Growth Modeling With Latent Variables Using Mplus

Linda K. Muthén
Bengt Muthén

www.statmodel.com

Table Of Contents

Statistical Analysis With Latent Variables A General Modeling Framework

Statistical Concepts Captured By Latent Variables

- Continuous Latent Variables
  - Measurement errors
  - Factors
  - Random effects
  - 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

Types of Variables

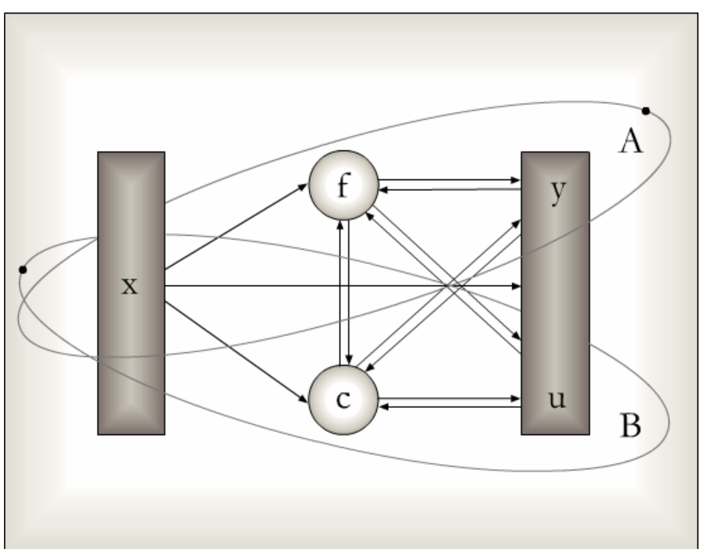
- Observed variables

  x background variables (no model structure)
  y continuous and censored outcome variables
  u categorical (dichotomous, ordinal, nominal) and count outcome variables

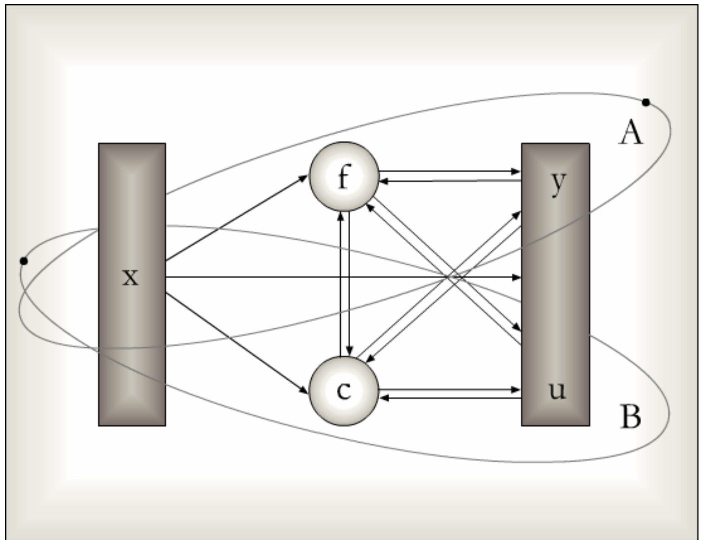
- Latent variables

  f continuous variables
  – interactions among f's
  c categorical variables
  – multiple c's

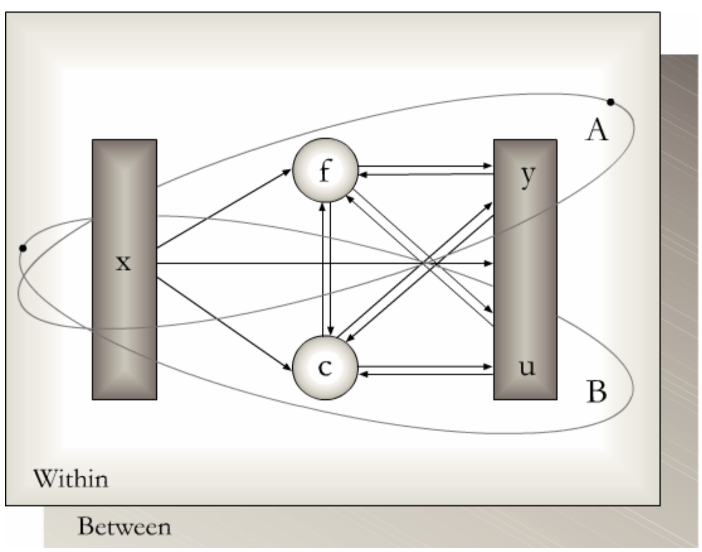
General Latent Variable Modeling Framework



General Latent Variable Modeling Framework



General Latent Variable Modeling Framework



Typical Examples Of Growth Modeling

LSAY Data

Longitudinal Study of American Youth (LSAY)

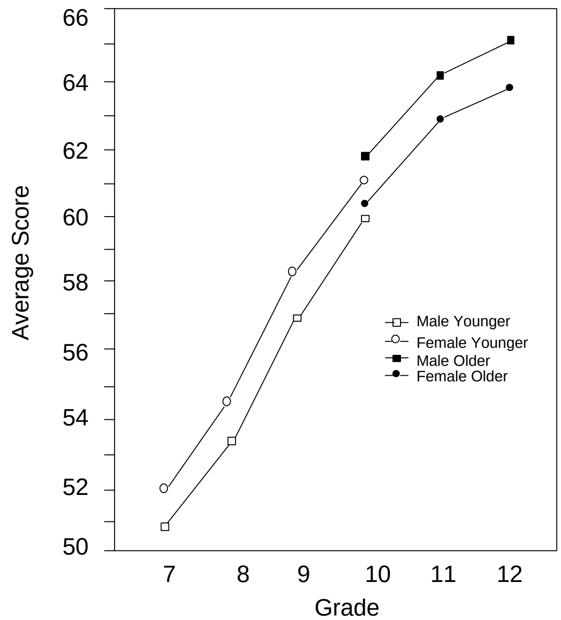
- Two cohorts measured each year beginning in 1987
  - Cohort 1 - Grades 10, 11, and 12
  - Cohort 2 - Grades 7, 8, 9, and 10

- Each cohort contains approximately 60 schools with approximately 60 students per school

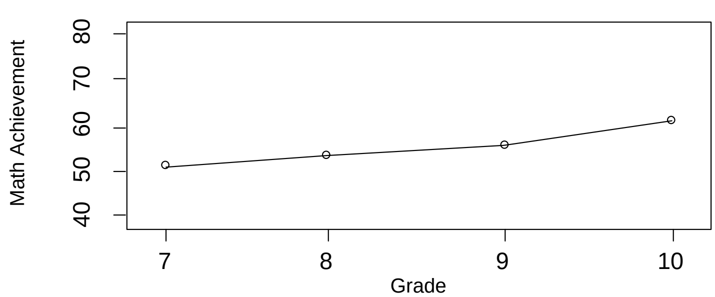
- Variables - math and science achievement items, math and science attitude measures, and background variable from parents, teachers, and school principals

- Approximately 60 items per test with partial item overlap across grades - adaptive tests

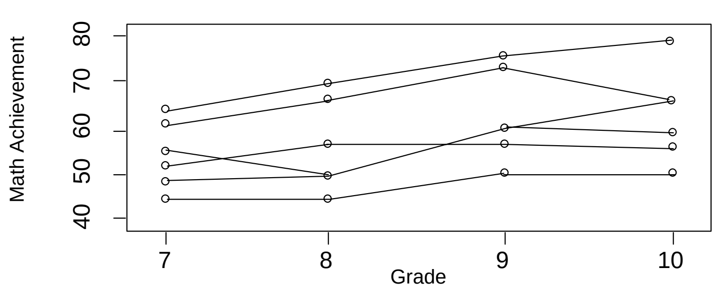
Math Total Score
LSAY Data

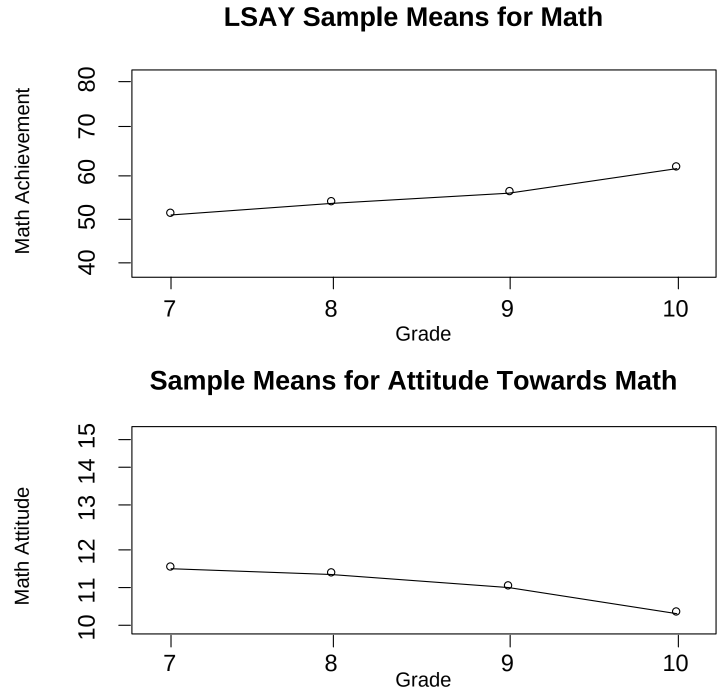


LSAY Mean Curve



Individual Curves



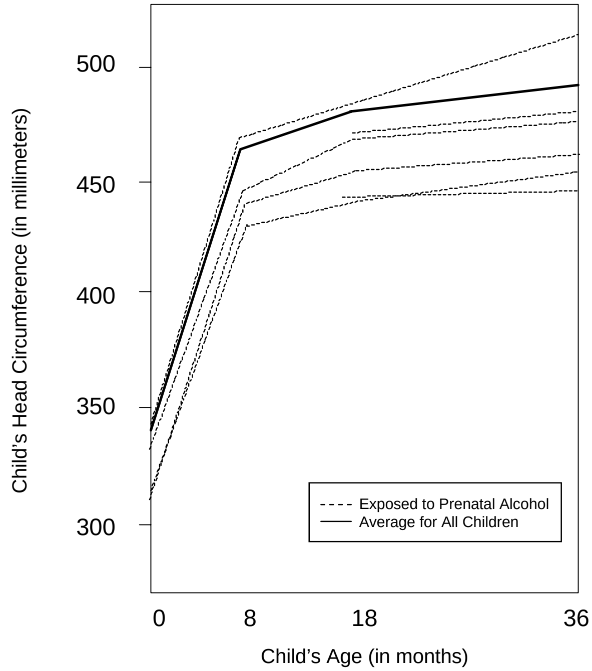


Maternal Health Project Data

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



Alternative Models For Longitudinal Data

- Growth curve model
- Auto-regressive model
- Hybrids

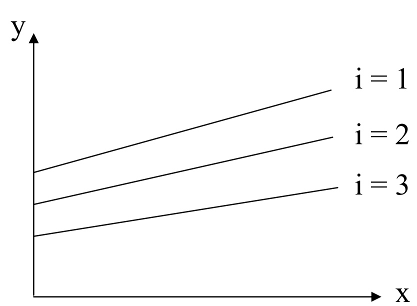
Basic Modeling Ideas

Individual Development Over Time

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

i = individual
t = timepoint
η_0 = growth intercept

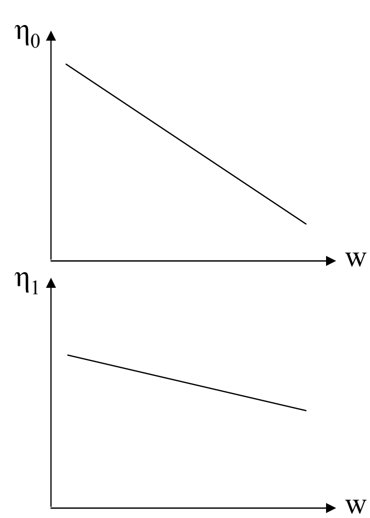
y = outcome
x = time score
η_1 = growth slope



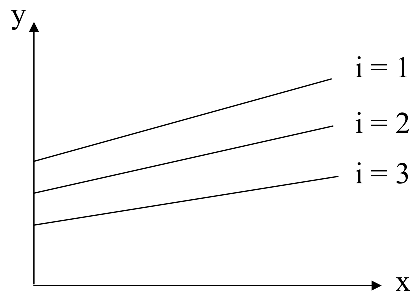
(2a) $\eta_{0i} = \alpha_0 + \gamma_0 \, w_i + \zeta_{0i}$

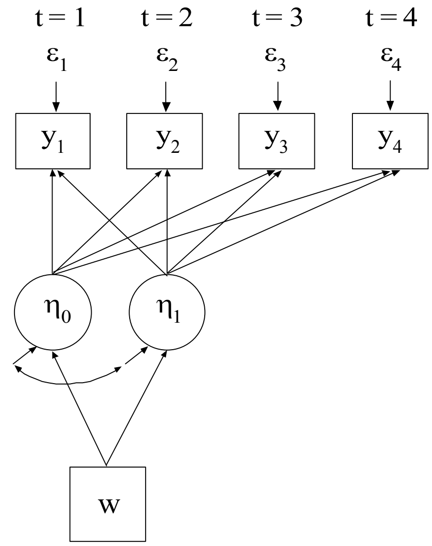
(2b) $\eta_{1i} = \alpha_1 + \gamma_1 \, w_i + \zeta_{1i}$

w = time-invariant covariate



Individual Development Over Time





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

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

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

Random Effects: Multilevel And Mixed Linear Modeling

Individual *i (i* = 1, 2, …, *n*) observed at time point *t* (*t* = 1, 2, … *T*).

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

- Level 1: $y_{it} = \eta_{0i} + \eta_{1i}\, x_{it} + \kappa_i\, w_{it} + \varepsilon_{it}$ (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)

  $\kappa_i = \alpha + \gamma\, w_i + \zeta_{1i}$ (42)

Mixed linear model:

$$y_{it} = \textit{fixed part} + \textit{random part} \quad (43)$$

$$= \alpha_0 + \gamma_0\, w_i + (\alpha_1 + \gamma_1\, w_i)\, x_{it} + (\alpha + \gamma\, w_i)\, w_{it} \quad (44)$$

$$+ \zeta_{0i} + \zeta_{1i} x_{it} + \zeta_i\, w_{it} + \varepsilon_{it}\,. \quad (45)$$

Random Effects: Multilevel And Mixed Linear Modeling (Continued)

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, 1995, MLn; 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:

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

Compare with level 1 of multilevel:

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

Multilevel approach:

- x_{it} 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):

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

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

κ_t not involved (parameter).

Random Effects: Mixed Linear Modeling and SEM

Mixed linear model in matrix form:

$$\mathrm{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iT})' \qquad (51)$$

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

Here, **X, Z** are design matrices with known values, $\boldsymbol{\alpha}$ contains fixed effects, and **b** contains random effects. Compare with (39) - (43).

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

Need for Λ_i, K_i, B_i, Γ_i.

Comparison Summary of Multilevel, Mixed Linear, and SEM Growth Models

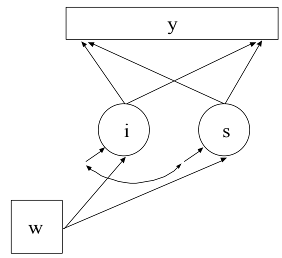
- Multilevel and mixed linear models are the same
- SEM differs from the multilevel and mixed linear models in two ways
  - Treatment of time scores
    - Time scores are data for multilevel and mixed linear models -- individuals can have different times of measurement
    - Time scores are parameters 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

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

- Wide: Multivariate, Single-Level Approach

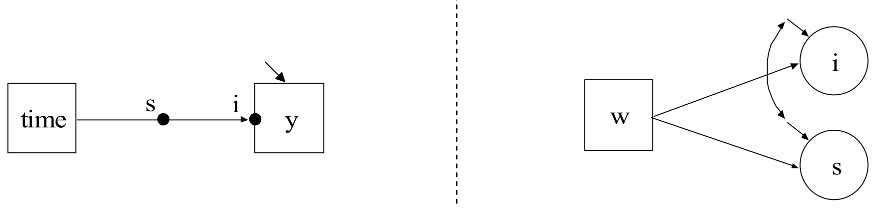
$y_{ti} = i_i + s_i \times time_{ti} + \varepsilon_{ti}$

i_i regressed on w_i
s_i regressed on w_i



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

Within | Between



Multilevel Modeling in a Latent Variable Framework

Integrating multilevel and SEM analyses (Asparouhov & Muthén, 2002).

