

Mplus Short Courses
Topic 8

**Multilevel Modeling With Latent
Variables Using Mplus:
Longitudinal Analysis**

Linda K. Muthén
Bengt Muthén

Copyright © 2011 Muthén & Muthén

www.statmodel.com

02/11/2011

1

Table Of Contents

General Latent Variable Modeling Framework	6
Multilevel Growth Models	10
Growth Models With Categorical Outcomes	16
Multivariate Approach To Multilevel Modeling	52
Multilevel Growth Modeling (3-Level Analysis)	57
Multilevel Growth Modeling Of Binary Outcomes (3-Level Analysis)	72
Two-Level Multiple Indicator Growth	92
Special Multilevel Growth Models	123
Multiple Indicator Growth Modeling As Two-Level Analysis	127
Multilevel Discrete-Time Survival Analysis	130
Two-Level Latent Transition Analysis	133
Multilevel Growth Mixture Modeling	139
Between-Level Latent Classes	157
Latent Transition Analysis	158
Growth Modeling	162
Growth Mixture Modeling	166
Technical Aspects Of Multilevel Modeling	172
References	181

2

Mplus Background

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

3

Mplus Background

- Mplus versions
 - V1: November 1998
 - V2: February 2001
 - V3: March 2004
 - V4: February 2006
 - V5: November 2007
 - V5.21: May 2009
 - V6: April, 2010
- Mplus team: Linda & Bengt Muthén, Thuy Nguyen, Tihomir Asparouhov, Michelle Conn, Jean Maninger

4

Statistical Analysis With Latent Variables A General Modeling Framework

Statistical Concepts Captured By Latent Variables

Continuous Latent Variables

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

Categorical Latent Variables

- Latent classes
- Clusters
- Finite mixtures
- Missing data

5

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

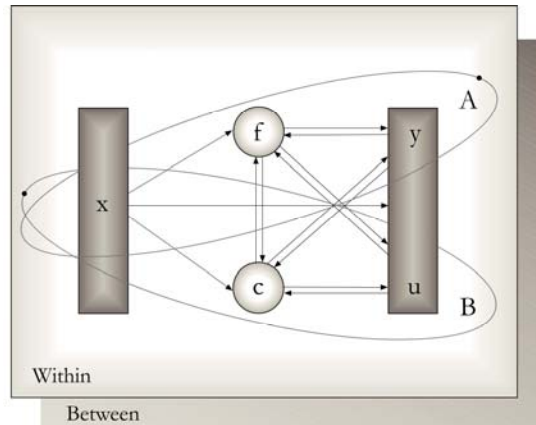
Categorical Latent Variables

- 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.

6

General Latent Variable Modeling Framework



7

Mplus

Several programs in one

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

Fully integrated in the general latent variable framework

8

Overview Of Mplus Courses

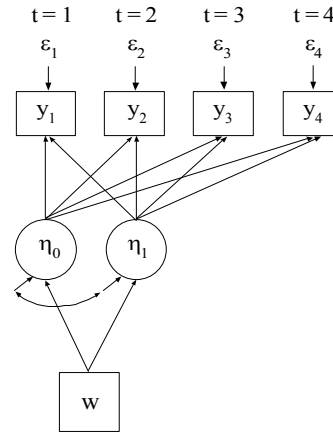
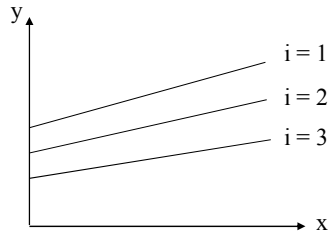
- **Topic 9.** Bayesian analysis using Mplus. University of Connecticut, May 24, 2011
- Courses taught by other groups in the US and abroad (see the Mplus web site)

9

Multilevel Growth Models

10

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

11

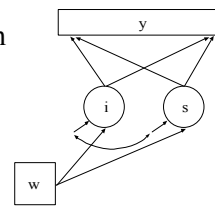
Growth Modeling Approached In Two Ways: Data Arranged As Wide Versus Long

- Wide: Multivariate, Single-Level Approach

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

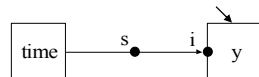
i_i regressed on w_i

s_i regressed on w_i

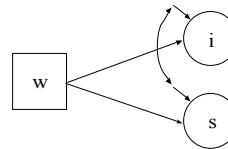


- Long: Univariate, 2-Level Approach (CLUSTER = id)

Within



Between



The intercept i is called y in Mplus

12

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

- Wide (one person):

	t1	t2	t3	t1	t2	t3		
Person i:	id	y1	y2	y3	x1	x2	x3	w

- Long (one cluster):

Person i:	t1	id	y1	x1	w
	t2	id	y2	x2	w
	t3	id	y3	x3	w

Mplus command: DATA LONGTOWIDE (UG ex 9.16)
DATA WIDETOLONG

13

Pros And Cons Of Wide Versus Long

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

14

Advantages Of Growth Modeling In A Latent Variable Framework

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

15

Growth Models With Categorical Outcomes

16

Growth Model With Categorical Outcomes

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

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

Level 1

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

$$\left. \begin{aligned} (2a) \quad \eta_{0i} &= \alpha_0 + \gamma_0 w_i + \zeta_{0i} \\ (2b) \quad \eta_{1i} &= \alpha_1 + \gamma_1 w_i + \zeta_{1i} \\ (2c) \quad \eta_{2i} &= \alpha_2 + \gamma_2 w_i + \zeta_{2i} \end{aligned} \right\} \text{Level 2}$$

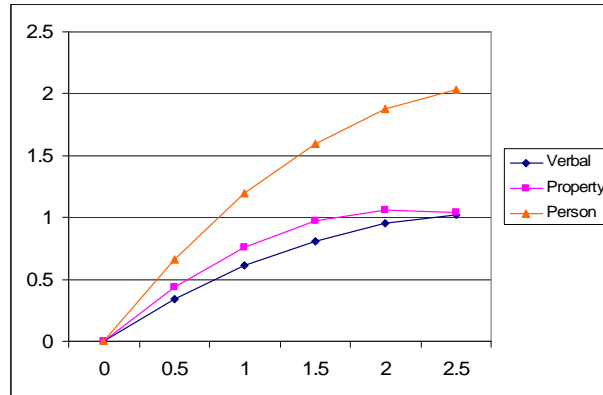
17

Aggression Growth Analysis

- Baltimore cohort 1: 1174 students in 41 classrooms (clustering due to classroom initially ignored)
- 8 time points over grades 1-7
- Quadratic growth
- Dichotomized items from the aggression instrument
- 4 analyses
 - Single item (“Breaks Things”), ignoring classroom clustering (ML uses 3 dimensions of numerical integration)
 - Multiple indicators (Property aggression factor), ignoring classroom clustering (ML: 8 dimensions; WLSM)
 - Single item with clustering (ML: 6 dimensions; WLSM)
 - Multiple indicators with clustering (ML: 16 dimensions; WLSM)

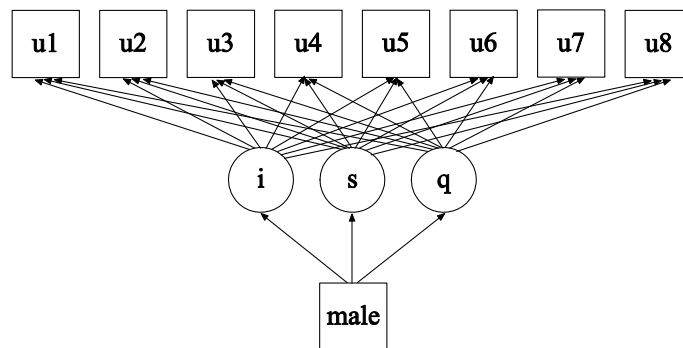
18

Different Growth Curves For Different Aspects Of Aggression



19

Binary Growth Modeling Of Aggression Item (No Clustering)



