical Association, 94, 766-776.

Satorra, A. & Saris, W. (1985). Power of the likelihood ratio test in covariance structure analysis. Psychometrika, 51, 83-90.