Flexible combination of random effects and other latent variables:

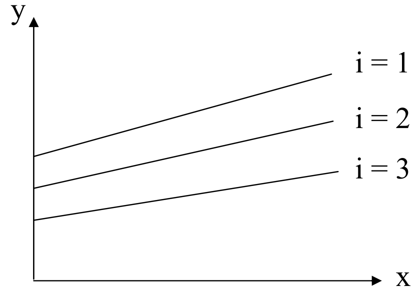
- Multilevel models with random effects (intercepts, slopes)
  - Individually-varying times of observation read as data
  - Random slopes for time-varying covariates
- SEM with factors on individual and cluster levels
- Models combining random effects and factors, e.g.
  - Cluster-level latent variable predictors with multiple indicators
  - Individual-level latent variable predictors with multiple indicators
- Special applications
  - Random coefficient regression (no clustering; heteroscedasticity)
  - Interactions between latent and observed continuous variables

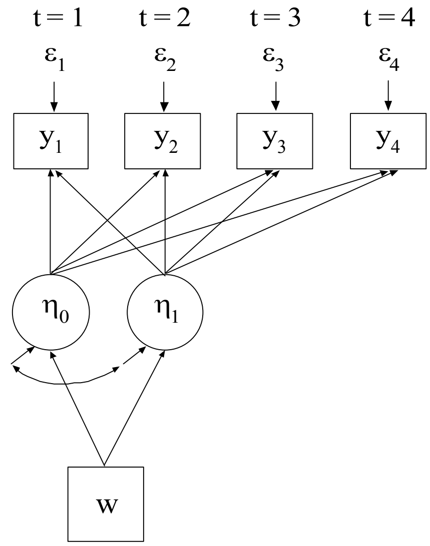
Advantages of Growth Modeling in a Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Random effects (intercepts, slopes) integrated with other latent variables
- Regressions among random effects
- Multiple processes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

The Latent Variable Growth Model in Practice

Individual Development Over Time





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

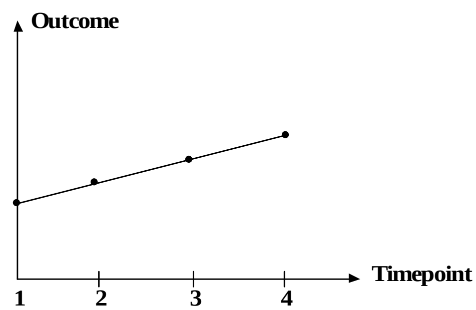
(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

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

Specifying Time Scores For Linear Growth Models

Linear Growth Model

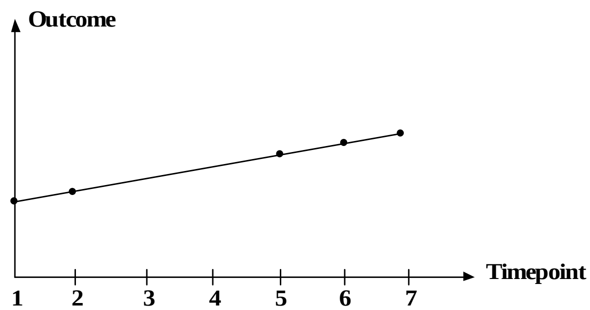
- Need two latent variables to describe a linear growth model: Intercept and Slope



- Equidistant time scores for slope:

Equidistant time scores	0	1	2	3
for slope:	0	.1	.2	.3

Specifying Time Scores For Linear Growth Models (Continued)



• Nonequidistant time scores for slope:	0	1	4	5	6
	0	.1	.4	.5	.6

Interpretation of the Linear Growth Factors

Model:

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

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

$$y_{i1} = \eta_{0i} + \eta_{1i}\, 0 + \varepsilon_{i1}, \qquad (18)$$

$$\eta_{0i} = y_{i1} - \varepsilon_{i1}, \qquad (19)$$

$$y_{i2} = \eta_{0i} + \eta_{1i}\, 1 + \varepsilon_{i2}, \qquad (20)$$

$$y_{i3} = \eta_{0i} + \eta_{1i}\, 2 + \varepsilon_{i3}, \qquad (21)$$

$$y_{i4} = \eta_{0i} + \eta_{1i}\, 3 + \varepsilon_{i4}, \qquad (22)$$

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.

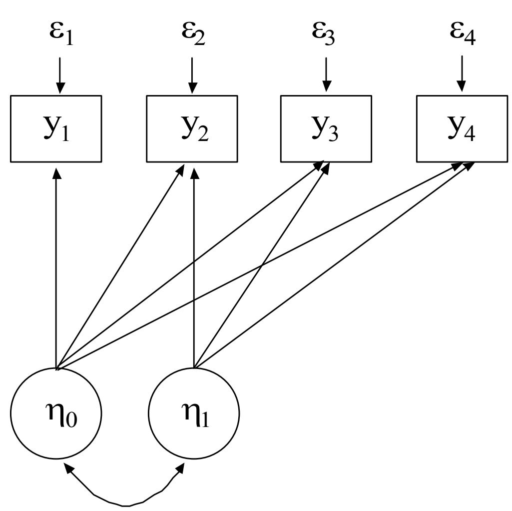
Interpreting Growth Model Parameters

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

Fixed parameters in the H_0 growth model:

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

Latent Growth Model Sources of Misfit

Sources of misfit:

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

Model modifications:

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

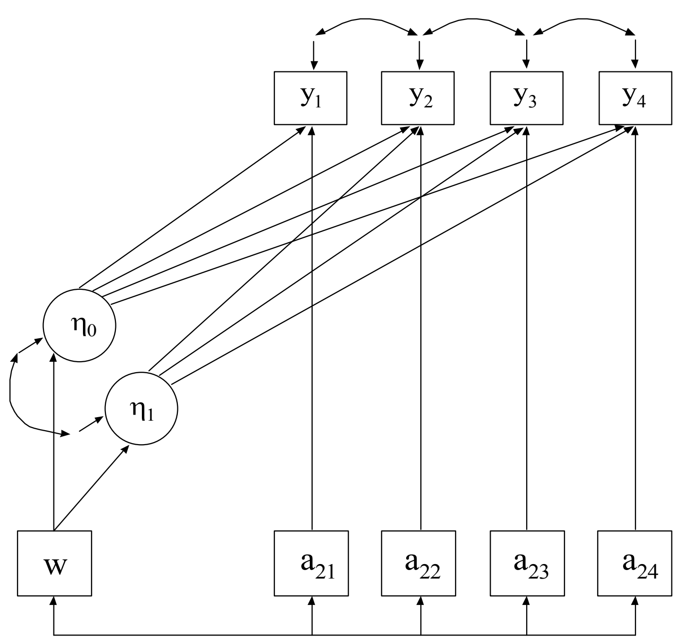
(8 parameters, 1 d.f.)

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

Fixed parameters in the H_0 growth model:

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

Time-Varying Covariates



Alternative Growth Model Parameterizations

Parameterization 1 – for continuous outcomes

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

$$\eta_{0i} = \boldsymbol{\alpha}_0 + \gamma_0\, w_i + \zeta_{0i}, \qquad (33)$$

$$\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}. \qquad (34)$$

Parameterization 2 – for categorical outcomes and multiple indicators

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

$$\eta_{0i} = \mathbf{0} + \gamma_0\, w_i + \zeta_{0i}, \qquad (36)$$

$$\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}, \qquad (37)$$

Simple Examples of Growth Modeling

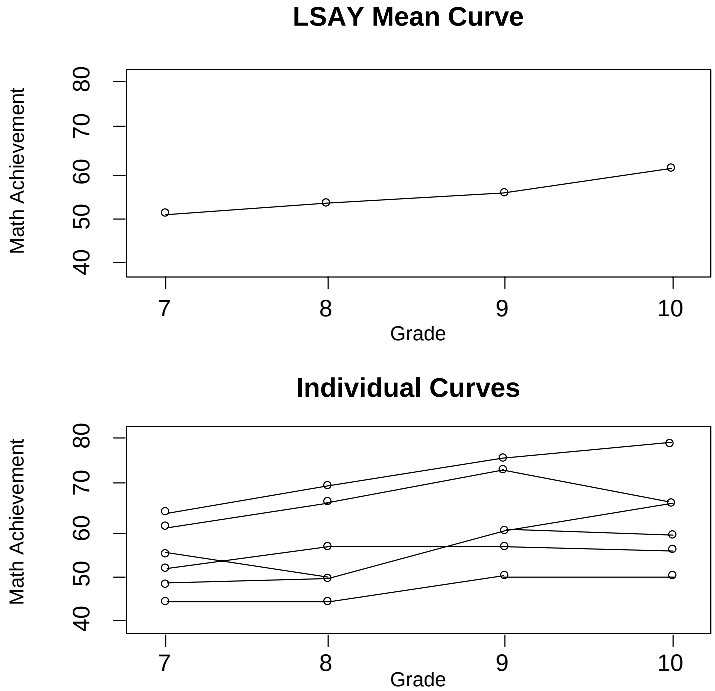
Steps in Growth Modeling

- 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 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 students per school. The variables measured included math and science achievement items, math and science attitude measures, and background information from parents, teachers, and school principals. There were 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 from Grades 7, 8, 9, and 10 and the background variables of mother’s education and home resources.



Input for LSAY TYPE=BASIC Analysis

```
TITLE:     LSAY For Younger Females With Listwise Deletion
           TYPE=BASIC Analysis
DATA:      FILE IS lsay.dat;
           FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:  NAMES ARE cohort id school weight math7 math8 math9
           math10 att7 att8 att9 att10 gender mothed homeres;
           USEOBS = (gender EQ 1 AND cohort EQ 2);
           MISSING = ALL (999);
           USEVAR = math7-math10;
ANALYSIS:  TYPE = BASIC;
PLOT:      TYPE = PLOT1;       !New in Mplus Version 3
```

Sample Statistics for LSAY Data

n = 984

Sample Statistics

Means

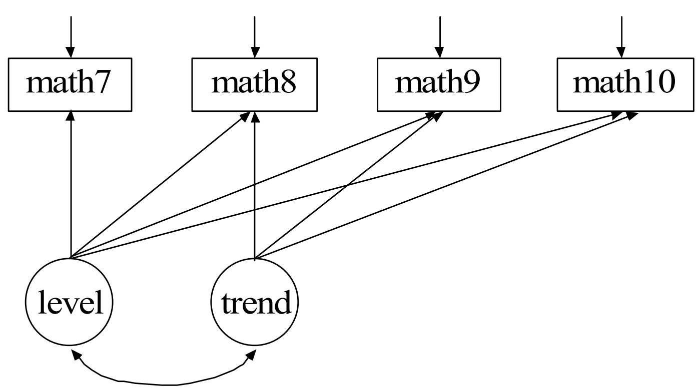
MATH7	MATH8	MATH9	MATH10
52.750	55.411	59.128	61.796

Covariances

	MATH7	MATH8	MATH9	MATH10
MATH7	81.107			
MATH8	67.663	82.829		
MATH9	73.150	76.513	100.986	
MATH10	77.952	82.668	95.158	131.326

Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	1.000			
MATH8	0.826	1.000		
MATH9	0.808	0.837	1.000	
MATH10	0.755	0.793	0.826	1.000



Input For LSAY Linear Growth Model Without Covariates

```
TITLE:      LSAY For Younger Females With Listwise Deletion
            Linear Growth Model Without Covariates
DATA:       FILE IS lsay.dat;
            FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:   NAMES ARE cohort id school weight math7 math8 math9
            math10 att7 att8 att9 att10 gender mothed homeres;
            USEOBS = (gender EQ 1 AND cohort EQ 2);
            MISSING = ALL (999);
            USEVAR = math7-math10;
ANALYSIS:   TYPE = MEANSTRUCTURE;
MODEL:      level BY math7-math10@1;
            trend BY math7@0 math8@1 math9@2 math10@3;
            [math7-math10@0];
            [level trend];
OUTPUT:     SAMPSTAT STANDARDIZED MODINDICES (3.84);

!New Version 3 Language For Growth Models
!MODEL: level trend | math7@0 math8@1 math9@2 math10@3;
```


Output Excerpts LSAY Linear Growth Model Without Covariates

Tests Of Model Fit

```
Chi-Square Test of Model Fit
          Value                         22.664
          Degrees of Freedom            5
          P-Value                       0.0004

CFI/TLI
          CFI                           0.995
          TLI                           0.994

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                      0.060
          90 Percent C.I.               0.036   0.086
          Probability RMSEA <= .05      0.223

SRMR (Standardized Root Mean Square Residual)
          Value                         0.025
```

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Modification Indices

	M.I.	E.P.C.	Std.E.P.C.	StdYX E.P.C.
TREND BY MATH7	6.793	0.185	0.254	0.029
TREND BY MATH8	14.694	-0.169	-0.233	-0.025
TREND BY MATH9	9.766	0.155	0.213	0.021

Output Excerpts LSAY Linear Growth Without Covariates

Model Results

	Estimates	S.E.	Est./S.E.	Std	StdYX
LEVEL BY					
MATH7	1.000	.000	.000	8.029	.906
MATH8	1.000	.000	.000	8.029	.861
MATH9	1.000	.000	.000	8.029	.800
MATH10	1.000	.000	.000	8.029	.708
TREND BY					
MATH7	.000	.000	.000	.000	.000
MATH8	1.000	.000	.000	1.377	.148
MATH9	2.000	.000	.000	2.753	.274
MATH10	3.000	.000	.000	4.130	.364

Output Excerpts LSAY Linear Growth Without Covariates (Continued)

LEVEL WITH					
TREND	3.491	.730	4.780	.316	.316
Residual Variances					
MATH7	14.105	1.253	11.259	14.105	.180
MATH8	13.525	.866	15.610	13.525	.156
MATH9	14.726	.989	14.897	14.726	.146
MATH10	25.989	1.870	13.898	25.989	.202
Variances					
LEVEL	64.469	3.428	18.809	1.000	1.000
TREND	1.895	.322	5.894	1.000	1.000

Output Excerpts LSAY Linear Growth Without Covariates (Continued)

Means					
LEVEL	52.623	.275	191.076	6.554	6.554
TREND	3.105	.075	41.210	2.255	2.255
Intercepts					
MATH7	.000	.000	.000	.000	.000
MATH8	.000	.000	.000	.000	.000
MATH9	.000	.000	.000	.000	.000
MATH10	.000	.000	.000	.000	.000

R-Square

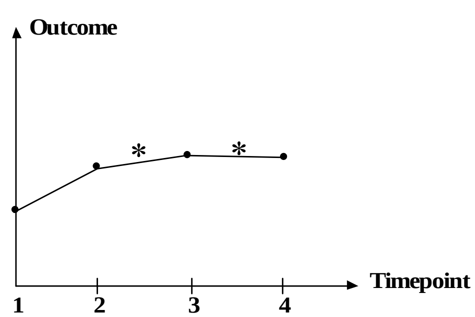
Observed Variable	R-Square
MATH7	0.820
MATH8	0.844
MATH9	0.854
MATH10	0.798

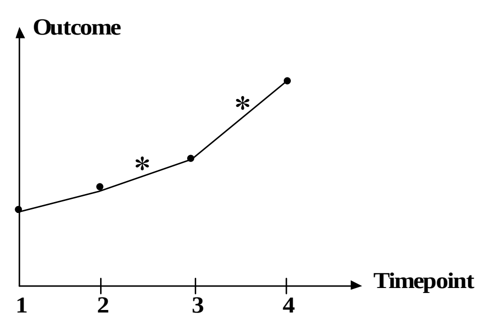
Growth Model With Free Time Scores

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

Interpretation of Slope Growth Factor Mean For Non-Linear Models

- The slope growth factor mean is the 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 change between grades 7 and 8
        - Time scores of 0 * * 1 – slope factor mean refers to change between grades 7 and 10

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
- Choice of time score starting values if needed
    - Means 52.75 55.41 59.13 61.80
    - Differences 2.66 3.72 2.67
    - Time scores 0 1 >2 >2+1

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

```
TITLE:      LSAY For Younger Females With Listwise Deletion
            Growth Model With Free Time Scores Without Covariates
DATA:       FILE IS lsay.dat;
            FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:   NAMES ARE cohort id school weight math7 math8 math9
            math10 att7 att8 att9 att10 gender mothed homeres;
            USEOBS = (gender EQ 1 AND cohort EQ 2);
            MISSING = ALL (999);
            USEVAR = math7-math10;
ANALYSIS:   TYPE = MEANSTRUCTURE;
MODEL:      level BY math7-math10@1;
            trend BY math7@0 math8@1 math9 math10;
            [math7-math10@0];
            [level trend];
OUTPUT:     RESIDUAL;

!New Version 3 Language For Growth Models
!MODEL: level trend | math7@0 math8@1 math9 math10;
```


Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates

n = 984

Tests Of Model Fit

```
Chi-Square Test of Model Fit
             Value                              4.222
             Degrees of Freedom                     3
             P-Value                           0.2373
CFI/TLI
             CFI                                1.000
             TLI                                0.999
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                           0.020
             90 Percent C.I.                    0.000   0.064
             Probability RMSEA <= .05           0.864
SRMR (Standardized Root Mean Square Residual)
             Value                              0.015
```

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

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
LEVEL BY					
MATH7	1.000	.000	.000	8.029	.903
MATH8	1.000	.000	.000	8.029	.870
MATH9	1.000	.000	.000	8.029	.797
MATH10	1.000	.000	.000	8.029	.708
TREND BY					
MATH7	.000	.000	.000	.000	.000
MATH8	1.000	.000	.000	1.377	.148
MATH9	**2.452**	**.133**	**18.442**	**2.780**	**.276**
MATH10	**3.497**	**.199**	**17.540**	**3.966**	**.350**

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

	(Estimates	S.E.	Est./S.E.	Std	StdYX)
TREND WITH					
LEVEL	3.110	.600	5.186	.342	.342
Variances					
LEVEL	64.470	3.394	18.994	1.000	1.000
TREND	**1.286**	**.265**	**4.853**	**1.000**	**1.000**
Means					
LEVEL	52.785	.283	186.605	6.574	6.574
TREND	**2.586**	**.167**	**15.486**	**2.280**	**2.280**

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

Residuals

Model Estimated Means / Intercepts / Thresholds

MATH7	MATH8	MATH9	MATH10
52.785	55.370	59.123	61.827

Residuals for Means / Intercepts / Thresholds

MATH7	MATH8	MATH9	MATH10
-.035	.041	.004	-.031

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

Model Estimated Covariances / Correlations / Residual Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	79.025			
MATH8	67.580	85.180		
MATH9	72.094	78.356	101.588	
MATH10	75.346	82.952	93.994	128.477