20

Input Binary Growth Modeling

```
TITLE: Hopkins Cohort 1 All time points with Classroom
Information
DATA: FILE = Cohort1_classroom_ALL.DAT;
VARIABLE: NAMES ARE PRCID

stub1F bkRule1F harm01F bkThin1F yell1F takeP1F fight1F
lies1F tease1F

stub1S bkRule1S harm01S bkThin1S yell1S takeP1S fight1S
lies1S tease1S

stub2S bkRule2S harm02S bkThin2S yell2S takeP2S fight2S
lies2S tease2S

stub3S bkRule3S harm03S bkThin3S yell3S takeP3S fight3S
lies3S tease3S
```

21

Input Binary Growth Modeling (Continued)

```
stub4S bkRule4S harm04S bkThin4S yell4S takeP4S
fight4S lies4S tease4S

stub5S bkRule5S harm05S bkThin5S yell5S takeP5S
fight5S lies5S tease5S

stub6S bkRule6S harm06S bkThin6S yell6S takeP6S
fight6S lies6S tease6S

stub7S bkRule7S harm07S bkThin7S yell7S takeP7S
fight7S lies7S tease7S

gender race des011 sch011 sec011 juv99 violchld
antisocr conductr

athort1F harmP1S athort1S harmP2S athort2S harmP3S
athort3S harmP4S athort4S harmP5S athort5S harmP6S
harmP7S athort7S
```

22

Input Binary Growth Modeling (Continued)

```
stubby2F bkRule2F harmO2F bkThin2F yell2F takeP2F  
fight2F harmP2F lies2F athort2F tease2F  
classrm;  
!  
    CLUSTER = classrm;  
USEVAR = bkthin1f bkthin1s bkthin2s bkthin3s bkthin4s  
bkthin5s bkthin6s bkthin7s male;  
CATEGORICAL = bkthin1f - bkthin7s;  
MISSING = ALL (999);  
DEFINE: CUT bkThin1f(1.5);  
CUT bkThin1s(1.5);  
CUT bkThin2s(1.5);  
CUT bkThin3s(1.5);  
CUT bkThin4s(1.5);  
CUT bkThin5s(1.5);  
CUT bkThin6s(1.5);  
CUT bkThin7s(1.5);  
male = 2 - gender;
```

23

Input Binary Growth Modeling (Continued)

```
ANALYSIS: ESTIMATOR = MLR;  
INTEGRATION = 10;  
PROCESS = 4;  
MODEL: i s q | bkthin1f@0 bkthin1s@.5 bkthin2s@1.5  
bkthin3s@2.5 bkthin4s@3.5 bkthin5s@4.5 bkthin6s@5.5  
bkthin7s@6.5;  
i-q ON male;  
OUTPUT: TECH1 TECH8 STANDARDIZED;
```

10 integration points per dimension: 1000 points, 5 seconds
15 integration points per dimension: 3375 points, 1 minute 50
seconds

24

Output Excerpts Binary Growth (No Clustering)

Tests of Model Fit

Loglikelihood

H0 Value	-3546.377
H0 Scaling Correction Factor for MLR	1.0005

Information Criteria

Number of Free Parameters	12
Akaike (AIC)	7116.754
Bayesian (BIC)	7177.572
Sample-Size Adjusted BIC	7139.455

$$(n^* = (n + 2) / 24)$$

25

Output Excerpts Binary Growth (No Clustering) (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
i				
bkthin1f	1.000	0.000	999.000	999.000
bkthin1s	1.000	0.000	999.000	999.000
bkthin2s	1.000	0.000	999.000	999.000
bkthin3s	1.000	0.000	999.000	999.000
bkthin4s	1.000	0.000	999.000	999.000
bkthin5s	1.000	0.000	999.000	999.000
bkthin6s	1.000	0.000	999.000	999.000
bkthin7s	1.000	0.000	999.000	999.000

26

**Output Excerpts Binary Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
s				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.500	0.000	999.000	999.000
bkthin2s	1.500	0.000	999.000	999.000
bkthin3s	2.500	0.000	999.000	999.000
bkthin4s	3.500	0.000	999.000	999.000
bkthin5s	4.500	0.000	999.000	999.000
bkthin6s	5.500	0.000	999.000	999.000
bkthin7s	6.500	0.000	999.000	999.000

27

**Output Excerpts Binary Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
q				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.250	0.000	999.000	999.000
bkthin2s	2.250	0.000	999.000	999.000
bkthin3s	6.250	0.000	999.000	999.000
bkthin4s	12.250	0.000	999.000	999.000
bkthin5s	20.250	0.000	999.000	999.000
bkthin6s	30.250	0.000	999.000	999.000
bkthin7s	42.250	0.000	999.000	999.000

28

Output Excerpts Binary Growth (No Clustering) (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
i ON				
male	1.109	0.186	5.960	0.000
s ON				
male	-0.137	0.121	-1.132	0.258
q ON				
male	0.028	0.019	1.494	0.135
s WITH				
i	-1.298	0.314	-4.137	0.000
q WITH				
i	0.138	0.040	3.435	0.001
s	-0.059	0.026	-2.260	0.024

29

Output Excerpts Binary Growth (No Clustering) (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Intercepts				
I	0.000	0.000	999.000	999.000
S	0.204	0.102	1.998	0.046
Q	-0.045	0.017	-2.689	0.007
Thresholds				
bkthin1f\$1	1.839	0.149	12.324	0.000
bkthin1s\$1	1.839	0.149	12.324	0.000
bkthin2s\$1	1.839	0.149	12.324	0.000
bkthin3s\$1	1.839	0.149	12.324	0.000
bkthin4s\$1	1.839	0.149	12.324	0.000

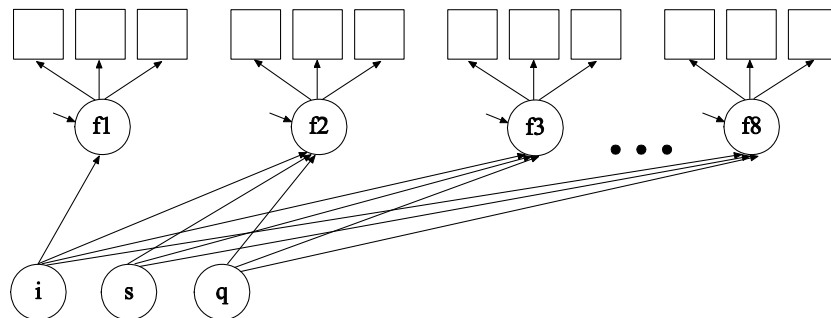
30

Output Excerpts Binary Growth (No Clustering) (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
bkthin5s\$1	1.839	0.149	12.324	0.000
bkthin6s\$1	1.839	0.149	12.324	0.000
bkthin7s\$1	1.839	0.149	12.324	0.000
Residual Variances				
i	3.803	0.625	6.083	0.000
s	0.545	0.190	2.868	0.004
q	0.007	0.004	1.781	0.075

31

Multiple Indicator Growth (No Clustering)



32

Input Excerpts Multiple Indicator Growth (No Clustering)

```
USEVAR = bkThin1f bkThin1s bkThin2s bkThin3s bkThin4s  
bkThin5s bkThin6s bkThin7s harm01f harm01s harm02s  
harm03s harm04s harm05s harm06s harm07s takeP1f takeP1s  
takeP2s takeP3s takeP4s takeP5s takeP6s takeP7s male;  
CATEGORICAL = bkThin1f - takeP7s;  
MISSING = ALL (999);  
  
DEFINE: CUT bkThin1f(1.5);  
CUT bkThin1s(1.5);  
CUT bkThin2s(1.5);  
CUT bkThin3s(1.5);  
CUT bkThin4s(1.5);  
CUT bkThin5s(1.5);  
CUT bkThin6s(1.5);  
CUT bkThin7s(1.5);  
  
CUT harm01f (1.5);  
CUT harm01s (1.5);  
CUT harm02s (1.5);  
CUT harm03s (1.5);
```

33

Input Excerpts Multiple Indicator Growth (No Clustering) (Continued)

```
CUT harm04s (1.5);  
CUT harm05s (1.5);  
CUT harm06s (1.5);  
CUT harm07s (1.5);  
  
CUT takeP1f(1.5);  
CUT takeP1s(1.5);  
CUT takeP2s(1.5);  
CUT takeP3s(1.5);  
CUT takeP4s(1.5);  
CUT takeP5s(1.5);  
CUT takeP6s(1.5);  
CUT takeP7s(1.5);  
  
male = 2 - gender;
```