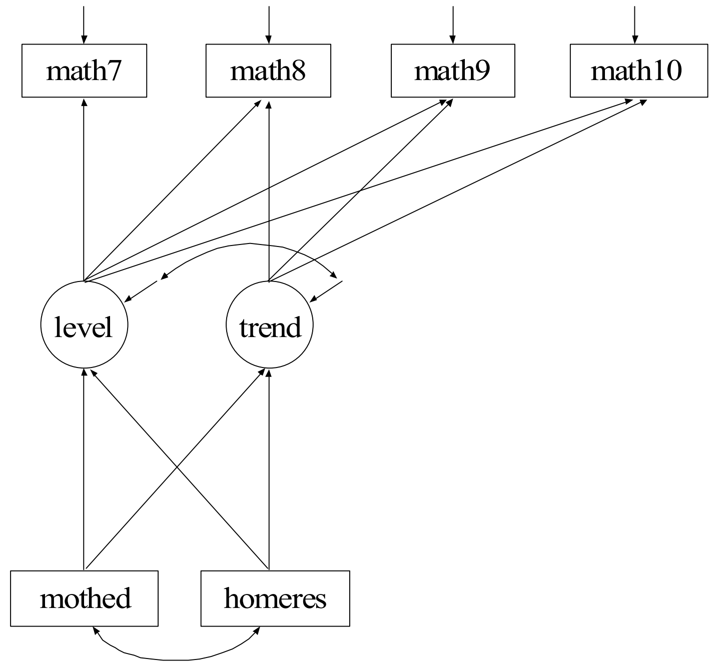
Residuals for Covariances / Correlations / Residual Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	1.999			
MATH8	.014	-2.436		
MATH9	.981	-1.921	-.705	
MATH10	2.527	-.368	1.067	2.715

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



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

```
TITLE:     LSAY For Younger Females With Listwise Deletion
           Growth Model With Free Time Scores and Covariates

DATA:      FILE IS lsay.dat;
           FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;

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

ANALYSIS:  TYPE = MEANSTRUCTURE; !ESTIMATOR = MLM;

MODEL:     level by math7-math10@1;
           trend BY math7@0 math8@1 math9 math10;
           [math7-math10@0];
           [level trend];
           level trend ON mothed homeres;

!New Version 3 Language For Growth Models
!MODEL: level trend | math7@0 math8@1 math9 math10;
!       level trend ON mothed homeres;
```

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates

n = 935

Tests Of Model Fit for ML

```
Chi-Square Test of Model Fit
             Value                            15.845
             Degrees of Freedom                    7
             P-Value                          0.0265
CFI/TLI
             CFI                               0.998
             TLI                               0.995
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                          0.037
             90 Percent C.I.                   0.012  0.061
             Probability RMSEA <= .05          0.794
SRMR (Standardized Root Mean Square Residual)
             Value                             0.015
```


Output Excerpts LSAY Growth Model With Free Time Scores And Covariates (Continued)

Tests Of Model Fit for MLM

```
Chi-Square Test of Model Fit
             Value                            8.554*
             Degrees of Freedom                    7
             P-Value                          0.2862
             Scaling Correction Factor         1.852
                 for MLM
CFI/TLI
             CFI                               0.999
             TLI                               0.999
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                          0.015
SRMR (Standardized Root Mean Square Residual)
             Value                             0.015
WRMR (Weighted Root Mean Square Residual)
             Value                             0.567
```

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates (Continued)

Selected Estimates For ML

	Estimates	S.E.	Est./S.E.	Std	StdYX
LEVEL ON					
MOTHED	**2.054**	**.281**	**7.322**	**.257**	**.247**
HOMERES	**1.376**	**.182**	**7.546**	**.172**	**.255**
TREND ON					
MOTHED	.103	.068	1.524	.094	.090
HOMERES	**.149**	**.045**	**3.334**	**.136**	**.201**
LEVEL WITH					
TREND	2.604	.559	4.658	.297	.297
Residual Variances					
LEVEL	53.931	2.995	18.008	.842	.842
TREND	1.134	.253	4.488	.942	.942
Intercepts					
LEVEL	43.877	.790	55.531	5.484	5.484
TREND	1.859	.221	8.398	1.695	1.695

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates (Continued)

R-Square

Observed Variable	R-Square
MATH7	0.813
MATH8	0.849
MATH9	0.861
MATH10	0.796

Latent Variable	R-Square
LEVEL	.158
TREND	.058

Model Estimated Average And Individual Growth Curves With Covariates

Model:

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

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

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

Estimated growth factor means:

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

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

Estimated outcome means:

$$\hat{E}(y_{it}) = \hat{E}(\eta_{0i}) + \hat{E}(\eta_{1i})\, x_t\, . \qquad (28)$$

Estimated outcomes for individual *i* :

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

where $\hat{\eta}_{0i}$ and $\hat{\eta}_{1i}$ are estimated factor scores. $\hat{y}_{it}$ 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 (Mothed)*
Sample Mean (Mothed) +
Estimated Slope (Homeres)*
Sample Mean (Homeres)

43.88 + 2.05*2.31 + 1.38*3.11 = 52.9

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

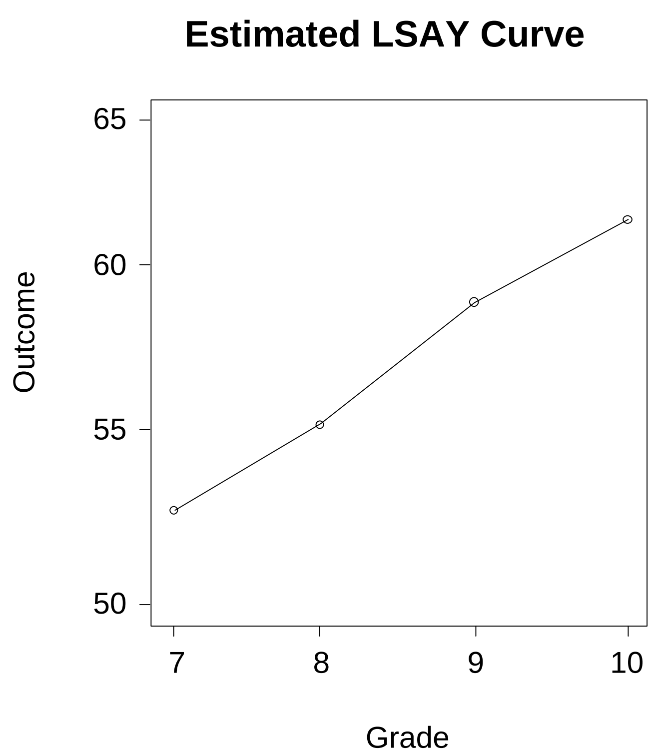
1.86 + .10*2.31 + .15*3.11 = 2.56

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 =
52.9 + 2.56 * (0) = **52.9**
Estimated Outcome Mean at Timepoint 2 =
52.9 + 2.56 * (1.00) = **55.46**
Estimated Outcome Mean at Timepoint 3 =
52.9 + 2.56 * (2.45) = **59.17**
Estimated Outcome Mean at Timepoint 4 =
52.9 + 2.56 * (3.50) = **61.86**



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

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 For LSAY Growth Model With Free Time Scores and Covariates Centered At Grade 10

```
TITLE:       LSAY For Younger Females With Listwise Deletion
             Growth Model With Free Time Scores and Covariates
             Centered at Grade 10
DATA:        FILE IS lsay.dat;
             FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:    NAMES ARE cohort id school weight math7 math8
             math9 math10 att7 att8 att9 att10 gender mothed homeres;
             USEOBS = (gender EQ 1 AND cohort EQ 2);
             MISSING = ALL (999);
             USEVAR = math7-math10 mothed homeres;
ANALYSIS:    TYPE = MEANSTRUCTURE;
MODEL:       level by math7-math10@1;
             trend BY math7*-3 math8*-2 math9@-1 math10@0;
             [math7-math10@0];
             [level trend];
             level trend ON mothed homeres;

!New Version 3 Language For Growth Models
!MODEL: level trend | math7*-3 math8*-2 math9@-1 math10@0;
!       level trend ON mothed homeres;
```


Output Excerpts LSAY Growth Model With Free Time Scores And Covariates Centered At Grade 10

n = 935

Tests of Model Fit

```
CHI-SQUARE TEST OF MODEL FIT

          Value                        15.845
          Degrees of Freedom                7
          P-Value                        .0265

RMSEA (ROOT MEAN SQUARE ERROR OF APPROXIMATION)

          Estimate                       .037
          90 Percent C.I.                .012     .061
          Probability RMSEA <= .05       .794
```

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates Centered At Grade 10 (Continued)

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
LEVEL ON					
MOTHED	2.418	0.353	6.851	0.238	0.229
HOMERES	1.903	0.229	8.294	0.187	0.277
TREND ON					
MOTHED	0.111	0.073	1.521	0.094	0.090
HOMERES	0.161	0.049	3.311	0.136	0.201

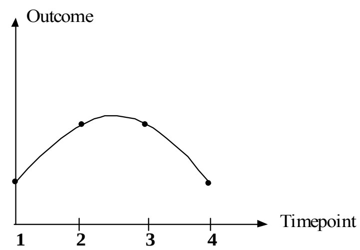
Further Practical Issues

Specifying Time Scores For Quadratic Growth Models

Quadratic Growth Model

$$y_{it} = \eta_{0i} + \eta_{1i} x_t + \eta_{2i} x_t^2 + \varepsilon_{it}$$

- Need three latent variables to describe a quadratic growth model: Intercept, Linear Slope, Quadratic Slope



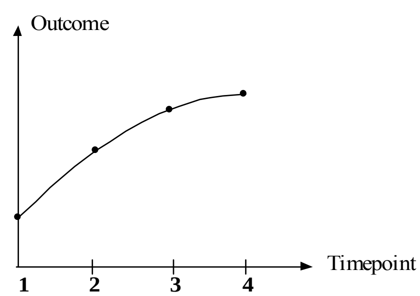
- Linear slope time scores: 0 1 2 3
  0 .1 .2 .3
- Quadratic slope time scores: 0 1 4 9
  0 .01 .04 .09

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

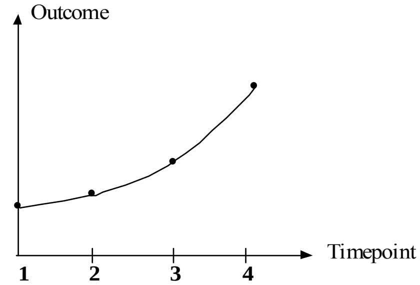
Growth model with a logarithmic growth curve--ln(t)



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

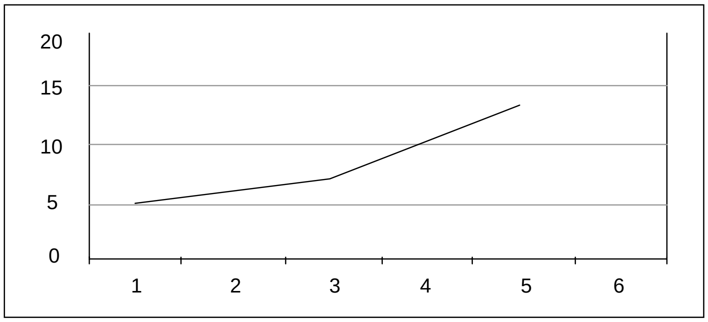


Time scores: 0 1.72 6.39 19.09

Piecewise Growth Modeling

Piecewise Growth Modeling

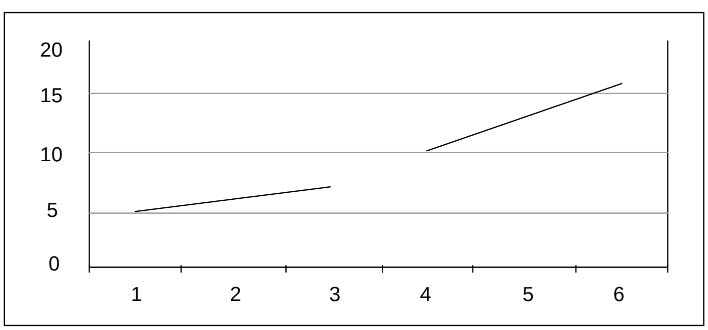
- Can be used to represent phases of development
- Can be used to capture non-linear growth
- Each piece has its own growth factor(s)
- Each piece can have its own coefficients for covariates



One intercept growth factor, two slope growth factors

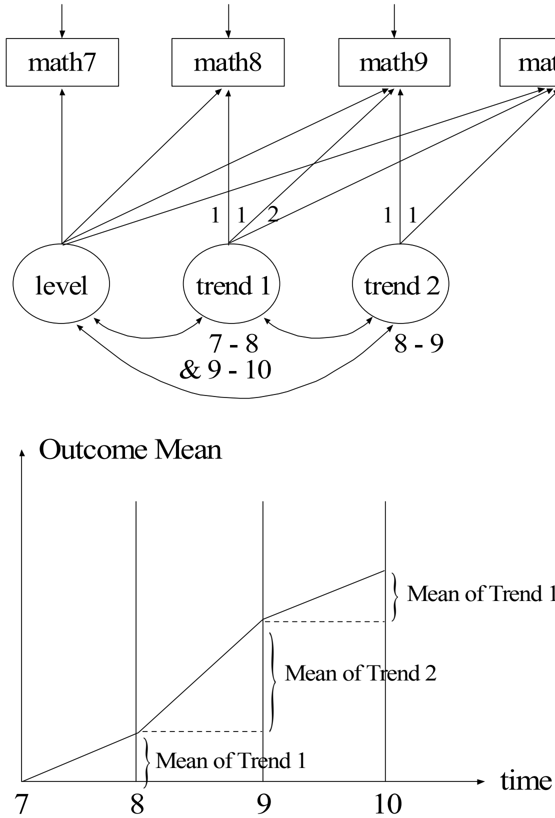
0	1	2	2	2	2	Time scores piece 1
0	0	0	1	2	3	Time scores piece 2

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



Input For LSAY Piecewise Growth Model With Covariates

```
TITLE:      LSAY For Younger Females With Listwise Deletion
            Piecewise Growth Model With Covariates
DATA:       FILE IS lsay.dat;
            FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:   NAMES ARE cohort id school weight math7 math8
            math9 math10 att7 att8 att9 att10 gender mothed homeres;
            USEOBS = (gender EQ 1 AND cohort EQ 2);
            MISSING = ALL (999);
            USEVAR = math7-math10 mothed homeres;
ANALYSIS:   TYPE = MEANSTRUCTURE;
MODEL:      level BY math7-math10@1;
            trend BY math7@0 math8@1 math9@1 math10@2;
            trend BY math7@0 math8@0 math9@1 math10@1;
            [math7-math10@0];
            [level trend1 trend2];
            level trend1 trend2 ON mothed homeres;
!New Version 3 Language For Growth Models
!MODEL: level trend1 | math7@0 math8@1 math9@1 math10@2;
!       level trend2 | math7@0 math8@0 math9@1 math10@1;
!       level trend1 trend2 ON mothed homeres;
```

Output Excerpts LSAY Piecewise Growth Model With Covariates

n = 935

Tests of Model Fit

```
CHI-SQUARE TEST OF MODEL FIT

          Value                          11.721
          Degrees of Freedom                  3
          P-Value                          .0083

RMSEA (ROOT MEAN SQUARE ERROR OF APPROXIMATION)

          Estimate                          .056
          90 Percent C.I.                   .025    .091
          Probability RMSEA <= .05          .331
```


Output Excerpts LSAY Piecewise Growth Model With Covariates (Continued)

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
LEVEL ON					
MOTHED	2.127	.284	7.488	.266	.256
HOMERES	1.389	.185	7.524	.174	.257
TREND1 ON					
MOTHED	-.126	.147	-.858	-.113	-.109
HOMERES	.091	.096	.950	.081	.120
TREND2 ON					
MOTHED	.436	.191	2.285	.185	.178
HOMERES	.289	.124	2.329	.123	.181

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 on the outcome, e.g. math achievement
a_{1ti} : time-related variable (time scores); 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:

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

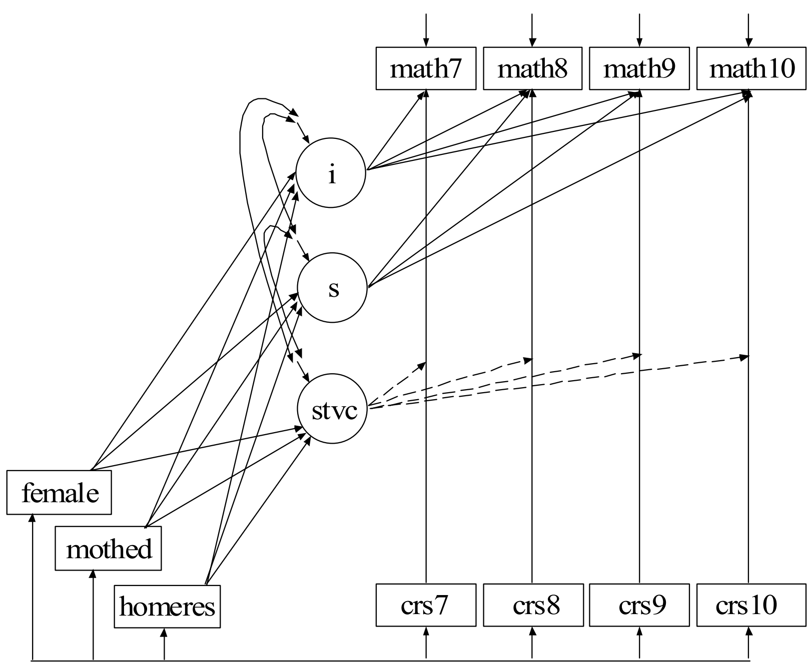
$$Level\ 2: \begin{cases} \pi_{0i} = \beta_{00} + \beta_{01}\, x_i + r_{0i}, \\ \pi_{1i} = \beta_{10} + \beta_{11}\, x_i + r_{1i}, \\ \pi_{2i} = \beta_{20} + \beta_{21}\, x_i + r_{2i}. \end{cases} \qquad (56)$$

Growth Modeling In Multilevel Terms (Continued)

Time scores a_{1ti} read in as data (not loading parameters).

- π_{2ti} possible with time-varying random slope variances
- Flexible correlation structure for $V(e) = \Theta$ (T x T)
- Regressions among random coefficients possible, e.g.

$$\pi_{1i} = \beta_{10} + \gamma_1 \, \pi_{0i} + \beta_{11} \, x_i + r_{1i}, \qquad (57)$$
$$\pi_{2i} = \beta_{20} + \gamma_2 \, \pi_{0i} + \beta_{21} \, x_i + r_{2i}. \qquad (58)$$



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

          HO 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
    MTH7                  0.000         0.000          0.000
    MTH8                  0.000         0.000          0.000
    MTH9                  0.000         0.000          0.000
    MTH10                 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
    MTH7                  0.185         0.011         16.464
    MTH8                  0.178         0.008         22.232
    MTH9                  0.156         0.008         18.497
    MTH10                 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
```


Random Slopes

- In single-level modeling random slopes β_i describe variation across individuals *i*,

$$y_i = \alpha_i + \beta_i x_i + \varepsilon_i, \quad (100)$$
$$\alpha_i = \alpha + \zeta_{0i}, \quad (101)$$
$$\beta_i = \beta + \zeta_{1i}, \quad (102)$$

Resulting in heteroscedastic residual variances

$$V(y_i \mid x_i) = V(\beta_i)\ x_i^2 + \theta \quad (103)$$

- In two-level modeling random slopes β_j describe variation across clusters *j*

$$y_{ij} = a_j + \beta_j x_{ij} + \varepsilon_{ij}, \quad (104)$$
$$a_j = a + \zeta_{0j}, \quad (105)$$
$$\beta_j = \beta + \zeta_{1j}. \quad (106)$$

A small variance for a random slope typically leads to slow convergence of the ML-EM iterations. This suggests respecifying the slope as fixed.

Mplus allows random slopes for predictors that are

- Observed covariates
- Observed dependent variables (Version 3)
- Continuous latent variables (Version 3)

Regressions Among Random Effects

Regressions Among Random Effects

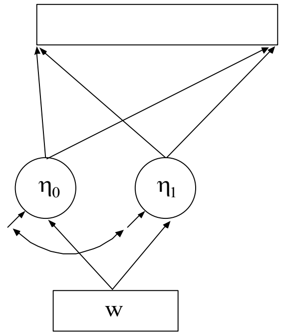
Standard multilevel model (where $x_t = 0, 1, \ldots, T$):

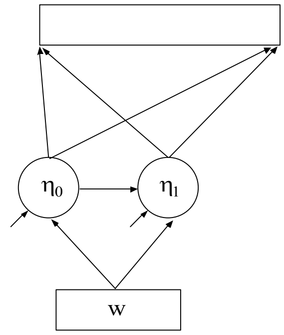
$$Level - 1: y_{it} = \eta_{0i} + \eta_i \, x_t + \varepsilon_{it}, \quad (1)$$
$$Level - 2a: \eta_{0i} = \alpha_0 + \gamma_0 \, w_i + \zeta_{0i}, \quad (2)$$
$$Level - 2b: \eta_{1i} = \alpha_1 + \gamma_1 \, w_i + \zeta_{1i}. \quad (3)$$

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 Modeling With Parallel Processes

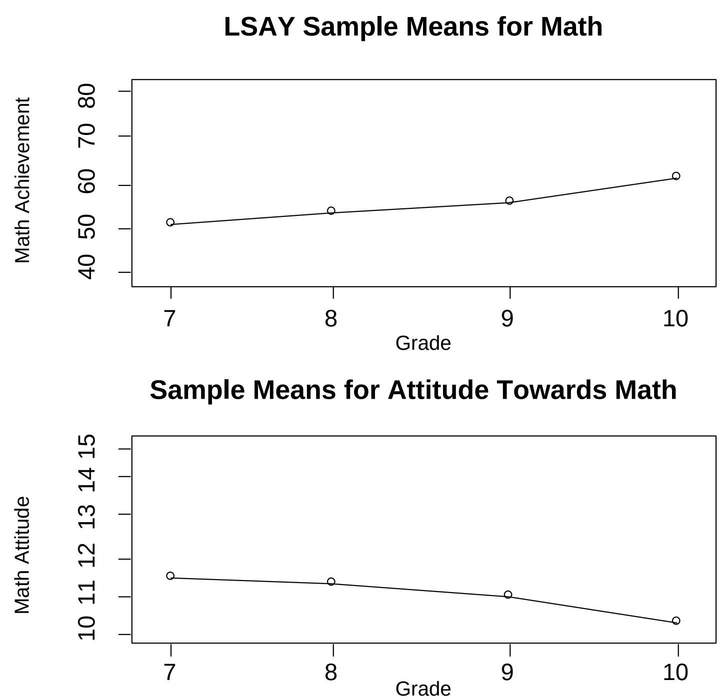
Growth Modeling With Parallel Processes

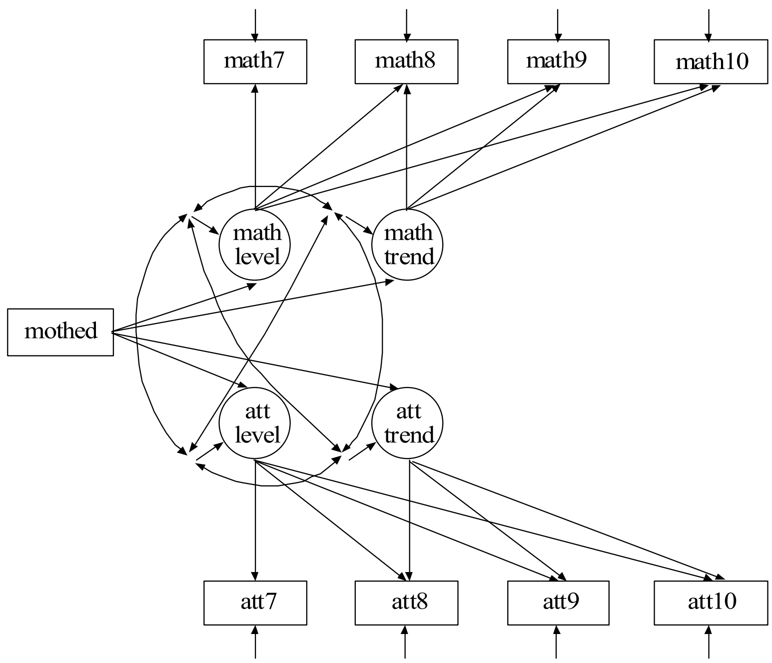
- Estimate a growth model for each process separately
  - Determine the shape of the growth curve
  - Fit model without covariates
  - Modify the model
- Joint analysis of both processes
- Add covariates

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 included math and science achievement items, math and science attitude measures, and background information from parents, teachers, and school principals. There were 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.





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;

ANALYSIS:  TYPE = MEANSTRUCTURE;
```

Input For LSAY Parallel Process Growth Model

```
MODEL:     levmath BY math7-math10@1;
           trndmath BY math7@0 math8@1 math9 math10;

           levatt BY att7-att10@1;
           trndatt BY att7@0 att8@1 att9@2 att10@3;

           [math7-math10@0 att7-att10@0];
           [levmath trndmath levatt trndatt];

           levmath-trndatt ON mothed;

OUTPUT     MODINDICES STANDARDIZED;

!New Version 3 Language For Growth Models
!MODEL: levmath trndmath | math7@0 math8@1 math9 math10;
!       levatt trndatt   | att7@0 att8@1 att9@2 att10@3;
!        levmath-trndatt 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)

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
LEVMATH ON					
MOTHED	**2.462**	**.280**	**8.798**	**.311**	**.303**
TRNDMATH ON					
MOTHED	**.145**	**.066**	**2.195**	**.132**	**.129**
LEVATT ON					
MOTHED	**.053**	**.086**	**.614**	**.025**	**.024**
TRNDATT ON					
MOTHED	**.012**	**.035**	**.346**	**.017**	**.017**

Output Excerpts LSAY Parallel Process Growth Model (Continued)

Selected Estimates (Continued)

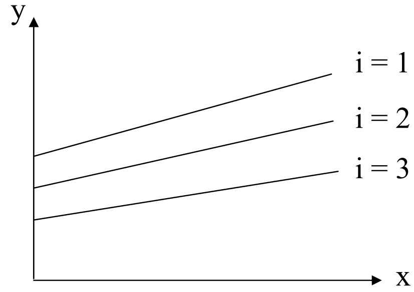
	Estimates	S.E.	Est./S.E.	Std	StdYX
TRNDMATH WITH					
LEVMATH	3.032	.580	5.224	.350	.350
LEVATT WITH					
LEVMATH	4.733	.702	6.738	.282	.282
TRNDMATH	.544	.164	3.312	.235	.235
TRNDATT WITH					
LEVMATH	-.276	.279	-.987	-.049	-.049
TRNDMATH	.130	.066	1.976	.168	.168
LEVATT	-.567	.115	-4.913	-.378	-.378

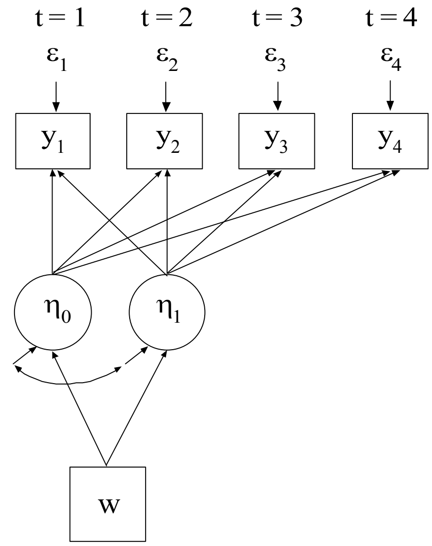
Computational Issues For Growth Models

- 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

Advanced Growth Models

Individual Development Over Time





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

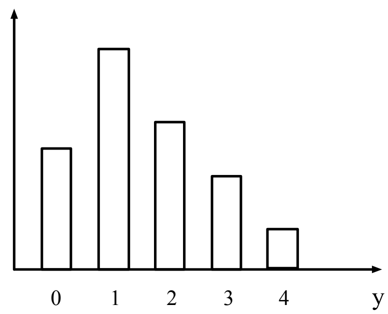
(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

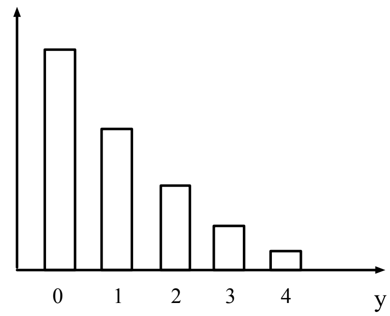
(2b) $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$

Advantages of Growth Modeling in a Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Random effects (intercepts, slopes) integrated with other latent variables
- Regressions among random effects
- Multiple processes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

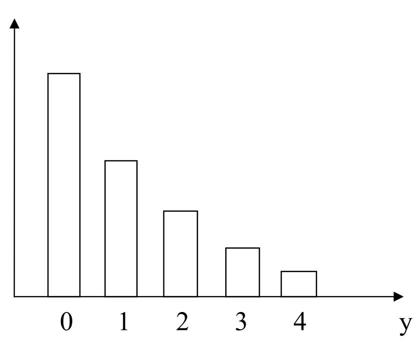
Modeling With A Preponderance Of Zeros

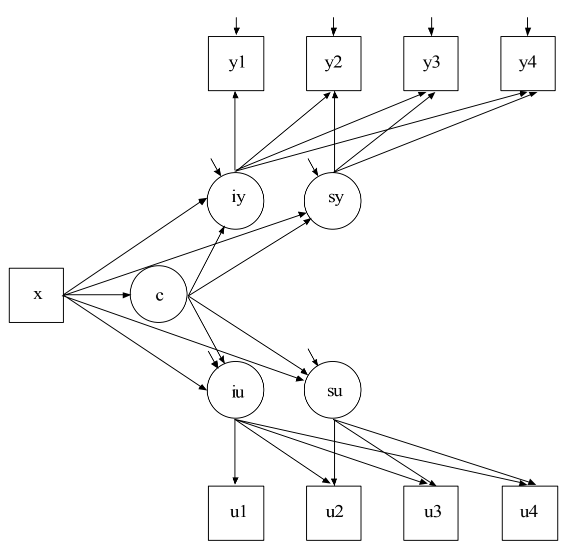




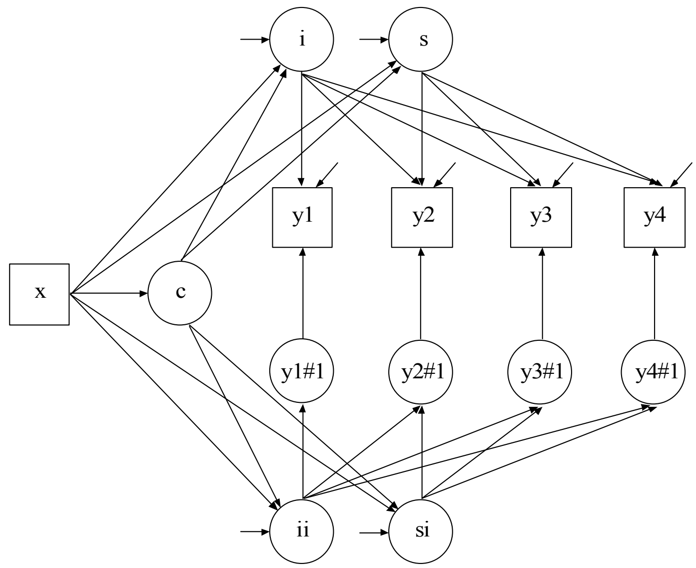
- Outcomes: non-normal continuous – count – categorical
- Censored-normal modeling
- Two-part (semicontinuous modeling): Duan et al. (1983), Olsen & Schafer (2001)
- Mixture models, e.g. zero-inflated (mixture) Poisson (Roeder et al., 1999), censored-inflated, mover-stayer latent transition models, growth mixture models
- Onset (survival) followed by growth: Albert & Shih (2003)

Two-Part (Semicontinuous) Growth Modeling

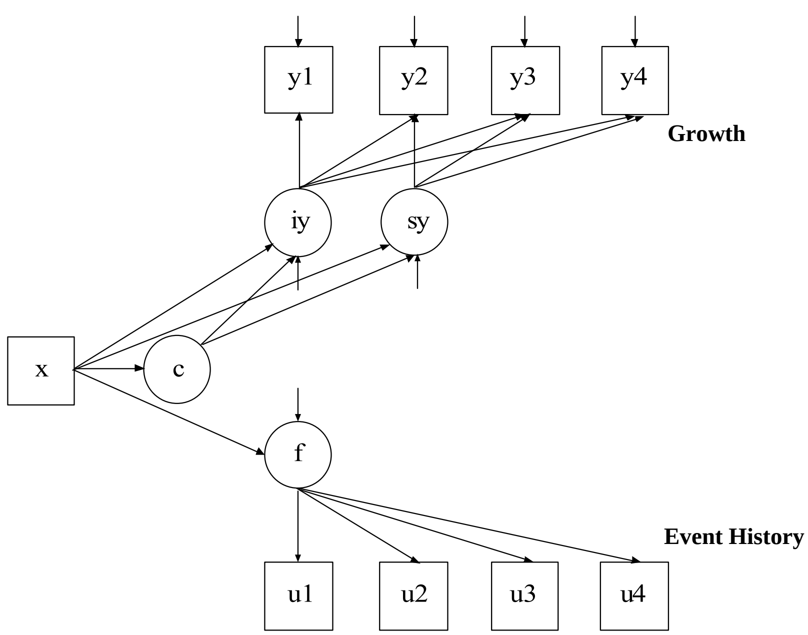




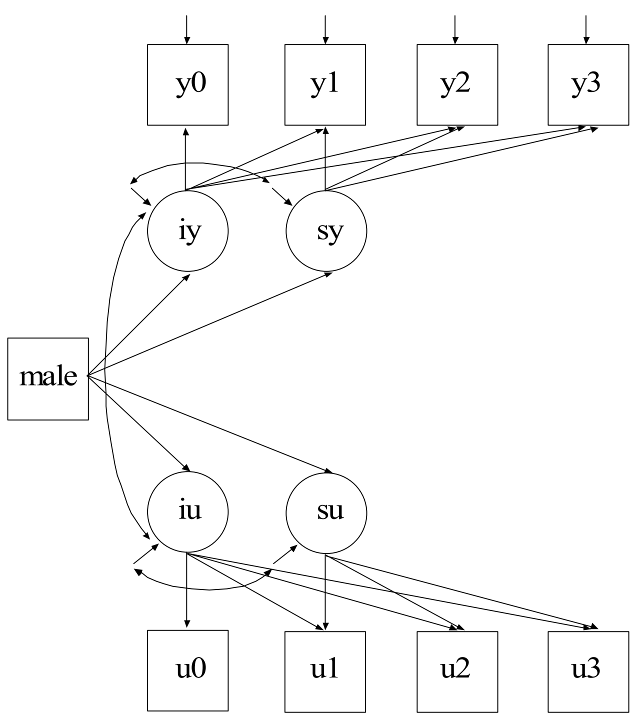
Inflated Growth Modeling (Two Classes At Each Time Point)



Onset (Survival) Followed By Growth



Two-Part Growth Modeling



Input For Step 1 Of A Two-Part Growth Model