34

Input Excerpts Multiple Indicator Growth (No Clustering)

```
ANALYSIS:  PROCESS = 4;
            ESTIMATOR = WLSM;
            PARAMETERIZATION = THETA;

MODEL:     f1 BY bkthin1f
            harmo1f (1)
            takep1f (2);
            f2 BY bkthin1s
            harmo1s (1)
            takep1s (2);
            f3 BY bkthin2s
            harmo2s (1)
            takep2s (2);
```

35

Input Excerpts Multiple Indicator Growth (No Clustering) (Continued)

```
            f4 BY bkthin3s
            harmo3s (1)
            takep3s (2);
            f5 BY bkthin4s
            harmo4s (1)
            takep4s (2);
            f6 BY bkthin5s
            harmo5s (1)
            takep5s (2);
            f7 BY bkthin6s
            harmo6s (1)
            takep6s (2);
```

36

Input Excerpts Multiple Indicator Growth (No Clustering) (Continued)

```
f8 BY bkthin7s
harmo7s (1)
takep7s (2);

[bkthin1f$1-bkthin7s$1] (11);
[harmo1f$1-harmo7s$1] (12);
[takep1f$1-takep7s$1] (13);

i s q | f1@0 f2@.5 f3@1.5 f4@2.5 f5@3.5 f6@4.5
f7@5.5 f8@6.5;
i-q ON male;
```

OUTPUT: TECH1 TECH8 STANDARDIZED;

37

Output Excerpts Multiple Indicator Growth (No Clustering)

Test of Model Fit

Chi-Square Test of Model Fit

Value	524.261
Degrees of Freedom	300
P-Value	0.0000
Scaling Correction Factor for WLSM	0.757

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference tests. MLM, MLR and WLSM chi-square difference testing is described in the Mplus Technical Appendices at www.statmodel.com. See chi-square difference testing in the index of the Mplus Users' Guide.

38

Output Excerpts Multiple Indicator Growth (No Clustering) (Continued)

Chi-Square Test of Model Fit for the Baseline Model

Value	37306.671
Degrees of Freedom	300
P-Value	0.0000

CFI/TLI

CFI	0.994
TLI	0.994

Number of Free Parameters 24

RMSEA (Root Mean Square Error of Approximation)

Estimate	0.025
----------	-------

39

Output Excerpts Multiple Indicator Growth (No Clustering) (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
i				
f1	1.000	0.000	999.000	999.000
f2	1.000	0.000	999.000	999.000
f3	1.000	0.000	999.000	999.000
f4	1.000	0.000	999.000	999.000
f5	1.000	0.000	999.000	999.000
f6	1.000	0.000	999.000	999.000
f7	1.000	0.000	999.000	999.000
f8	1.000	0.000	999.000	999.000

40

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
s				
f1	0.000	0.000	999.000	999.000
f2	0.500	0.000	999.000	999.000
f3	1.500	0.000	999.000	999.000
f4	2.500	0.000	999.000	999.000
f5	3.500	0.000	999.000	999.000
f6	4.500	0.000	999.000	999.000
f7	5.500	0.000	999.000	999.000
f8	6.500	0.000	999.000	999.000

41

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
α				
f1	0.000	0.000	999.000	999.000
f2	0.250	0.000	999.000	999.000
f3	2.250	0.000	999.000	999.000
f4	6.250	0.000	999.000	999.000
f5	12.250	0.000	999.000	999.000
f6	20.250	0.000	999.000	999.000
f7	30.250	0.000	999.000	999.000
f8	42.250	0.000	999.000	999.000

42

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f1 BY				
bkthin1f	1.000	0.000	999.000	999.000
harmol1f	1.239	0.112	11.089	0.000
takep1f	1.045	0.078	13.432	0.000
f2 BY				
bkthin1s	1.000	0.000	999.000	999.000
harmol1s	1.239	0.112	11.089	0.000
takep1s	1.045	0.078	13.432	0.000
f3 BY				
bkthin2s	1.000	0.000	999.000	999.000
harmo2s	1.239	0.112	11.089	0.000

43

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
takep2s	1.045	0.078	13.432	0.000
f4 BY				
bkthin3s	1.000	0.000	999.000	999.000
harmo3s	1.239	0.112	11.089	0.000
takep3s	1.045	0.078	13.432	0.000
f5 BY				
bkthin4s	1.000	0.000	999.000	999.000
harmo4s	1.239	0.112	11.089	0.000
takep4s	1.045	0.078	13.432	0.000
f6 BY				
bkthin5s	1.000	0.000	999.000	999.000

44

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
harmo5s	1.239	0.112	11.089	0.000
takep5s	1.045	0.078	13.432	0.000
f7 BY				
bkthin6s	1.000	0.000	999.000	999.000
harmo6s	1.239	0.112	11.089	0.000
takep6s	1.045	0.078	13.432	0.000
f8 BY				
bkthin7s	1.000	0.000	999.000	999.000
harmo7s	1.239	0.112	11.089	0.000
takep7s	1.045	0.078	13.432	0.000

45

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
i ON				
male	1.122	0.169	6.634	0.000
s ON				
male	0.046	0.108	0.424	0.671
q ON				
male	-0.006	0.016	-0.350	0.727
s WITH				
i	-1.225	0.227	-5.405	0.000
q WITH				
i	0.113	0.028	4.098	0.000
s	-0.086	0.019	-4.493	0.000

46

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Intercepts				
f1	0.000	0.000	999.000	999.000
f2	0.000	0.000	999.000	999.000
f3	0.000	0.000	999.000	999.000
f4	0.000	0.000	999.000	999.000
f5	0.000	0.000	999.000	999.000
f6	0.000	0.000	999.000	999.000
f7	0.000	0.000	999.000	999.000
f8	0.000	0.000	999.000	999.000
i	0.000	0.000	999.000	999.000
s	-0.071	0.090	-0.789	0.430
q	0.005	0.014	0.362	0.717

47

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Thresholds				
bthin1f\$1	1.944	0.150	12.995	0.000
bthin1s\$1	1.944	0.150	12.995	0.000
bthin2s\$1	1.944	0.150	12.995	0.000
bthin3s\$1	1.944	0.150	12.995	0.000
bthin4s\$1	1.944	0.150	12.995	0.000
bthin5s\$1	1.944	0.150	12.995	0.000
bthin6s\$1	1.944	0.150	12.995	0.000
bthin7s\$1	1.944	0.150	12.995	0.000
harmof\$1	1.477	0.175	8.437	0.000
harmols\$1	1.477	0.175	8.437	0.000
harmo2s\$1	1.477	0.175	8.437	0.000

48

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
harmo3s\$1	1.477	0.175	8.437	0.000
harmo4s\$1	1.477	0.175	8.437	0.000
harmo5s\$1	1.477	0.175	8.437	0.000
harmo6f\$1	1.477	0.175	8.437	0.000
harmo7s\$1	1.477	0.175	8.437	0.000
takep1f\$1	1.288	0.142	9.075	0.000
takep1s\$1	1.288	0.142	9.075	0.000
takep2s\$1	1.288	0.142	9.075	0.000
takep3s\$1	1.288	0.142	9.075	0.000
takep4s\$1	1.288	0.142	9.075	0.000
takep5s\$1	1.288	0.142	9.075	0.000

49

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
takep6s\$1	1.288	0.142	9.075	0.000
takep7s\$1	1.288	0.142	9.075	0.000
Residual variances				
f1	0.564	0.329	1.714	0.087
f2	1.648	0.396	4.161	0.000
f3	7.274	1.488	4.888	0.000
f4	2.278	0.510	4.462	0.000
f5	1.911	0.414	4.615	0.000

50

**Output Excerpts Multiple Indicator Growth
(No Clustering) (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f6	2.281	0.463	4.926	0.000
f7	2.462	0.526	4.682	0.000
f8	1.536	0.447	3.437	0.001
i	4.234	0.556	7.621	0.000
s	0.749	0.142	5.291	0.000
q	0.011	0.003	3.854	0.000