```
TITLE:     step 1 of a two-part growth model
              Amover u    y
               >0    1   >0
                0    0   999
               999  999  999
DATA:      FILE = amp.dat;
VARIABLE:  NAMES ARE caseid
           amover0 ovrdrnk0 illdrnk0 vrydrn0
           amover1 ovrdrnk1 illdrnk1 vrydrn1
           amover2 ovrdrnk2 illdrnk2 vrydrn2
           amover3 ovrdrnk3 illdrnk3 vrydrn3
           amover4 ovrdrnk4 illdrnk4 vrydrn4
           amover5 ovrdrnk5 illdrnk5 vrydrn5
           amover6 ovrdrnk6 illdrnk6 vrydrn6
           tfq0-tfq6 v2 sex race livewith
           agedrnk0-agedrnk6 grades0-grades6;
           USEV = amover0 amover1 amover2 amover3
           sex race u0-u3 y0-y3;
           !MISSING = ALL (999);
```


Input For Step 1 Of A Two-Part Growth Model (Continued)

```
DEFINE:     u0 = 1;                    !binary part of variable
            IF(amover0 eq 0) THEN u0 = 0;
            IF(amover0 eq 999) THEN u0 = 999;
            y0 = amover0;              !continuous part of variable
            IF (amover0 eq 0) THEN y0 = 999;
            u1 = 1;
            IF(amover1 eq 0) THEN u1 = 0;
            IF(amover1 eq 999) THEN u1 = 999:
            y1 = amover1;
            IF(amover1 eq 0) THEN y1 = 999;
            u2 = 1;
            IF(amover2 eq 0) THEN u2 = 0;
            IF(amover2 eq 999) THEN u2 = 999;
            y2 = amover2;
            IF(amover2 eq 0) THEN y2 = 999;
            u3 = 1;
            IF(amover3 eq 0) THEN u3 = 0;
            IF(amover3 eq 999) THEN u3 = 999;
            y3 = amover3;
            IF(amover3 eq 0) THEN y3 = 999;
ANALYSIS:   TYPE = BASIC;
SAVEDATA:   FILE = ampyu.dat;
```

Output Excerpts Step 1 Of A Two-Part Growth Model

SAVEDATA Information

```
Order and format of variables
   AMOVER0   F10.3
   AMOVER1   F10.3
   AMOVER2   F10.3
   AMOVER3   F10.3
   SEX       F10.3
   RACE      F10.3
   U0        F10.3
   U1        F10.3
   U2        F10.3
   U3        F10.3
   Y0        F10.3
   Y1        F10.3
   Y2        F10.3
   Y3        F10.3
Save file
     ampyu.dat
Save file format
     14F10.3
Save file record length    1000
```


Input For Step 2 Of A Two-Part Growth Model

```
TITLE:       two-part growth model with linear growth for both
             parts
DATA:        FILE = ampyau.dat;
VARIABLE:    NAMES = amover0-amover3 sex race u0-u3 y0-y3;
             USEV = u0-u3 y0-y3 male;
             USEOBS = u0 NE 999;
             MISSING = ALL (999);
             CATEGORICAL = u0-u3;
DEFINE:      Male = 2-sex;
```

Input For Step 2 Of A Two-Part Growth Model (Continued)

```
ANALYSIS:  TYPE = MISSING;
           ESTIMATOR = ML;
           ALGORITHM = INTEGRATION;
           COVERAGE = .09;
MODEL:     iu su | u0@0 u1@0.5 u2@1.5 u3@2.5;
           iy sy | y0@0 y1@0.5 y2@1.5 y3@2.5;
           iu-sy ON male;
           ! estimate the residual covariances
           ! iu with su, iy with sy, and iu with iy
           iu WITH sy@0;
           su WITH iy-sy@0;
OUTPUT:    PATTERNS SAMPSTAT STANDARDIZED TECH1 TECH4 TECH8;
PLOT:      TYPE = PLOT3;
           SERIES = u0-u3(su) | y0-y3(sy);
```


Output Excerpts Step 2 Of A Two-Part Growth Model

Tests of Model Fit

```
Loglikelihood

          H0 Value                     -3277.101

Information Criteria

          Number of Free parameters          19
          Akaike (AIC)                 6592.202
          Bayesian (BIC)               6689.444
          Sample-Size Adjusted BIC     6629.092
             (n* = (n + 2) / 24)
```

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Model Results

	Estimates	S.E.	Est./S.E.	Std	StdYX
IU |					
U0	1.000	0.000	0.000	2.839	0.843
U1	1.000	0.000	0.000	2.839	0.882
U2	1.000	0.000	0.000	2.839	0.926
U3	1.000	0.000	0.000	2.839	0.905
SU |					
U0	0.000	0.000	0.000	0.000	0.000
U1	0.500	0.000	0.000	0.416	0.129
U2	1.500	0.000	0.000	1.249	0.407
U3	2.500	0.000	0.000	2.082	0.664

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

IY |					
Y0	1.000	0.000	0.000	0.534	0.787
Y1	1.000	0.000	0.000	0.534	0.738
Y2	1.000	0.000	0.000	0.534	0.740
Y3	1.000	0.000	0.000	0.534	0.644
SY |					
Y0	0.000	0.000	0.000	0.000	0.000
Y1	0.500	0.000	0.000	0.117	0.162
Y2	1.500	0.000	0.000	0.351	0.487
Y3	2.500	0.000	0.000	0.586	0.707

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

```
IU       ON
   MALE            0.569    0.234      2.433    0.200    0.100
SU       ON
   MALE           -0.181    0.119     -1.518   -0.218   -0.109
IY       ON
   MALE            0.149    0.061      2.456    0.279    0.139
SY       ON
   MALE           -0.068    0.038     -1.790   -0.290   -0.145

IU       WITH
   SU             -1.144    0.326     -3.509   -0.484   -0.484
   IY              1.193    0.134      8.897    0.788    0.788
   SY              0.000    0.000      0.000    0.000    0.000
IY       WITH
   SY             -0.039    0.019     -2.109   -0.316   -0.316
SU       WITH
   IY              0.000    0.000      0.000    0.000    0.000
   SY              0.000    0.000      0.000    0.000    0.000
```


Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

```
Intercepts
   Y0                0.000    0.000    0.000    0.000    0.000
   Y1                0.000    0.000    0.000    0.000    0.000
   Y2                0.000    0.000    0.000    0.000    0.000
   Y3                0.000    0.000    0.000    0.000    0.000
   IU                0.000    0.000    0.000    0.000    0.000
   SU                0.855    0.098    8.716    1.027    1.027
   IY                0.232    0.059    3.901    0.435    0.435
   SY                0.240    0.031    7.830    1.025    1.025