51

Multivariate Approach To Multilevel Modeling

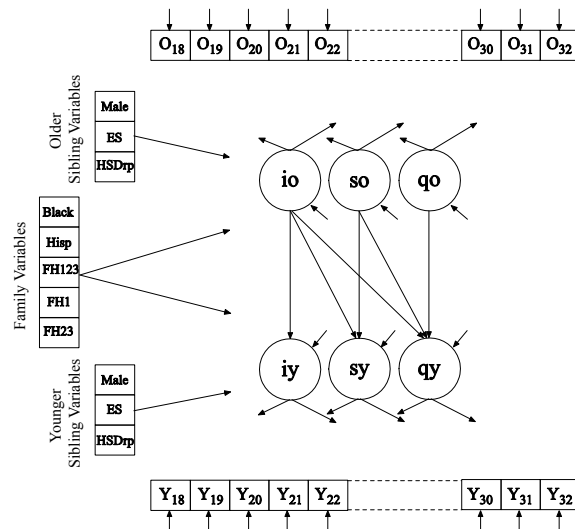
52

Multivariate Modeling Of Family Members

- Multilevel modeling: clusters independent, model for between- and within-cluster variation, members of a cluster statistically equivalent
- Multivariate approach: clusters independent, model for all variables for each cluster member, different parameters for different cluster members.
 - Used in latent variable growth modeling where the cluster members are the repeated measures over time
 - Allows for different cluster sizes by missing data techniques
 - More flexible than the multilevel approach, but computationally convenient only for applications with small cluster sizes (e.g. twins, spouses)

53

Figure 1. A Longitudinal Growth Model of Heavy Drinking for Two-Sibling Families



Source: Khoo, S.T. & Muthen, B. (2000). Longitudinal data on families: Growth modeling alternatives. Multivariate Applications in Substance Use Research, J. Rose, L. Chassin, C. Presson & J. Sherman (eds.), Hillsdale, N.J.: Erlbaum, pp. 43-78. 54

Three-Level Modeling As Single-Level Analysis

- The sibling growth model has 3 levels:
 - Time
 - Individual
 - Family
- Analyzed as doubly multivariate:
 - Repeated measures in wide, multivariate form
 - Siblings in wide, multivariate form

It is possible to do four-level by TYPE = TWOLEVEL, for instance families within geographical segments

55

Input For Multivariate Modeling Of Family Data

```
TITLE:      Multivariate modeling of family data
            one observation per family

DATA:      FILE IS multi.dat;

VARIABLE:  NAMES ARE o18-o32 y18-y32 omale oes ohdrop ymale yes
            yhsdrop black hisp fh123 fh1 fh23;

MODEL:    io so qo | o18@0 o19@1 o20@2 o21@3 o22@4
            o23@5 o24@6 o25@7 o26@8 o27@9 o28@10 o29@11 o30@12
            o31@13 o32@14;
            iy sy qy | y18@0 y19@1 y20@2 y21@3 y22@4 y23@5 y24@6
            y25@7 y26@8 y27@9 y28@10 y29@11 y30@12
            y31@13 y32@14;
            io ON omale oes ohdrop black hisp fh123 fh1 fh23;
            iy ON ymale yes yhsdrop black hisp fh123 fh1 fh23;
            io ON io;
            sy ON io so;
            qy ON io so qo;
```

56

Multilevel Growth Modeling (3-Level Analysis)

57

Three-Level Analysis In Multilevel Terms

Time point t , individual i , cluster j .

- y_{tij} : individual-level, outcome variable
- a_{1tij} : individual-level, time-related variable (age, grade)
- a_{2tij} : individual-level, time-varying covariate
- x_{ij} : individual-level, time-invariant covariate
- w_j : cluster-level covariate

Three-level analysis (Mplus considers Within and Between)

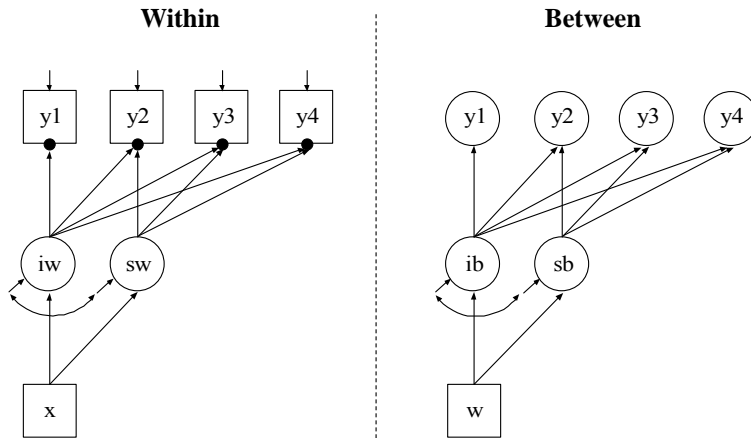
$$\text{Level 1 (Within)} : y_{tij} = \pi_{0ij} + \pi_{1ij} a_{1tij} + \pi_{2ij} a_{2tij} + e_{tij}, \quad (1)$$

$$\text{Level 2 (Within)} : \begin{cases} \pi_{0ij} = \beta_{00j} + \beta_{01j} x_{ij} + r_{0ij} \rightarrow iw \\ \pi_{1ij} = \beta_{10j} + \beta_{11j} x_{ij} + r_{1ij} \rightarrow sw \\ \pi_{2ij} = \beta_{20j} + \beta_{21j} x_{ij} + r_{2ij}. \end{cases} \quad (2)$$

$$\text{Level 3 (Between)} : \begin{cases} \beta_{00j} = \gamma_{000} + \gamma_{001} w_j + u_{00j} \rightarrow ib \\ \beta_{10j} = \gamma_{100} + \gamma_{101} w_j + u_{10j} \rightarrow sb \\ \beta_{20j} = \gamma_{200} + \gamma_{201} w_j + u_{20j}, \\ \beta_{01j} = \gamma_{010} + \gamma_{011} w_j + u_{01j}, \\ \beta_{11j} = \gamma_{110} + \gamma_{111} w_j + u_{11j}, \\ \beta_{21j} = \gamma_{210} + \gamma_{211} w_j + u_{21j}. \end{cases} \quad (3)$$

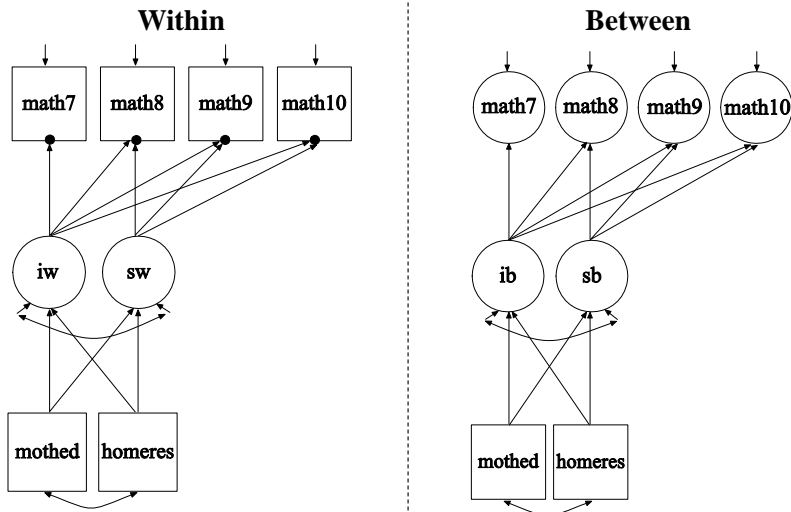
58

Two-Level Growth Modeling (Three-Level Analysis)



59

LSAY Two-Level Growth Model



60

Input For LSAY Two-Level Growth Model With Free Time Scores And Covariates

```
TITLE:      LSAY two-level growth model with free time scores
            and covariates

DATA:      FILE IS lsay98.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;
            USEOBS = (gender EQ 1 AND cohort EQ 2);
            MISSING = ALL (999);
            USEVAR = math7-math10 mothed homeres;
            CLUSTER = school;

ANALYSIS:  TYPE = TWOLEVEL;
            ESTIMATOR = MUML;
```

61

Input For LSAY Two-Level Growth Model With Free Time Scores And Covariates (Continued)

```
MODEL:     %WITHIN%
            iw sw | math7@0 math8@1
            math9*2 (1)
            math10*3 (2);
            iw sw ON mothed homeres;

            %BETWEEN%
            ib sb | math7@0 math8@1
            math9*2 (1)
            math10*3 (2);
            ib sb ON mothed homeres;

OUTPUT     SAMPSTAT STANDARDIZED RESIDUAL;
```

62

Output Excerpts LSAY Two-Level Growth Model With Free Time Scores And Covariates

Summary of Data

Number of clusters 50

Size (s) Cluster ID with Size s

1	114	
2	136	
6	132	304

34	104
39	309
40	302

Average cluster size 18.627

Estimated Intraclass Correlations for the Y Variables

Variable	Intraclass Correlation	Variable	Intraclass Correlation	Variable	Intraclass Correlation
MATH7	0.199	MATH8	0.149	MATH9	0.168
MATH10	0.165				63

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

Tests Of Model Fit

Chi-square Test of Model Fit

Value	24.058*
Degrees of Freedom	14
P-Value	0.0451

CFI / TLI

CFI	0.997
TLI	0.995

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.028
----------	-------

SRMR (Standardized Root Mean Square Residual)

Value for Between	0.048
Value for Within	0.007

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

Model Results

Within Level	Estimates	S.E.	Est./S.E.	Std	StdYX
SW					
MATH8	1.000	0.000	0.000	1.073	0.128
MATH9	2.487	0.163	15.220	2.670	0.288
MATH10	3.589	0.223	16.076	3.853	0.368
IW ON					
MOTHED	1.780	0.232	7.665	0.246	0.226
HOMERES	0.892	0.221	4.031	0.124	0.173
SW ON					
MOTHED	0.053	0.063	0.836	0.049	0.045
HOMERES	0.135	0.044	3.047	0.125	0.176

65

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

	Estimates	S.E.	Est./S.E.	Std	StdYX
SW WITH					
IW	2.112	0.522	4.044	0.273	0.273
HOMERES WITH					
MOTHED	0.261	0.039	6.709	0.261	0.203
Residual Variances					
MATH7	12.748	1.434	8.888	12.748	0.197
MATH8	12.298	0.893	13.771	12.298	0.174
MATH9	14.237	1.132	12.578	14.237	0.166
MATH10	24.829	2.230	11.133	24.829	0.226
IW	47.060	3.069	15.333	0.903	0.903
SW	1.110	0.286	3.879	0.964	0.964
Variances					
MOTHED	0.841	0.049	17.217	0.841	1.000
HOMERES	1.970	0.069	28.643	1.970	1.000

66

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

	Estimates	S.E.	Est./S.E.	Std	StdYX
Between Level					
SB					
MATH8	1.000	0.000	0.000	0.196	0.052
MATH9	2.487	0.163	15.220	0.488	0.119
MATH10	3.589	0.223	16.076	0.704	0.115
IB ON					
MOTHEd	-1.225	2.587	-0.474	-0.362	-0.107
HOMERES	7.160	1.847	3.876	2.117	1.011
SB ON					
MOTHEd	0.995	0.647	1.538	5.073	1.493
HOMERES	0.017	0.373	0.045	0.086	0.041
SB WITH					
IB	0.382	0.248	1.538	0.575	0.575

67

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

HOMERES WITH					
MOTHEd	0.103	0.019	5.488	0.103	0.733
Residual Variances					
MATH7	2.059	0.552	3.732	2.059	0.153
MATH8	0.544	0.268	2.033	0.544	0.039
MATH9	0.105	0.213	0.493	0.105	0.006
MATH10	1.395	0.504	2.767	1.395	0.067
IB	1.428	1.690	0.845	0.125	0.125
SB	-0.051	0.071	-0.713	-1.321	-1.321
Variances					
MOTHEd	0.087	0.023	3.801	0.087	1.000
HOMERES	0.228	0.056	4.066	0.228	1.000
Means					
MOTHEd	2.307	0.043	53.277	2.307	7.838
HOMERES	3.108	0.062	50.375	3.108	6.509
Intercepts					
IB	33.510	2.678	12.512	9.909	9.909
SB	0.163	0.776	0.210	0.830	0.830

68

**Output Excerpts LSAY Two-Level Growth Model
With Free Time Scores And Covariates (Continued)**

R-Square

Within Level

Observed Variable	R-Square
----------------------	----------

MATH7	0.803
MATH8	0.826
MATH9	0.834
MATH10	0.774

Latent Variable	R-Square
--------------------	----------

IW	0.097
SW	0.036

69

**Output Excerpts LSAY Two-Level Growth Model
With Free Time Scores And Covariates (Continued)**

R-Square

Between Level

Observed Variable	R-Square
----------------------	----------

MATH7	0.847
MATH8	0.961
MATH9	0.994
MATH10	0.933

Latent Variable	R-Square
--------------------	----------

IB	0.875
SB	Undefined 0.23207E+01

70

Further Readings On Three-Level Growth Analysis

- Muthén, B. (1997). Latent variable modeling with longitudinal and multilevel data. In A. Raftery (ed), Sociological Methodology (pp. 453-480). Boston: Blackwell Publishers. (#73)
- Muthén, B. & Asparouhov, T. (2011). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. In J. Hox & J.K. Roberts (eds), The Handbook of Advanced Multilevel Analysis, pp 15-40. New York: Taylor and Francis.
- Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods. Second edition. Newbury Park, CA: Sage Publications.
- Snijders, T. & Bosker, R. (1999). Multilevel analysis. An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.

71

Multilevel Growth Modeling Of Binary Outcomes (3-Level Analysis)

72

Multilevel Growth Model With Binary Outcomes

Time point t , individual i , cluster j .

Logit-linear growth.

Level 1
(Within Cluster): $\log \left[\frac{P(u_{it} = 1 | \eta_{0ij}, \eta_{1ij}, a_{it})}{P(u_{it} = 0 | \eta_{0ij}, \eta_{1ij}, a_{it})} \right] = \eta_{0ij} + \eta_{1ij} \cdot (a_{it} - c) \quad (1)$

Level 2
(Within Cluster): $\begin{cases} \eta_{0ij} = \beta_{00j} + \beta_{01} x_{ij} + r_{0ij}, & \rightarrow \eta_{0w} \\ \eta_{1ij} = \beta_{10j} + \beta_{11} x_{ij} + r_{1ij}, & \rightarrow \eta_{1w} \end{cases} \quad (2)$

Level 3
(Between Cluster): $\begin{cases} \beta_{00j} = \gamma_{000} + \gamma_{001} w_j + r_{00j}, & \rightarrow \eta_{0b} \\ \beta_{10j} = \gamma_{100} + \gamma_{101} w_j + r_{10j}, & \rightarrow \eta_{1b} \end{cases} \quad (3)$

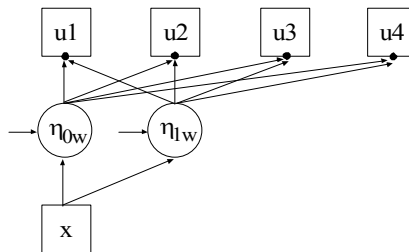
Fitzmaurice, Laird, & Ware (2004). Applied longitudinal analysis. Wiley & Sons.

Fitzmaurice, Davidian, Verbeke, & Molenberghs (2009). Longitudinal Data Analysis. Chapman & Hall/CRC Press.