Thresholds
   U0$1              2.655    0.206   12.877
   U1$1              2.655    0.206   12.877
   U2$1              2.655    0.206   12.877
   U3$1              2.655    0.206   12.877
```

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Residual Variances

Y0	0.175	0.032	5.470	0.175	0.380
Y1	0.266	0.029	9.159	0.266	0.509
Y2	0.238	0.027	8.810	0.238	0.457
Y3	0.269	0.054	5.014	0.269	0.392
IU	7.982	1.086	7.351	0.990	0.990
SU	0.685	0.202	3.400	0.988	0.988
IY	0.279	0.040	7.019	0.981	0.981
SY	0.054	0.017	3.224	0.979	0.979

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

R-Square

Observed Variable	R-Square
U0	0.710
U1	0.682
U2	0.650
U3	0.666
Y0	0.620
Y1	0.491
Y2	0.543
Y3	0.608

Latent Variable	R-Square
IU	0.010
SU	0.012
IY	0.019
SY	0.021

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Technical 4 Output

ESTIMATED MEANS FOR THE LATENT VARIABLES

	IU	SU	IY	SY	MALE
1	0.305	0.758	0.312	0.204	0.536

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	IU	SU	IY	SY	MALE
IU	8.062				
SU	-1.170	0.694			
IY	1.214	-0.007	0.285		
SY	-0.010	0.003	-0.042	0.055	
MALE	0.142	-0.045	0.037	-0.017	0.249

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

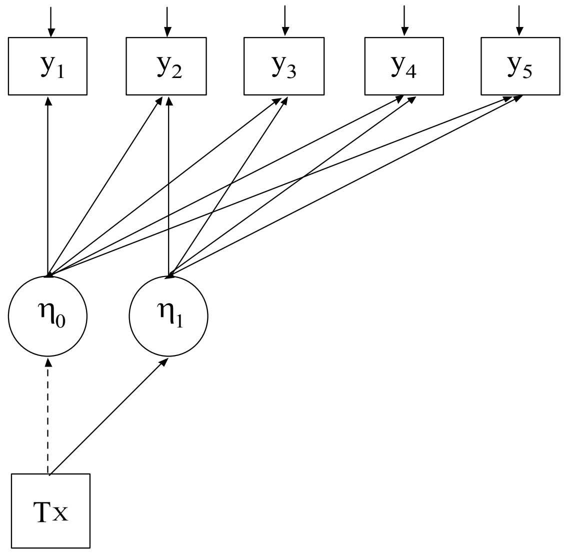
	IU	SU	IY	SY	MALE
IU	1.000				
SU	-0.495	1.000			
IY	0.801	-0.015	1.000		
SY	-0.014	0.016	-0.336	1.000	
MALE	0.100	-0.109	0.139	-0.145	1.000

Multiple Populations

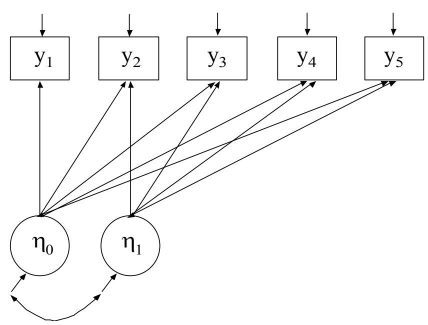
Multiple Population Growth Modeling

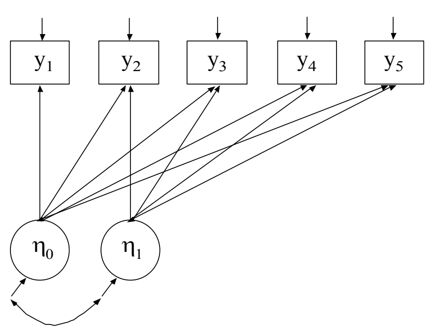
- Group as a dummy variable
- Multiple-group analysis
- Multiple-group analysis of randomized interventions

Group Dummy Variable as a Covariate



Two-Group Model





NLSY: Multiple Cohort Structure

Birth Year Cohort	Age[a]																			
	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
57								82	83	84	85	86	87	88	89	90	91	92	93	94
58							82	83	84	85	86	87	88	89	90	91	92	93	94	
59						82	83	84	85	86	87	88	89	90	91	92	93	94		
60					82	83	84	85	86	87	88	89	90	91	92	93	94			
61				82	83	84	85	86	87	88	89	90	91	92	93	94				
62			82	83	84	85	86	87	88	89	90	91	92	93	94					
63		82	83	84	85	86	87	88	89	90	91	92	93	94						
64	82	83	84	85	86	87	88	89	90	91	92	93	94							

[a] Non-shaded areas represent years in which alcohol measures were obtained

Preventive Interventions Randomized Trials

Prevention Science Methodology Group (PSMG)

Developmental Epidemiological Framework:

- Determining the levels and variation in risk and protective factors as well as developmental paths within a defined population in the absence of intervention
- Directing interventions at these risk and protective factors in an effort to change the developmental trajectories in a defined population
- Evaluating variation in intervention impact across risk levels and contexts on proximal and distal outcomes, thereby empirically testing the developmental model

Aggressive Classroom Behavior: The GBG Intervention

Muthén & Curran (1997, Psychological Methods)

The Johns Hopkins Prevention Center carried out a school-based preventive intervention randomized trial in Baltimore public schools starting in grade 1. One of the interventions tested was the Good Behavior Game intervention, a classroom based behavior management strategy promoting good behavior. It was designed specifically to reduce aggressive behavior of first graders and was aimed at longer term impact on aggression through middle school.

One first grade classroom in a school was randomly assigned to receive the Good Behavior Game intervention and another matched classroom in the school was treated as control. After an initial assessment in fall of first grade, the intervention was administered during the first two grades.

The GBG Aggression Example (Continued)

The outcome variable of interest was teacher ratings (TOCA-R) of each child's aggressive behavior (Breaks rules, harms property, fights) in the classroom through grades 1 – 6. Eight teacher ratings were made from fall and spring for the first two grades and every spring in grades 3 – 6.

The most important scientific question was whether the Good Behavior Game reduces the slope of the aggression trajectory across time. It was also of interest to know whether the intervention varies in impact for children who started out as high aggressive versus low aggressive.

Analyses in Muthén-Curran (1997) were based on data for 75 boys in the GBG group who stayed in the intervention condition for two years and 111 boys in the control group.

The GBG Aggression Example: Analysis Results

Muthén & Curran (1997):

- Step 1: Control group analysis;
- Step 2: Treatment group analysis;
- Step 3: Two-group analysis w/out interactions;
- Step 4: Two-group analysis with interactions;
- Step 5: Sensitivity analysis of final model
- Step 6: Power analysis

MUTHÉN AND CURRAN

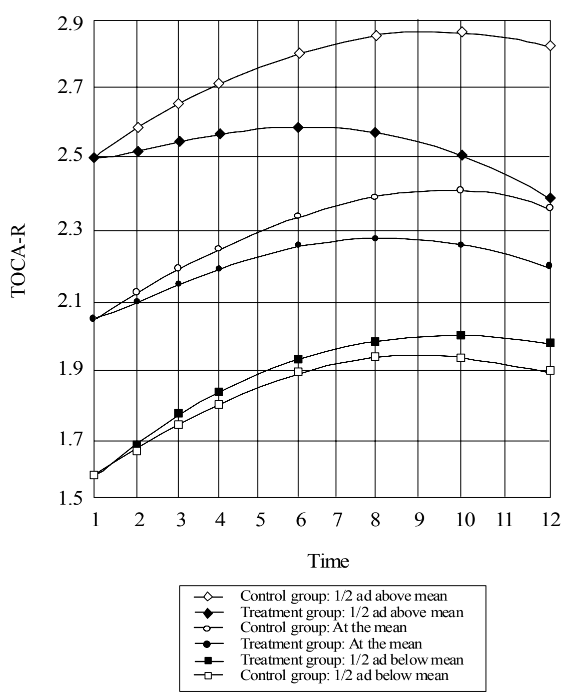
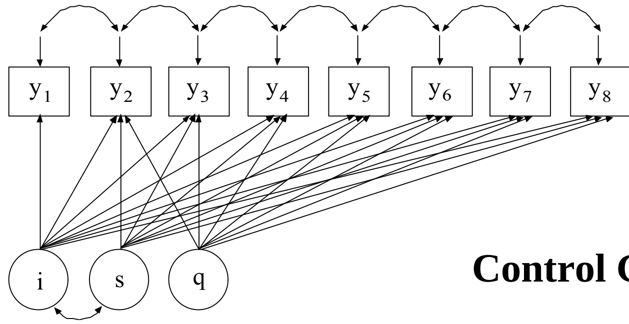
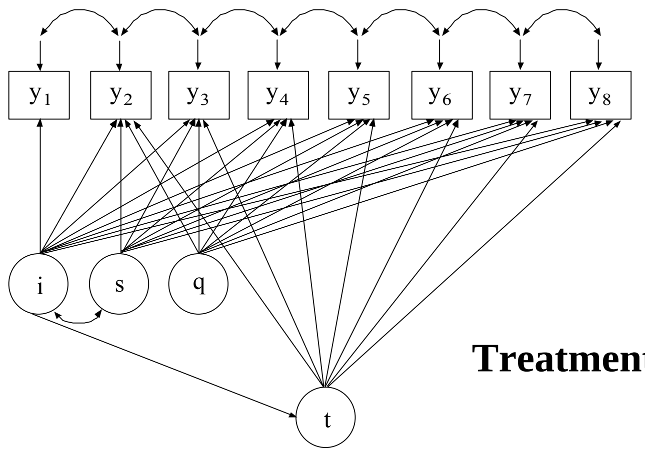


Figure 15. Model implied growth trajectories of Teacher Observation of Classroom Behavior—Revised (TOCA-R) scores as a function of initial status. Each timepoint represents one 6-month interval.



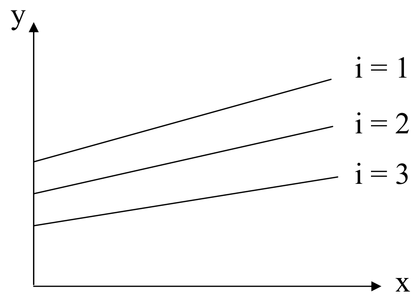
Control Group

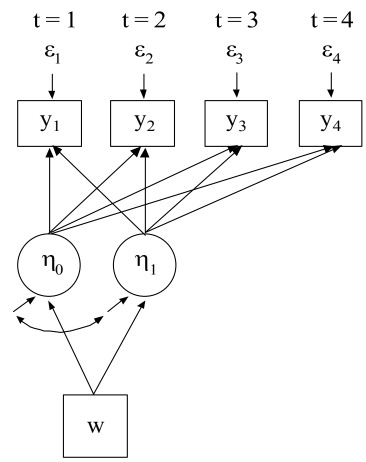


Treatment Group

Growth Mixture Modeling

Individual Development Over Time





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

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

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

Mixtures and Latent Trajectory Classes

Modeling motivated by substantive theories of:

- Multiple Disease Processes: Prostate cancer (Pearson et al.)
- Multiple Pathways of Development: Adolescent-limited versus life-course persistent antisocial behavior (Moffitt), crime curves (Nagin), alcohol development (Zucker, Schulenberg)
- Subtypes: Subtypes of alcoholism (Cloninger, Zucker)

Example: Mixed-effects Regression Models for Studying the Natural History of Prostate Disease

Pearson, Morrell, Landis and Carter (1994).
Statistics in Medicine

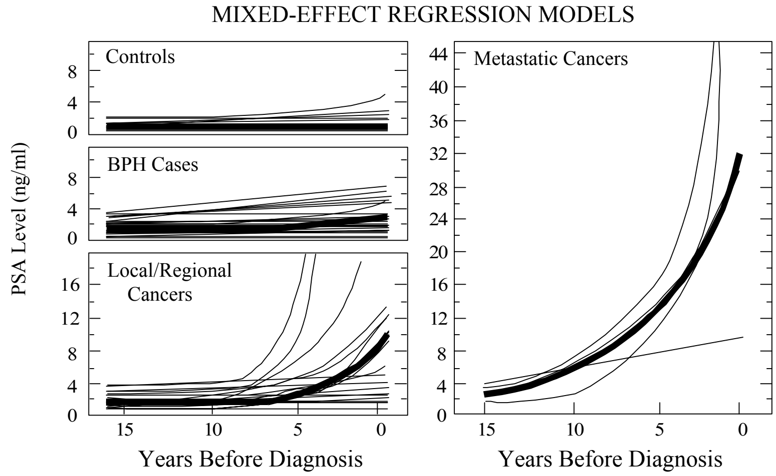
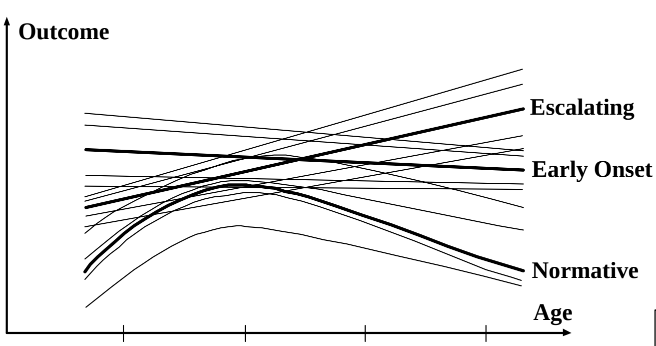
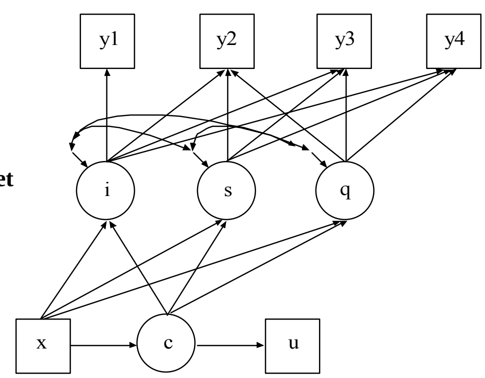


Figure 2. Longitudinal PSA curves estimated from the linear mixed-effects model for the group average (thick solid line) and for each individual in the study (thin solid lines)

Growth Mixture Modeling Of Developmental Pathways





Growth Mixture Modeling In Randomized Trials

- Growth mixture modeling of control group describing normative growth
- Class-specific treatment effects in terms of changed trajectories
- Muthén, Brown et al. (2002) in Biostatistics – application to an aggressive behavior preventive intervention in Baltimore public schools

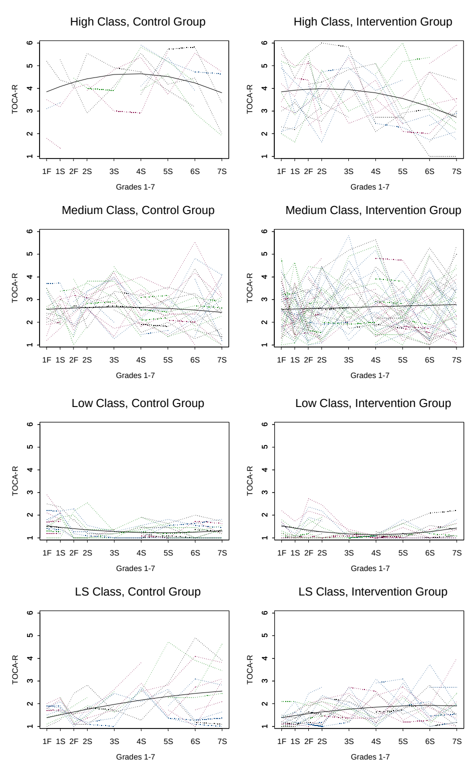
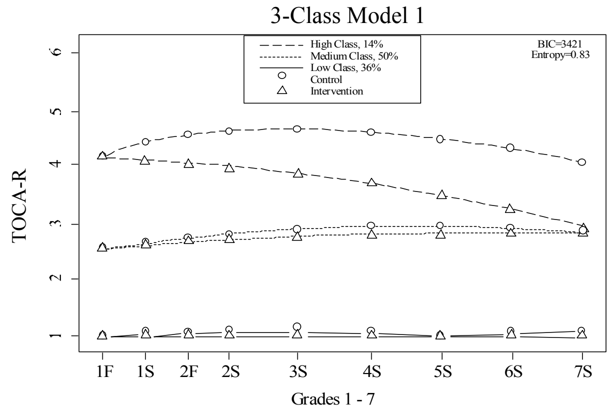
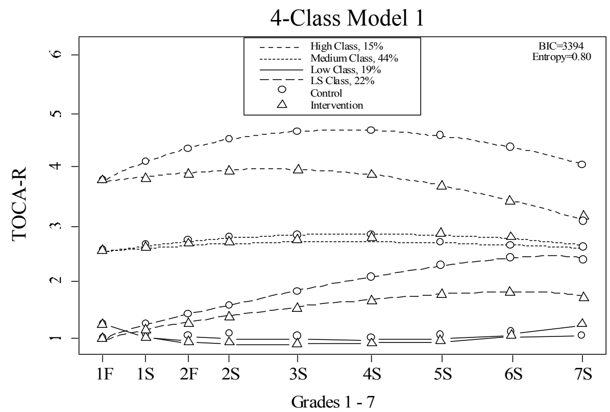


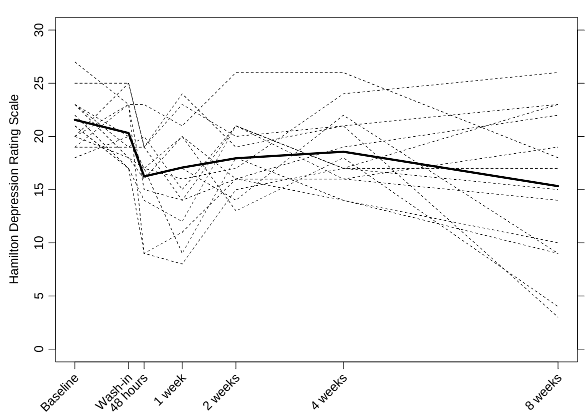
Figure 6. *Estimated Mean Growth Curves and Observed Trajectories for 4-Class model 1 by Class and Intervention Status*



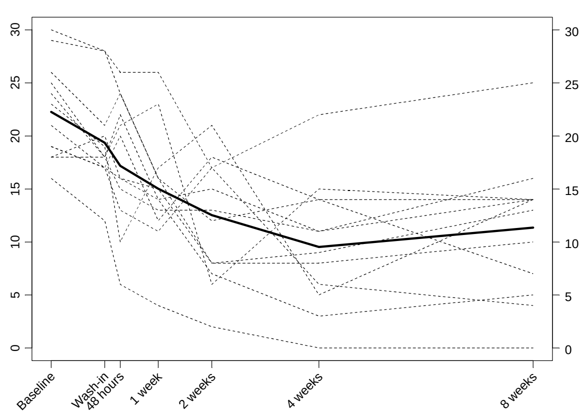


A Clinical Trial of Depression Medication: 2-Class Growth Mixture Modeling

Placebo Non-Responders, 55%

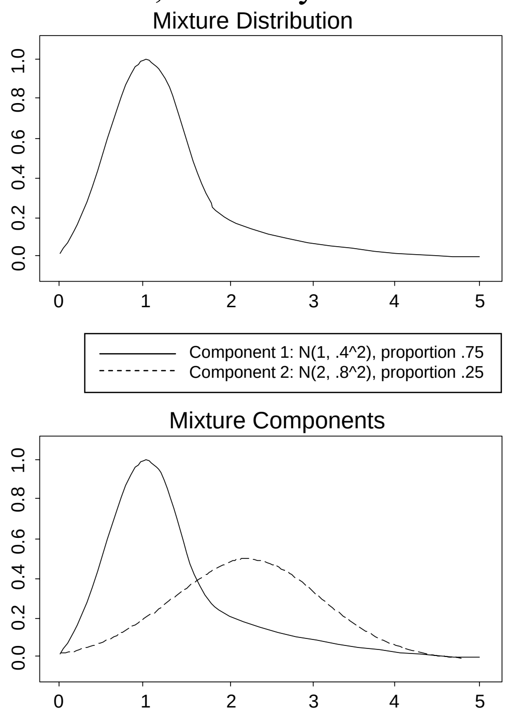


Placebo Responders, 45%



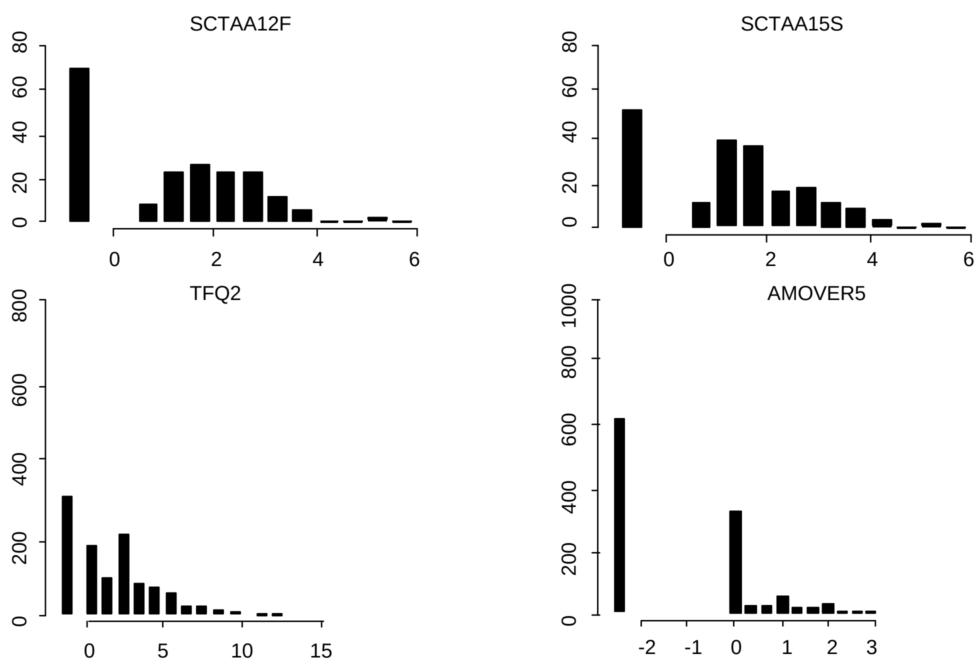
Mixture Distributions

Non-normality for mixture, normality for mixture components.



Two Types of Distribution

(1) Normal mixture components, (2) Preponderance of zeroes (Muthén, 2001, Two-part growth mixture modeling).



Growth Mixture Analysis

Generalization of conventional random effect growth modeling (multilevel modeling) to include qualitatively different developments (Muthén & Shedden, 1999 in Biometrics; Muthén 2004 in Handbook chapter).
Combination of conventional growth modeling and cluster analysis (finite mixture analysis).

- Setting
  - Longitudinal data
  - A single or multiple items measured repeatedly
  - Hypothesized trajectory classes (categorical latent variable)
  - Individual trajectory variation within classes

Growth Mixture Analysis

- Aim
  - Estimate trajectory shapes
  - Estimate trajectory class probabilities
  - Estimate variation within class
  - Relate class probabilities to covariates
  - Classify individuals into classes (posterior prob's)
  - Relate within-class variation to covariates

Application: Mathematics achievement, grades 7 – 10 (LSAY) related to mother's education and home resources.
National sample, n = 846.

A Strategy For Finding The Number Of Classes In Growth Mixture Modeling

- **Comparing models with different numbers of classes**
  - BIC – low BIC value corresponds to a high loglikelihood value and a parsimonious model
  - TECH11 – Lo-Mendell-Rubin likelihood ratio test (Biometrika, 2001)
- **Residuals and model tests**
  - TECH7 – class-specific comparisons of model-estimated means, variances, and covariances versus posterior probability-weighted sample statistics
  - TECH12 – class-mixed residuals for univariate skewness and kurtosis
  - TECH13 – multivariate skew and kurtosis model tests

A Strategy For Finding The Number Of Classes In Growth Mixture Modeling

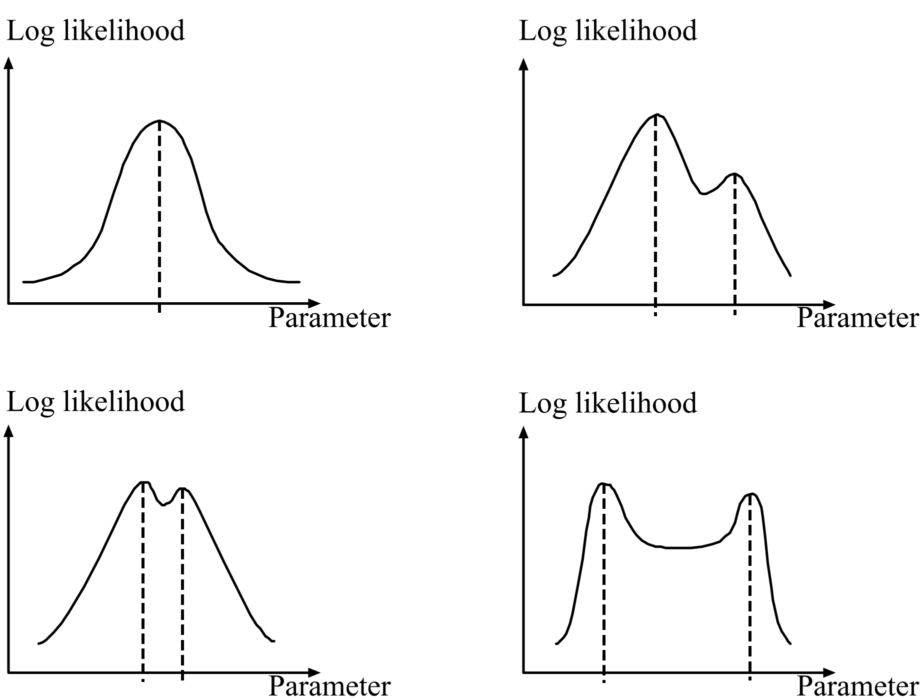
- **Classification quality**
  - Posterior probabilities – classification table and entropy
  - Individual trajectory classification using pseudo classes (Bandeen-Roche et al., 1997; Muthén et al. in Biostatistics, 2002)
- **Interpretability and usefulness of the latent classes**
  - Trajectory shapes
  - Number of individuals in each class
  - Number of estimated parameters
  - Substantive theory
  - Auxiliary (external) variables – predictive validity

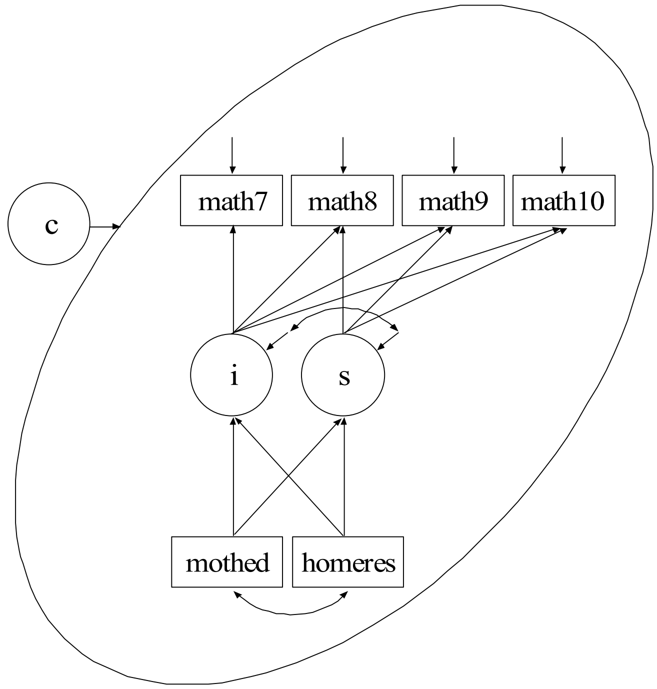
Strategies For Finding Starting Values In Growth Mixture Modeling

- **Strategy 1**
  - Do a conventional one-class analysis
  - Use estimated growth factor means and standard deviations as growth factor mean starting values in a multi-class model – mean plus and minus .5 standard deviation
- **Strategy 2**
  - Estimate a multi-class model with the variances and covariances of the growth factors fixed to zero
  - Use the estimated growth factor means as growth factor mean starting values for a model with growth factor variances and covariances free

New in Version 3 – random starts makes it unnecessary to give starting values: starts = 50 5

Global and Local Solutions



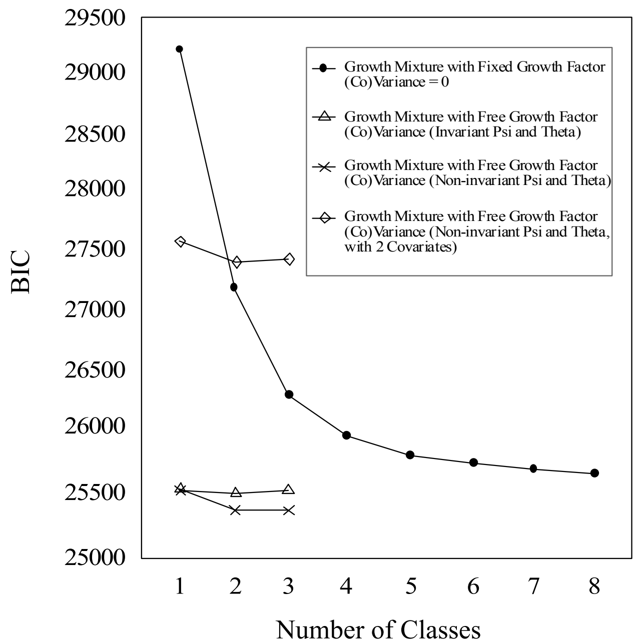


Deciding On The Number Of Classes For The LSAY Growth Mixture Model

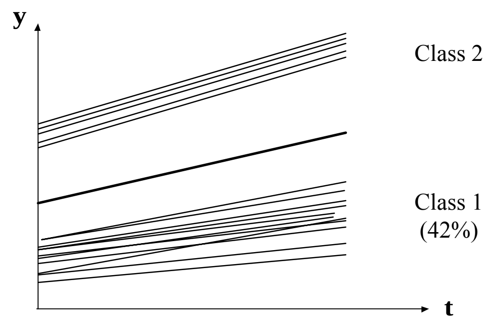
n = 935

Number of Classes	1	2	3
Loglikelihood	-11,997.653	-11,864.826	-11,856.220
# parameters	15	29	36
BIC	24,098	23,928	23,959
AIC	24,025	23,788	23,784
Entropy	NA	.468	.474
LRT p-value for k-1 classes	NA	.0000	.4041
Multivariate skew p-value	.00	.34	.26
Multivariate kurtosis p-value	.00	.10	.05

Model Fit by BIC: LSAY



LSAY: Estimated Two-Class Growth Mixture Model



Class 1

Intercept	ON Mothed*	Homeres
Slope	ON Mothed	Homeres*

Class 2

Intercept	ON Mothed*	Homeres*
Slope	ON Mothed	Homeres

Conventional Single-Class Analysis

Intercept	ON Mothed*	Homeres*
Slope	ON Mothed	Homeres*

Input For LSAY 2-Class Growth Mixture Model