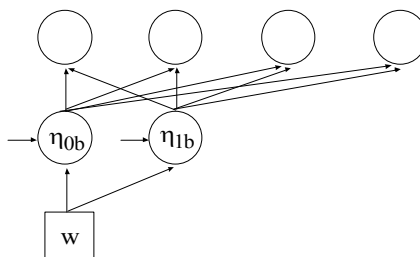
73

Growth Model With Binary Outcomes

Within Cluster



Between Cluster



74

Observed Data Likelihood

Individual i in cluster j

$$\prod_j \int \phi_j(\eta_{bj}) \prod_i \left(\int f_{ij}(u_{ij} | \eta_{bj}, \eta_{wij}) \phi_{ij}(\eta_{wij}) d\eta_{wij} \right) d\eta_{bj}$$

- Maximum likelihood estimation
- Numerical integration
- The logit-linear binary growth model results in 4 dimensions of integration

75

Input Excerpts Two-Level Binary Growth

```
USEVAR = bkthin1f bkthin1s bkthin2s bkthin3s bkthin4s
bkthin5s bkthin6s bkthin7s male;
CATEGORICAL = bkthin1f - bkthin7s;
MISSING = ALL (999);
CLUSTER = classrm;
WITHIN = male;

DEFINE: CUT bkThin1f(1.5);
CUT bkThin1s(1.5);
CUT bkThin2s(1.5);
CUT bkThin3s(1.5);
CUT bkThin4s(1.5);
CUT bkThin5s(1.5);
CUT bkThin6s(1.5);
CUT bkThin7s(1.5);
male = 2 - gender;

ANALYSIS: TYPE = TWOLEVEL;
PROCESS = 4;
ESTIMATOR = WLSM;
```

76

Input Excerpts Two-Level Binary Growth

```
MODEL:      %WITHIN%

            iw sw qw | bkthin1f@0 bkthin1s@.5 bkthin2s@1.5
            bkthin3s@2.5 bkthin4s@3.5 bkthin5s@4.5 bkthin6s@5.5
            bkthin7s@6.5;

            iw-qw on male;

            %BETWEEN%

            ib sb qb |bkthin1f@0 bkthin1s@.5 bkthin2s@1.5
            bkthin3s@2.5 bkthin4s@3.5 bkthin5s@4.5 bkthin6s@5.5
            bkthin7s@6.5;

            bkthin1f-bkthin7s@0;

OUTPUT:     TECH1 TECH8 STANDARDIZED;
```

77

Output Excerpts Two-Level Binary Growth

```
Number of clusters                41
Average cluster size              28.634
```

Estimated Intraclass Correlations for the Y Variables

Intraclass		Intraclass		Intraclass	
Variable	Correlation	Variable	Correlation	Variable	Correlation
bkthin1f	0.470	bkthin1s	0.484	bkthin2s	0.385
bkthin3s	0.089	bkthin4s	0.052	bkthin5s	0.094
bkthin6s	0.137	bkthin7s	0.104		

78

Output Excerpts Two-Level Binary Growth (Continued)

Tests of Model Fit

Chi-Square Test of Model Fit

Value	101.742*
Degrees of Freedom	62
P-Value	0.0011

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference tests. MLM, MLR and WLSM chi-square difference testing is described in the Mplus Technical Appendices at www.statmodel.com. See chi-square difference testing in the index of the Mplus Users' Guide.

79

Output Excerpts Two-Level Binary Growth (Continued)

Chi-Square Test of Model Fit for the Baseline Model

Value	924.374
Degrees of Freedom	64
P-Value	0.0000

CFI/TLI

CFI	0.954
TLI	0.952

Number of Free Parameters 18

RMSEA (Root Mean Square Error Of Approximation)

Estimate 0.023

SRMR (Standardized Root Mean Square Residual)

Value for Within	0.056
Value for Between	0.513

80

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
iw				
bkthin1f	1.000	0.000	999.000	999.000
bkthin1s	1.000	0.000	999.000	999.000
bkthin2s	1.000	0.000	999.000	999.000
bkthin3s	1.000	0.000	999.000	999.000
bkthin4s	1.000	0.000	999.000	999.000
bkthin5s	1.000	0.000	999.000	999.000
bkthin6s	1.000	0.000	999.000	999.000
bkthin7s	1.000	0.000	999.000	999.000

81

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
sw				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.500	0.000	999.000	999.000
bkthin2s	1.500	0.000	999.000	999.000
bkthin3s	2.500	0.000	999.000	999.000
bkthin4s	3.500	0.000	999.000	999.000
bkthin5s	4.500	0.000	999.000	999.000
bkthin6s	5.500	0.000	999.000	999.000
bkthin7s	6.500	0.000	999.000	999.000

82

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
qw				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.250	0.000	999.000	999.000
bkthin2s	2.250	0.000	999.000	999.000
bkthin3s	6.250	0.000	999.000	999.000
bkthin4s	12.250	0.000	999.000	999.000
bkthin5s	20.250	0.000	999.000	999.000
bkthin6s	30.250	0.000	999.000	999.000
bkthin7s	42.250	0.000	999.000	999.000

83

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
iw ON				
male	0.980	0.131	7.496	0.000
sw ON				
male	-0.177	0.087	-2.032	0.042
qw ON				
male	0.023	0.013	1.707	0.088
sw WITH				
iw	-0.290	0.121	-2.388	0.017
qw WITH				
iw	0.022	0.016	1.434	0.151
sw	-0.010	0.010	-0.987	0.324

84

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Residual Variances				
iw	1.315	0.284	4.637	0.000
sw	0.102	0.071	1.430	0.153
qw	0.001	0.001	0.683	0.495
Between Level				
ib				
bkthin1f	1.000	0.000	999.000	999.000
bkthin1s	1.000	0.000	999.000	999.000
bkthin2s	1.000	0.000	999.000	999.000
bkthin3s	1.000	0.000	999.000	999.000

85

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
bkthin4s	1.000	0.000	999.000	999.000
bkthin5s	1.000	0.000	999.000	999.000
bkthin6s	1.000	0.000	999.000	999.000
bkthin7s	1.000	0.000	999.000	999.000
sb				
bkthin1f	1.000	0.000	999.000	999.000
bkthin1s	0.500	0.000	999.000	999.000
bkthin2s	1.500	0.000	999.000	999.000
bkthin3s	2.500	0.000	999.000	999.000
bkthin4s	3.500	0.000	999.000	999.000
bkthin5s	4.500	0.000	999.000	999.000

86

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
bkthin6s	5.500	0.000	999.000	999.000
bkthin7s	6.500	0.000	999.000	999.000
qb				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.250	0.000	999.000	999.000
bkthin2s	2.250	0.000	999.000	999.000
bkthin3s	6.250	0.000	999.000	999.000
bkthin4s	12.250	0.000	999.000	999.000
bkthin5s	20.250	0.000	999.000	999.000
bkthin6s	30.250	0.000	999.000	999.000
bkthin7s	42.250	0.000	999.000	999.000

87

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
sw WITH				
ib	-0.395	0.122	-3.246	0.001
qb WITH				
ib	0.043	0.014	3.069	0.002
sb	-0.018	0.007	-2.451	0.014
Means				
ib	0.000	0.000	999.000	999.000
sb	0.219	0.133	1.648	0.099
qb	-0.028	0.018	-1.614	0.107

88

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Thresholds				
bkthin1f\$1	1.592	0.253	6.304	0.000
bkthin1s\$1	1.592	0.253	6.304	0.000
bkthin2s\$1	1.592	0.253	6.304	0.000
bkthin3s\$1	1.592	0.253	6.304	0.000
bkthin4s\$1	1.592	0.253	6.304	0.000
bkthin5s\$1	1.592	0.253	6.304	0.000
bkthin6s\$1	1.592	0.253	6.304	0.000
bkthin7s\$1	1.592	0.253	6.304	0.000

89

Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Variiances				
ib	0.935	0.285	3.278	0.001
sb	0.165	0.058	2.842	0.004
qb	0.002	0.001	2.071	0.038
Residual Variiances				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.000	0.000	999.000	999.000
bkthin2s	0.000	0.000	999.000	999.000

90

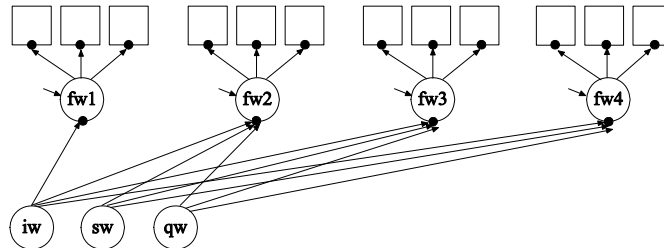
Output Excerpts Two-Level Binary Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
bkthin3s	0.000	0.000	999.000	999.000
bkthin4s	0.000	0.000	999.000	999.000
bkthin5s	0.000	0.000	999.000	999.000
bkthin6s	0.000	0.000	999.000	999.000
bkthin7s	0.000	0.000	999.000	999.000

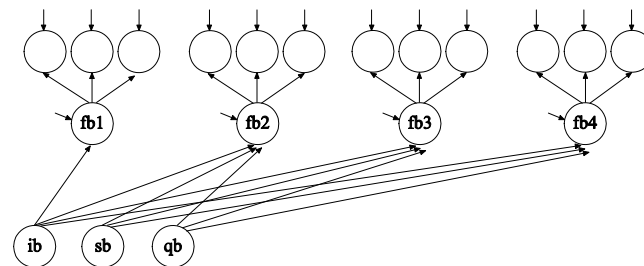
91

Two-Level Multiple Indicator Growth

Within (individual variation)



Between (classroom variation)



92

Multilevel Multiple Indicator Growth In Other Software

Treated as 4-level analysis:

- Level 1: Indicators
- Level 2: Time points
- Level 3: Individuals
- Level 4: Clusters

This approach is

- Cumbersome requiring dummy variables to indicate indicators
- Limiting in that only intercepts/thresholds and not loadings are allowed to vary across indicators (Rasch model for categorical outcomes).

93

Multilevel Multiple Indicator Growth In Other Software (Continued)

In contrast, Mplus treats such analysis as two-level modeling with

- Within: Indicators, time points, individuals (doubly wide format: indicators x time points)
- Between: Clusters

and allowing intercepts/thresholds and loadings to vary across indicators and, if needed, across time.

94

Input Excerpts Two-Level Multiple Indicator Growth

```
ANALYSIS:  TYPE = TWOLEVEL;  
           PROCESS = 4;  
           ESTIMATOR = WLSM;  
  
MODEL:    %WITHIN%  
          f1 BY bkthin1f  
          harmolf (1)  
          takep1f (2);  
          f2 BY bkthin1s  
          harmo1s (1)  
          takep1s (2);  
          f3 BY bkthin2s  
          harmo2s (1)  
          takep2s (2);
```

95

Input Excerpts Two-Level Multiple Indicator Growth (Continued)

```
          f4 BY bkthin3s  
          harmo3s (1)  
          takep3s (2);  
          f5 BY bkthin4s  
          harmo4s (1)  
          takep4s (2);  
          f6 BY bkthin5s  
          harmo5s (1)  
          takep5s (2);  
          f7 BY bkthin6s  
          harmo6s (1)  
          takep6s (2);
```

96

Input Excerpts Two-Level Multiple Indicator Growth (Continued)

```
f8 BY bkthin7s
harmo7s (1)
takep7s (2);

iw sw qw | f1@0 f2@.5 f3@1.5 f4@2.5 f5@3.5 f6@4.5
f7@5.5 f8@6.5;

iw-qw ON male;

%BETWEEN%

f1b BY bkthin1f
harmo1f (3)
takep1f (4);

f2b BY bkthin1s
harmo1s (3)
takep1s (4);
```

97

Input Excerpts Two-Level Multiple Indicator Growth (Continued)

```
f3b BY bkthin2s
harmo2s (3)
takep2s (4);

f4b BY bkthin3s
harmo3s (3)
takep3s (4);

f5b BY bkthin4s
harmo4s (3)
takep4s (4);

f6b BY bkthin5s
harmo5s (3)
takep5s (4);
```

98

Input Excerpts Two-Level Multiple Indicator Growth (Continued)

```
f7b BY bkthin6s
harmo6s (3)
takep6s (4);
f8b BY bkthin7s
harmo7s (3)
takep7s (4);
[bkthin1f$1-bkthin7s$1] (11);
[harmolf$1-harmo7s$1] (12);
[takep1f$1-takep7s$1] (13);
ib sb qb |f1b@0 f2b@.5 f3b@1.5 f4b@2.5 f5b@3.5 f6b@4.5
f7b@5.5 f8b@6.5;
bkthin1f-takep7s@0;
OUTPUT: TECH1 TECH8 STANDARDIZED;
```

99

Output Excerpts Two-Level Multiple Indicator Growth

Tests of Model Fit

Chi-Square Test of Model Fit

Value	767.730*
Degrees of Freedom	584
P-Value	0.0000

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference tests. MLM, MLR and WLSM chi-square difference testing is described in the Mplus Technical Appendices at www.statmodel.com. See chi-square difference testing in the index of the Mplus Users' Guide.

100

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Chi-Square Test of Model Fit for the Baseline Model

Value	28765.253
Degrees of Freedom	576
P-Value	0.0000
CFI/TLI	
CFI	0.993
TLI	0.994
Number of Free Parameters	40
RMSEA (Root Mean Square Error of Approximation)	
Estimate	0.016
SRMR (Standardized Root Mean Square Residual)	
Value for Within	0.065
Value for Between	0.232

101

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

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

Within Level

iw				
f1	1.000	0.000	999.000	999.000
f2	1.000	0.000	999.000	999.000
f3	1.000	0.000	999.000	999.000
f4	1.000	0.000	999.000	999.000
f5	1.000	0.000	999.000	999.000
f6	1.000	0.000	999.000	999.000
f7	1.000	0.000	999.000	999.000
f8	1.000	0.000	999.000	999.000

102

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
sw				
f1	0.000	0.000	999.000	999.000
f2	0.500	0.000	999.000	999.000
f3	1.500	0.000	999.000	999.000
f4	2.500	0.000	999.000	999.000
f5	3.500	0.000	999.000	999.000
f6	4.500	0.000	999.000	999.000
f7	5.500	0.000	999.000	999.000
f8	6.500	0.000	999.000	999.000
qw				
f1	0.000	0.000	999.000	999.000

103

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f2	0.250	0.000	999.000	999.000
f3	2.250	0.000	999.000	999.000
f4	6.250	0.000	999.000	999.000
f5	12.250	0.000	999.000	999.000
f6	20.250	0.000	999.000	999.000
f7	30.250	0.000	999.000	999.000
f8	42.250	0.000	999.000	999.000
f1 BY				
bkthinlf	1.000	0.000	999.000	999.000
harmolf	1.143	0.088	12.932	0.000

104

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
harmo3s	1.143	0.088	12.932	0.000
takep3s	0.993	0.067	14.782	0.000
f5 BY				
bkthin4s	1.000	0.000	999.000	999.000
harmo4s	1.143	0.088	12.932	0.000
takep4s	0.993	0.067	14.782	0.000
f6 BY				
bkthin5s	1.000	0.000	999.000	999.000
harmo5s	1.143	0.088	12.932	0.000
takep5s	0.993	0.067	14.782	0.000

105

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f7 BY				
bkthin6s	1.000	0.000	999.000	999.000
harmo6s	1.143	0.088	12.932	0.000
takep6s	0.993	0.067	14.782	0.000
f8 BY				
bkthin7s	1.000	0.000	999.000	999.000
harmo7s	1.143	0.088	12.932	0.000
takep7s	0.993	0.067	14.782	0.000
iw ON				
male	1.449	0.175	8.258	0.000

106

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
sw ON				
male	-0.091	0.129	-0.705	0.481
qw ON				
male	0.011	0.021	0.553	0.580
sw WITH				
iw	-0.394	0.141	-2.793	0.005
qw WITH				
iw	0.017	0.020	0.835	0.404
sw	-0.031	0.016	-1.973	0.049
Residual Variances				
f1	1.093	0.259	4.221	0.000

107

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f2	0.739	0.236	3.131	0.002
f3	2.850	0.497	5.731	0.000
f4	1.985	0.561	3.541	0.000
f5	2.255	0.487	4.630	0.000
f6	2.581	0.317	8.150	0.000
f7	2.171	0.383	5.667	0.000
f8	2.332	0.557	4.188	0.000
iw	2.812	0.366	7.679	0.000
sw	0.277	0.110	2.510	0.012
qw	0.004	0.002	1.797	0.072

108

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Between Level				
ib				
f1b	1.000	0.000	999.000	999.000
f2b	1.000	0.000	999.000	999.000
f3b	1.000	0.000	999.000	999.000
f4b	1.000	0.000	999.000	999.000
f5b	1.000	0.000	999.000	999.000
f6b	1.000	0.000	999.000	999.000
f7b	1.000	0.000	999.000	999.000
f8b	1.000	0.000	999.000	999.000