```
TITLE:       2-class varying slopes on mothed and homeres
             varying Psi varying Theta
DATA:        FILE IS lsay.dat;
             FORMAT is 3f8 f8.4 8f8.2 2f8.2;
VARIABLE:    NAMES ARE cohort id school weight math7
             math8 math9 math10 att7 att8 att9 att10 gender mothed
             homeres;
             USEOBS = (gender EQ 1 AND cohort EQ 2);
             MISSING = ALL (999);
             USEVAR = math7-math10 mothed homeres;
             CLASSES = c(2);
ANALYSIS:    TYPE = MIXTURE;
```


Input For LSAY 2-Class Growth Mixture Model (Continued)

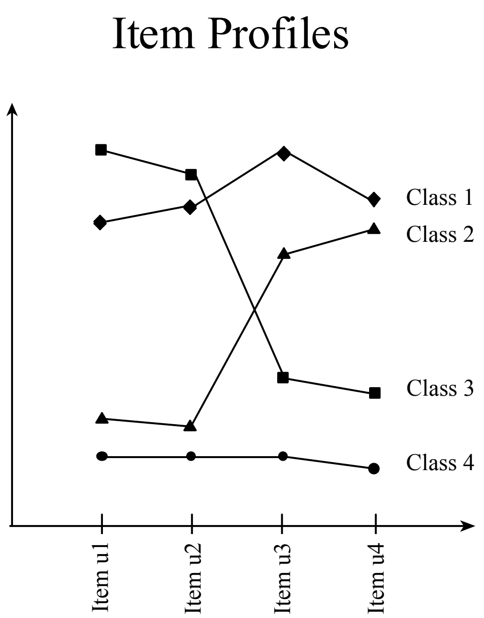
```
MODEL:       %OVERALL%
             intercpt BY math7-math10@1;
             slope BY math7@0 math8@1 math9 math10;
             [math7-math10@0];
             intercpt slope ON  mothed homeres;
             %c#1%                            !Not needed in Version 3
             [intercpt*42.8 slope*.6];        !Not needed in Version 3
             %c#2%
             [intercpt*62.8 slope*3.6];       !Not needed in Version 3
             intercpt slope ON mothed homeres;
             math7-math10 intercpt slope;
             slope WITH intercpt;
OUTPUT:      TECH8 TECH11 TECH12 TECH13 RESIDUAL;
```

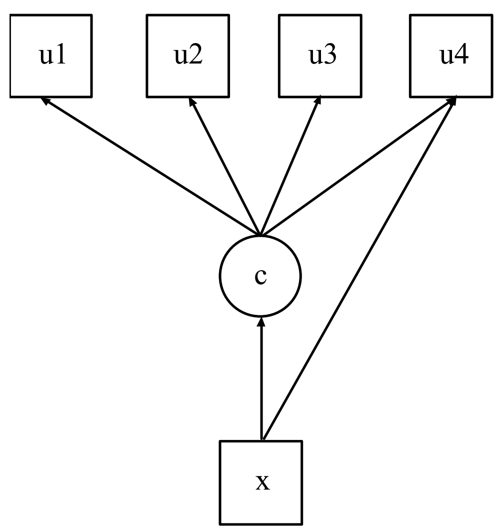
New Version 3 Language For Growth Models

```
MODEL:       %OVERALL%
             intercpt slope | math7@0 math8@1 math9 math10;
             intercpt slope ON mothed homeres;
             %c#2%
             intercpt slope ON mothed homeres;
             math7-math10 intercpt slope;
             slope WITH intercpt;
```

Latent Class Models

Latent Class Analysis





Confirmatory Latent Class Analysis With Several Latent Class Variables

Introduced by Goodman (1974); Bartholomew (1987). Special case for longitudinal data: latent transition analysis, introduced by Collins; latent Markov models.

- Setting
  - Cross-sectional or longitudinal data
  - Multiple items measuring several different constructs
  - Hypothesized simple structure for measurements
  - Hypothesized constructs represented as latent class variables (categorical latent variables)

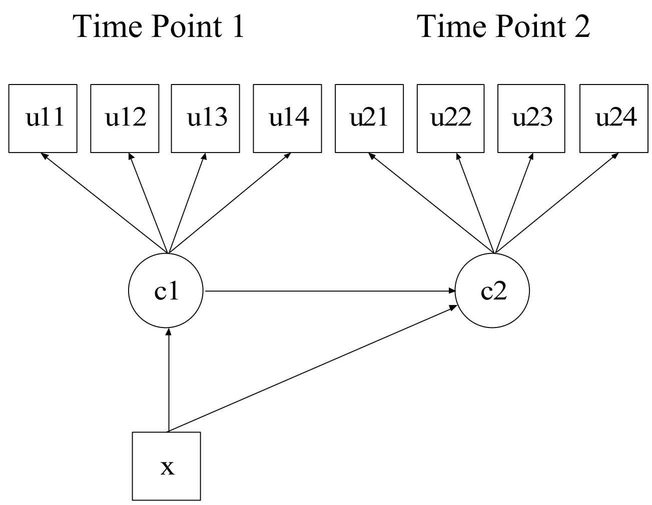
Confirmatory Latent Class Analysis With Several Latent Class Variables (Continued)

- Aim
  - Identify items that indicate classes well
  - Test simple measurement structure
  - Study relationships between latent class variables
  - Estimate class probabilities
  - Relate class probabilities to covariates
  - Classify individuals into classes (posterior probabilities)
- Applications
  - Latent transition analysis with four latent class indicators at two time points and a covariate

Latent Transition Analysis

Transition Probabilities

		c2: 1	c2: 2
c1	1	0.8	0.2
	2	0.4	0.6



Input For LTA With Two Time Points And A Covariate

```
TITLE:      Latent transition analysis for two time points and a
            covariate using Mplus Version 3
DATA:       FILE = mc2tx.dat;
VARIABLE:   NAMES ARE u11-u14 u21-u24 x c1 c2;
            USEV = u11-u14 u21-u24 x;
            CATEGORICAL = u11-u24;
            CLASSES = c1(2) c2(2);
ANALYSIS:   TYPE = MIXTURE;
MODEL:
            %OVERALL%
            c2#1 ON c1#1 x;
            c1#1 on x;
```

Input For LTA With Two Time Points And A Covariate (Continued)

```
MODEL c1:
            %c1#1%
            [u11$1] (1);
            [u12$1] (2);
            [u13$1] (3);
            [u14$1] (4);
            %c1#2%
            [u11$1] (5);
            [u12$1] (6);
            [u13$1] (7);
            [u14$1] (8);
MODEL c2:
            %c2#1%
            [u21$1] (1);
            [u22$1] (2);
            [u23$1] (3);
            [u24$1] (4);
            %c2#2%
            [u21$1] (5);
            [u22$1] (6);
            [u23$1] (7);
            [u24$1] (8);
OUTPUT:     TECH1 TECH8;
```


Output Excerpts LTA With Two Time Points And A Covariate

Tests Of Model Fit

```
Loglikelihood
             HO Value                      -3926.187
Information Criteria
             Number of Free Parameters            13
             Akaike (AIC)                   7878.374
             Bayesian (BIC)                 7942.175
             Sample-Size Adjusted BIC       7900.886
                (n* = (n + 2) / 24)
             Entropy                           0.902
```

Output Excerpts LTA With Two Time Points And A Covariate (Continued)

```
Chi-Square Test of Model Fit for the Latent Class Indicator Model
Part
            Pearson Chi-Square

            Value                          250.298
            Degrees of Freedom                 244
            P-Value                         0.3772

            Likelihood Ratio Chi-Square

            Value                          240.811
            Degrees of Freedom                 244
            P-Value                         0.5457
```

Final Class Counts

```
FINAL  CLASS COUNTS AND PROPORTIONS OF TOTAL SAMPLE SIZE BASED
ON ESTIMATED POSTERIOR PROBABILITIES
 Class 1         328.42644          0.32843
 Class 2         184.43980          0.18444
 Class 3         146.98726          0.14699
 Class 4         340.14650          0.34015
```


Output Excerpts LTA With Two Time Points And A Covariate (Continued)

Model Results

```
                        Estimates      S.E.    Est./S.E.

LATENT CLASS INDICATOR MODEL PART

Class 1-C1, 1-C2

  Thresholds
     U11$1                 -2.020      0.110     -18.353
     U12$1                 -2.003      0.106     -18.919
     U13$1                 -1.776      0.098     -18.046
     U14$1                 -1.861      0.101     -18.396
     U21$1                 -2.020      0.110     -18.353
     U22$1                 -2.003      0.106     -18.919
     U23$1                 -1.776      0.098     -18.046
     U24$1                 -1.861      0.101     -18.396
```

Output Excerpts LTA With Two Time Points And A Covariate (Continued)

Class 1-C1, 2-C2

Thresholds

U11$1	-2.020	0.110	-18.353
U12$1	-2.003	0.106	-18.919
U13$1	-1.776	0.098	-18.046
U14$1	-1.861	0.101	-18.396
U21$1	1.964	0.111	17.736
U22$1	2.164	0.119	18.113
U23$1	1.864	0.100	18.704
U24$1	2.107	0.112	18.879

Class 2-C1, 1-C2

Thresholds

U11$1	1.964	0.111	17.736
U12$1	2.164	0.119	18.113
U13$1	1.864	0.100	18.704
U14$1	2.107	0.112	18.879
U21$1	-2.020	0.110	-18.353
U22$1	-2.003	0.106	-18.919
U23$1	-1.776	0.098	-18.046
U24$1	-1.861	0.101	-18.396

Output Excerpts LTA With Two Time Points And A Covariate (Continued)

Model Results (Continued)

Class 2-C1, 2-C2

Thresholds

	Estimates	S.E.	Est./S.E.
U11$1	1.964	0.111	17.736
U12$1	2.164	0.119	18.113
U13$1	1.864	0.100	18.704
U14$1	2.107	0.112	18.879
U21$1	1.964	0.111	17.736
U22$1	2.164	0.119	18.113
U23$1	1.864	0.100	18.704
U24$1	2.107	0.112	18.879

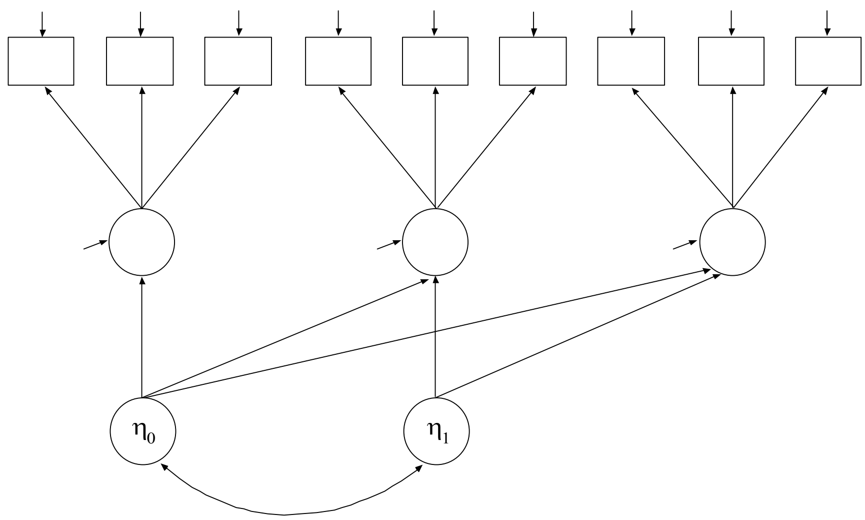
Output Excerpts LTA With Two Time Points And A Covariate (Continued)

```
LATENT CLASS REGRESSION MODEL PART

C2#1        ON
   C1#1                        0.530        0.180        2.953
C2#1        ON
   X                          -1.038        0.107       -9.703
C1#1        ON
   X                          -1.540        0.112      -13.761
Intercepts
   C1#1                        0.065        0.082        0.797
   C2#1                       -0.407        0.120       -3.381
```


Growth Modeling With Multiple Indicators

Growth Of Latent Variable Construct Measured By Multiple Indicators



Steps in Growth Modeling With Multiple Indicators

- Exploratory factor analysis of indicators for each timepoint
- Determine the shape of the growth curve for each indicator and the sum of the indicators
- Fit a growth model for each indicator—must be the same
- Confirmatory factor analysis of all timepoints together
  - Covariance structure analysis without measurement parameter invariance
  - Covariance structure analysis with invariant loadings
  - Mean and covariance structure analysis with invariant measurement intercepts and loadings
- Growth model with measurement invariance across timepoints

Advantages Of Using Multiple Indicators Instead Of An Average

- Estimation of unequal weights
- Partial measurement invariance—changes across time in individual item functioning
- No confounding of time-specific variance and measurement error variance
- Smaller standard errors for growth factor parameters (more power)

Classroom Aggression Data (TOCA)

The classroom aggression data are from an intervention study in Baltimore public schools carried out by the Johns Hopkins Prevention Research Center. Subjects were randomized into treatment and control conditions. The TOCA-R instrument was used to measure 10 aggression items at multiple timpoints. The TOCA-R is a teacher rating of student behavior in the classroom. The items are rated on a six-point scale from almost never to almost always.

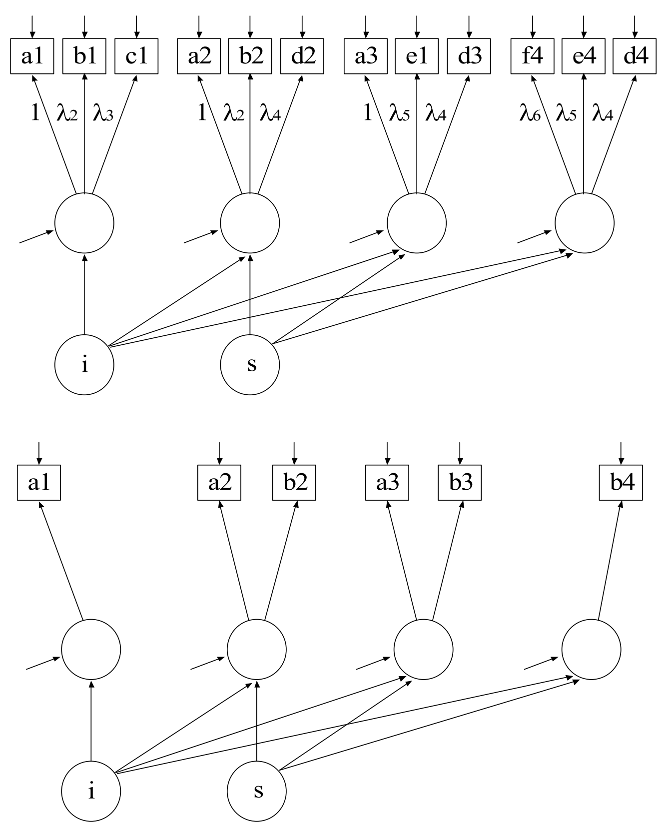
Data for this analysis include the 342 boys in the control group. Four time points are examined: Spring Grade 1, Spring Grade 2, Spring Grade 3, and Spring Grade 4.

Seven aggression items are used in the analysis:

- Break rules
- Fights
- Harms others
- Lies
- Stubborn
- Teasing classmates
- Yells at others

Degrees Of Invariance Across Time

- Case 1
  - Same items
  - All items invariant
  - Same construct
- Case 2
  - Same items
  - Some items non-invariant
  - Same construct
- Case 3
  - Different items
  - Some items invariant
  - Same construct
- Case 4
  - Different items
  - Some items invariant
  - Different construct

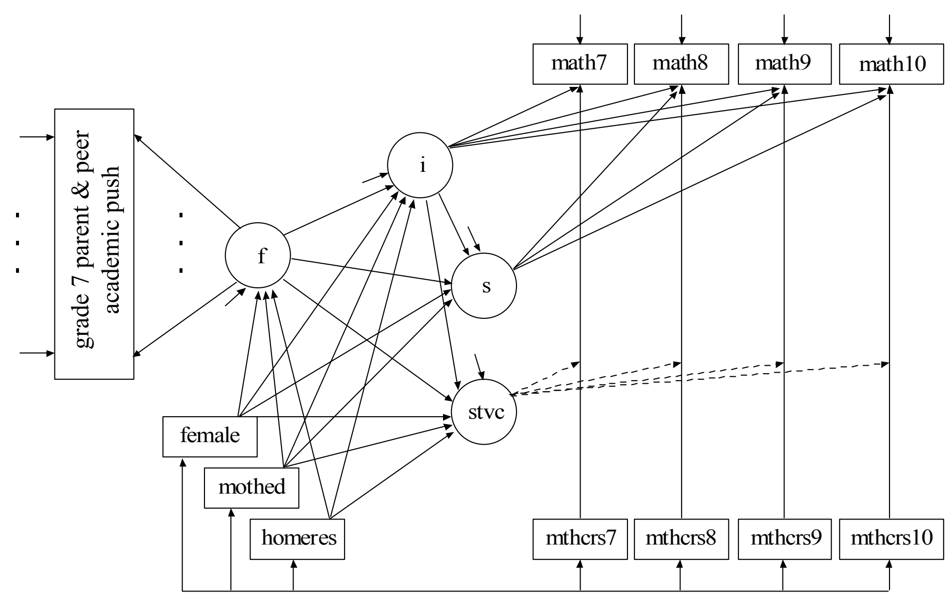


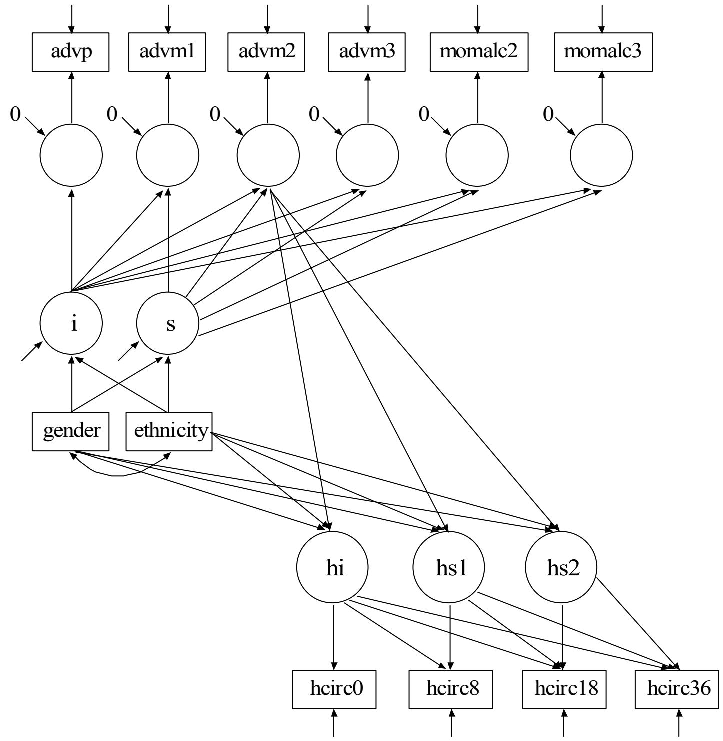
Power For Growth Models

Designing Future Studies: Power

- Computing power for growth models using Satorra-Saris (Muthén & Curran, 1997; examples)
- Computing power using Monte Carlo studies (Muthén & Muthén, 2002)
- Power calculation web site – PSMG
- Multilevel power (Miyazaki & Raudenbush, 2000; Moerbeek, Breukelen & Berger, 2000; Raudenbush, 1997; Raudenbush & Liu, 2000)
- School-based studies (Brown & Liao, 1999: Principles for designing randomized preventive trials)
- Multiple- (sequential-) cohort power
- Designs for follow-up (Brown, Indurkhia, Kellam, 2000)

Embedded Growth Models





References

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

Analysis With Longitudinal Data

Introductory

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.E., Duncan, S.C., Strycker, L.A., Li, F., & Alpert, A. (1999). An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications. Mahwah, NJ: Lawrence Erlbaum Associates.

Goldstein, H. (1995). Multilevel statistical models. Second edition. London: Edward Arnold.

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

References (Continued)

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

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.

References (Continued)

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

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

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. & 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.

References (Continued)

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.

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.

References (Continued)

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.

Advanced

Albert, P.A., McShane, L.M., Shih, J.H. & The US SCI Bladder Tumor Marker Network (2001). Latent class modeling approaches for assessing diagnostic error without a gold standard: With applications to p.53 immunohistochemical assays in bladder tumors. Biometrics, 57, 610-619.

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.

References (Continued)

Collins, L.M. & Sayer, A. (Eds.), New Methods for the Analysis of Change. Washington, D.C.: American Psychological Association.

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.

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

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.

References (Continued)

Mixture References

Analysis With Categorical Latent Variables (Mixture Modeling)

General

Agresti, A. (1990). Categorical data analysis. New York: John Wiley & Sons.

Everitt, B.S. & Hand, D.J. (1981). Finite mixture distributions. London: Chapman and Hall.

McLachlan, G.J. & Peel, D. (2000). Finite mixture models. New York: Wiley & Sons.

Muthén, L.K. & Muthén, B. (1998-2001). Mplus User's Guide. Los Angeles, CA: Muthén & Muthén.

Schwartz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 6, 461-464.

Titterington, D.M., Smith, A.F.M., & Makov, U.E. (1985). Statistical analysis of finite mixture distributions. Chichester, U.K.: John Wiley & Sons.

Lo, Y., Mendell, N.R. & Rubin, D.B. (2001). Testing the number of components in a normal mixture. Biometrika, 88, 767-778.

References (Continued)

Vuong, Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, 57, 307-333.

Asparouhov, T. & Muthén, B. (2002). Skew and kurtosis tests in mixture modeling.

Latent Class Analysis

Bandeen-Roche, K., Miglioretti, D.L., Zeger, S.L. & Rathouz, P.J. (1997). Latent variable regression for multiple discrete outcomes. Journal of the American Statistical Association, 92, 1375-1386.

Bartholomew, D.J. (1987). Latent variable models and factor analysis. New York: Oxford University Press.

Bucholz, K.K., Heath, A.C., Reich, T., Hesselbrock, V.M., Kramer, J.R., Nurnberger, J.I., & Schuckit, M.A. (1996). Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multi-center family study of alcoholism. Alcohol Clinical Experimental Research, 20, 1462-1471.

Clogg, C.C. (1995). Latent class models. In G. Arminger, C.C. Clogg & M.E. Sobel (eds.), Handbook of statistical modeling for the social and behavioral sciences (pp 331-359). New York: Plenum Press.

References (Continued)

Clogg, C.C. & Goodman, L.A. (1985). Simultaneous latent structural analysis in several groups. In Tuma, N.B. (ed.), Sociological Methodology, 1985 (pp. 18-110). San Francisco: Jossey-Bass Publishers.

Dayton, C.M. & Macready, G.B. (1988). Concomitant variable latent class models. Journal of the American Statistical Association, 83, 173-178.

Formann, A.K. (1992). Linear logistic latent class analysis for polytomous data. Journal of the American Statistical Association, 87, 476-486.

Goodman, L.A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika, 61, 215-231.

Hagenaars, J.A. & McCutcheon, A.L. (2002). Applied latent class analysis. Cambridge, UK: Cambridge University Press.

Heijden, P.G.M., Dressens, J. & Bockenholt, U. (1996). Estimating the concomitant-variable latent-class model with the EM algorithm. Journal of Educational and Behavioral Statistics, 21, 215-229.

Lazarsfeld, P.F. & Henry, N.W. (1968). Latent structure analysis. New York: Houghton Mifflin.

Muthén, B. (2001). Latent variable mixture modeling. In G.A. Marcoulides & R.E. Schumacker (eds.), New Developments and Techniques in Structural Equation Modeling (pp. 1-33). Lawrence Erlbaum Associates.

References (Continued)

Muthén, B. & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. Alcoholism: Clinical and Experimental Research, 24, 882-891.

Nestadt, G., Hanfelt, J., Liang, K.Y., Lamacz, M., Wolyniec, P., Pulver, A.E. (1994). An evaluation of the structure of schizophrenia spectrum personality disorders. Journal of Personality Disorders, 8, 288-298.

Rindskopf, D. (1990). Testing developmental models using latent class analysis. In A. von Eye (ed.), Statistical methods in longitudinal research: Time series and categorical longitudinal data (Vol 2, pp. 443-469). Boston: Academic Press.

Rindskopf, D., & Rindskopf, W. (1986). The value of latent class analysis in medical diagnosis. Statistics in Medicine, 5, 21-27.

Stoolmiller, M. (2001). Synergistic interaction of child manageability problems and parent-discipline tactics in predicting future growth in externalizing behavior for boys. Developmental Psychology, 37, 814-825.

Uebersax, J.S., & Grove, W.M. (1990). Latent class analysis of diagnostic agreement. Statistics in Medicine, 9, 559-572.

References (Continued)

Latent Transition Analysis

Collins, L.M. & Wugalter, S.E. (1992). Latent class models for stage-sequential dynamic latent variables. Multivariate Behavioral Research, 27, 131-157.

Collins, L.M., Graham, J.W., Rouscult, S.S., & Hansen, W.B. (1997). Heavy caffeine use and the beginning of the substance use onset process: An illustration of latent transition analysis. In K. Bryant, M. Windle, & S. West (Eds.), The Science of Prevention: Methodological Advances from Alcohol and Substance Use Research. Washington DC: American Psychological Association. pp. 79-99.

Graham, J.W., Collins, L.M., Wugalter, S.E., Chung, N.K., & Hansen, W.B. (1991). Modeling transitions in latent stage- sequential processes: A substance use prevention example. Journal of Consulting and Clinical Psychology, 59, 48-57.

Kandel, D.B., Yamaguchi, K., & Chen, K. (1992). Stages of progression in drug involvement from adolescence to adulthood: Further evidence for the gateway theory. Journal of Studies of Alcohol, 53, 447-457.

Reboussin, B.A., Reboussin, D.M., Liang, K.Y., & Anthony, J.C. (1998). Latent transition modeling of progression of health-risk behavior. Multivariate Behavioral Research, 33, 457-478.

References (Continued)

Noncompliance (CACE)

Angrist, J.D., Imbens, G.W., Rubin. D.B. (1996). Identification of causal effects using instrumental variables. Journal of the American Statistical Association, 91, 444-445.

Jo, B. (1999). Estimation of intervention effects with noncompliance: Alternative model specifications. Forthcoming in Journal of Educational and Behavioral Statistics (will appear with comments).

Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. Psychological Methods, 7, 178-193.

Jo, B. (2002). Model misspecification sensitivity analysis in estimating causal effects of interventions with noncompliance. Statistics in Medicine.

Jo, B. & Muthén, B. (2000). Longitudinal studies with intervention and noncompliance: Estimation of casual effects in growth mixture modeling. To appear in N. Duan and S. Reise (Eds.), Multilevel Modeling: Methodological Advances, Issues, and Applications, Multivariate Applications Book Series, Lawrence Erlbaum Associates.

Jo, B. & Muthén, B. (2001). Modeling of intervention effects with noncompliance: A latent variable approach for randomized trials. In G.A. Marcoulides & R.E. Schumacker (eds.), New Developments and Techniques in Structural Equation Modeling (pp. 57-87). Lawrence Erlbaum Associates.

Little, R.J. & Yau, L.H.Y. (1998). Statistical techniques for analyzing data from prevention trials: treatment of no-shows using Rubin's causal model. Psychological Methods, 3, 147-159.

Factor Mixture Modeling, SEMM

Jedidi, K., Jagpal, H.S. & DeSarbo, W.S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved geterogeneity. Marketing Science, 16, 39-59.

Lubke, G. & Muthén, B. (2003). Performance of factor mixture models. Under review, Multivariate Behavioral Research.

Growth Mixtures, Latent Class Growth Analysis

Jones, B.L., Nagin, D.S. & Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. Sociological Methods & Research, 29, 374-393.

Land, K.C. (2001). Introduction to the special issue on finite mixture models. Sociological Methods & Research, 29, 275-281.

References (Continued)

Li, F., Duncan, T.E., Duncan, S.C. & Acock, A. (2001). Latent growth modeling of longitudinal data: a finite growth mixture modeling approach. Structural Equation Modeling, 8, 493-530.

Moffitt, T.E. (1993). Adolescence-limited and life-course persistent

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. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Collins, L.M. & Sayer, A. (eds.), New Methods for the Analysis of Change (pp. 291-322). Washington, D.C.: APA.

Muthén, B. (2001). Latent variable mixture modeling. In G.A. Marcoulides & R.E. Schumacker (eds.), New Developments and Techniques in Structural Equation Modeling (pp. 1-33). Lawrence Erlbaum Associates.

Muthén, B. (2001). Two-part growth mixture modeling. University of California, Los Angeles.

References (Continued)

Muthén, B. (2002). Beyond SEM: General latent variable modeling. Behaviormetrika, 29, 81-117. Muthén, B. & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. Alcoholism: Clinical and Experimental Research, 24, 882-891.

Muthén, B. & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. Alcoholism: Clinical and Experimental Research, 24, 882-891.

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

Muthén, B., Brown, C.H., Masyn, K., Jo, B., Khoo, S.T., Yang, C.C., Wang, C.P., Kellam, S., Carlin, J., & Liao, J. (2000). General growth mixture modeling for randomized preventive interventions. Biostatistics, 3, 459-475.

Muthén, B., Khoo, S.T., Francis, D. & Kim Boscardin, C. (2002). Analysis of reading skills development from Kindergarten through first grade: An application of growth mixture modeling to sequential processes. In S.R. Reise & N. Duan (eds), Multilevel Modeling: Methodological Advances, Issues, and Applications (pp. 71-89). Mahaw, NJ: Lawrence Erlbaum Associates.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semi-parametric, group-based approach. Psychological Methods, 4, 139-157.

References (Continued)

Nagin, D.S. (2005). Group-based modeling of development. Cambridge: Harvard University Press.

Nagin, D.S. & Land, K.C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. Criminology, 31, 327-362.

Nagin, D.S. & Tremblay, R.E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and non violent juvenile delinquency. Child Development, 70, 1181-1196.

Nagin, D.S. & Tremblay, R.E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. Psychological Methods, 6, 18-34.

Nagin, D.S., Farrington, D. & Moffitt, T. (1995). Life-course trajectories of different types of offenders. Criminology, 33, 111-139.

Nagin, D.S. & Tremblay, R.E. (2001). Analyzing developmental trajectories of distinct but related behaviors: a group-based method. Psychological Methods, 6, 18-34.

Pearson, J.D., Morrell, C.H., Landis, P.K., Carter, H.B., & Brant, L.J. (1994). Mixed-effect regression models for studying the natural history of prostate disease. Statistics in Medicine, 13, 587-601.

References (Continued)

Porjesz, B., & Begleiter, H. (1995). Event-related potentials and cognitive function in alcoholism. Alcohol Health & Research World, 19, 108-112.

Roeder, K., Lunch, 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.

Schulenberg, J., O'Malley, P.M., Bachman, J.G., Wadsworth, K.N., & Johnston, L.D. (1996). Getting drunk and growing up: Trajectories of frequent binge drinking during the transition to young adulthood. Journal of Studies on Alcohol, May, 289-304.

Verbeke, G., & Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. Journal of the American Statistical Association, 91, 217-221.

Zucker, R.A. (1994). Pathways to alcohol problems and alcoholism: A developmental account of the evidence for multiple alcoholisms and for contextual contributions to risk. In: R.A. Zucker, J. Howard & G.M. Boyd (Eds.), The development of alcohol problems: Exploring the biopsychosical matrix of risk (pp. 255-289) (NIAAA Research Monograph No. 26). Rockville, MD: U.S. Department of Health and Human Services.

References (Continued)

Discrete-Time Survival Analysis

Allison, P.D. (1984). Event History Analysis. Regression for Longitudinal Event Data. Quantitative Applications in the Social Sciences, No. 46. Thousand Oaks: Sage Publications.

Lin, H., Turnbull, B.W., McCulloch, C.E. & Slate, E. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. Journal of the American Statistical Association, 97, 53-65.

Muthén, B. & Masyn, K. (2001). Discrete-time survival mixture analysis.

Singer, J.D., and Willett, J.B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. Journal of Educational Statistics, 18(20), 155-195.

Vermunt, J.K. (1997). Log-linear models for event histories. Advanced quantitative techniques in the social sciences, vol. 8. Thousand Oaks: Sage Publications.