109

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
sb				
f1b	0.000	0.000	999.000	999.000
f2b	0.500	0.000	999.000	999.000
f3b	1.500	0.000	999.000	999.000
f4b	2.500	0.000	999.000	999.000
f5b	3.500	0.000	999.000	999.000
f6b	4.500	0.000	999.000	999.000
f7b	5.500	0.000	999.000	999.000
f8b	6.500	0.000	999.000	999.000
qb				
f1b	0.000	0.000	999.000	999.000

110

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f2b	0.250	0.000	999.000	999.000
f3b	2.250	0.000	999.000	999.000
f4b	6.250	0.000	999.000	999.000
f5b	12.250	0.000	999.000	999.000
f6b	20.250	0.000	999.000	999.000
f7b	30.250	0.000	999.000	999.000
f8b	42.250	0.000	999.000	999.000
f1b BY				
bkthin1f	1.000	0.000	999.000	999.000
harmolf	1.099	0.145	7.603	0.000
takeplf	0.979	0.126	7.798	0.000

111

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f2b BY				
bkthin1s	1.000	0.000	999.000	999.000
harmols	1.099	0.145	7.603	0.000
takepls	0.979	0.126	7.798	0.000
f3b BY				
bkthin2s	1.000	0.000	999.000	999.000
harmo2s	1.099	0.145	7.603	0.000
takep2s	0.979	0.126	7.798	0.000
f4b BY				
bkthin3s	1.000	0.000	999.000	999.000
harmo3s	1.099	0.145	7.603	0.000

112

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
takep3s	0.979	0.126	7.798	0.000
f5b BY				
bkthin4s	1.000	0.000	999.000	999.000
harmo4s	1.099	0.145	7.603	0.000
takep4s	0.979	0.126	7.798	0.000
f6b BY				
bkthin5s	1.000	0.000	999.000	999.000
harmo5s	1.099	0.145	7.603	0.000
takep5s	0.979	0.126	7.798	0.000
f7b BY				
bkthin6s	1.000	0.000	999.000	999.000

113

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
harmo6s	1.099	0.145	7.603	0.000
takep6s	0.979	0.126	7.798	0.000
f8b BY				
bkthin7s	1.000	0.000	999.000	999.000
harmo7s	1.099	0.145	7.603	0.000
takep7s	0.979	0.126	7.798	0.000
sb WITH				
ib	-0.563	0.200	-2.813	0.005
qb WITH				
ib	0.066	0.023	2.927	0.003
sb	-0.031	0.012	-2.530	0.011

114

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Means				
f1	0.000	0.000	999.000	999.000
f2	0.000	0.000	999.000	999.000
f3	0.000	0.000	999.000	999.000
f4	0.000	0.000	999.000	999.000
f5	0.000	0.000	999.000	999.000
f6	0.000	0.000	999.000	999.000
f7	0.000	0.000	999.000	999.000
f8	0.000	0.000	999.000	999.000
ib	0.000	0.000	999.000	999.000
sb	0.063	0.170	0.371	0.711

115

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
qb	-0.014	0.023	-0.615	0.538
Intercepts				
f1b	0.000	0.000	999.000	999.000
f2b	0.000	0.000	999.000	999.000
f3b	0.000	0.000	999.000	999.000
f4b	0.000	0.000	999.000	999.000
f5b	0.000	0.000	999.000	999.000
f6b	0.000	0.000	999.000	999.000
f7b	0.000	0.000	999.000	999.000
f8b	0.000	0.000	999.000	999.000

116

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Thresholds				
bkthin1f\$1	2.364	0.357	6.625	0.000
bkthin1s\$1	2.364	0.357	6.625	0.000
bkthin2s\$1	2.364	0.357	6.625	0.000
bkthin3s\$1	2.364	0.357	6.625	0.000
bkthin4s\$1	2.364	0.357	6.625	0.000
bkthin5s\$1	2.364	0.357	6.625	0.000
bkthin6s\$1	2.364	0.357	6.625	0.000
bkthin7s\$1	2.364	0.357	6.625	0.000
harmo1f\$1	1.676	0.353	4.746	0.000
harmo1s\$1	1.676	0.353	4.746	0.000

117

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
harmo2s\$1	1.676	0.353	4.746	0.000
harmo3s\$1	1.676	0.353	4.746	0.000
harmo4s\$1	1.676	0.353	4.746	0.000
harmo5s\$1	1.676	0.353	4.746	0.000
harmo6s\$1	1.676	0.353	4.746	0.000
harmo7s\$1	1.676	0.353	4.746	0.000
takep1f\$1	1.612	0.288	5.595	0.000
takep1s\$1	1.612	0.288	5.595	0.000
takep2s\$1	1.612	0.288	5.595	0.000
takep3s\$1	1.612	0.288	5.595	0.000
takep4s\$1	1.612	0.288	5.595	0.000

118

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
takep5s\$1	1.612	0.288	5.595	0.000
takep6s\$1	1.612	0.288	5.595	0.000
takep7s\$1	1.612	0.288	5.595	0.000
Variances				
ib	1.328	0.515	2.580	0.010
sb	0.258	0.097	2.669	0.008
qb	0.004	0.002	2.298	0.022
Residual Variances				
bkthin1f	0.000	0.000	999.000	999.000
bkthin1s	0.000	0.000	999.000	999.000
bkthin2s	0.000	0.000	999.000	999.000

119

Output Excerpts Two-Level Multiple Indicator Growth (Continued)

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
bkthin3s	0.000	0.000	999.000	999.000
bkthin4s	0.000	0.000	999.000	999.000
bkthin5s	0.000	0.000	999.000	999.000
bkthin6s	0.000	0.000	999.000	999.000
bkthin7s	0.000	0.000	999.000	999.000
harmo1f	0.000	0.000	999.000	999.000
harmo1s	0.000	0.000	999.000	999.000
harmo2s	0.000	0.000	999.000	999.000
harmo3s	0.000	0.000	999.000	999.000
harmo4s	0.000	0.000	999.000	999.000
harmo5s	0.000	0.000	999.000	999.000

120

**Output Excerpts Two-Level
Multiple Indicator Growth (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
harmo6s	0.000	0.000	999.000	999.000
harmo7s	0.000	0.000	999.000	999.000
takep1f	0.000	0.000	999.000	999.000
takep1s	0.000	0.000	999.000	999.000
takep2s	0.000	0.000	999.000	999.000
takep3s	0.000	0.000	999.000	999.000
takep4s	0.000	0.000	999.000	999.000
takep5s	0.000	0.000	999.000	999.000
takep6s	0.000	0.000	999.000	999.000
takep7s	0.000	0.000	999.000	999.000
f1b	2.258	0.786	2.871	0.004

121

**Output Excerpts Two-Level
Multiple Indicator Growth (Continued)**

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
f2b	2.206	0.821	2.685	0.007
f3b	3.002	1.062	2.827	0.005
f4b	0.377	0.237	1.594	0.111
f5b	0.051	0.117	0.440	0.660
f6b	0.251	0.214	1.176	0.239
f7b	0.543	0.303	1.790	0.073
f8b	-0.468	0.298	-1.567	0.117

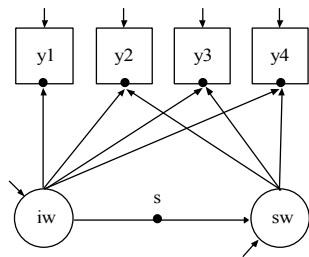
122

Special Multilevel Growth Models

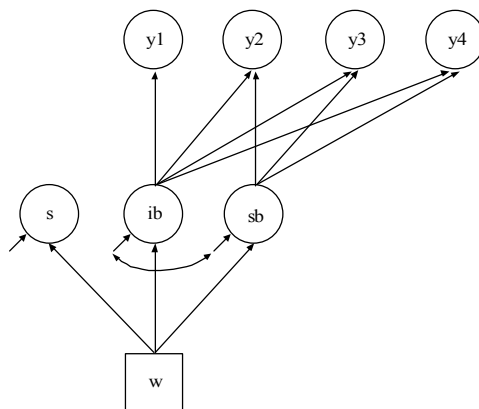
123

Multilevel Modeling With A Random Slope For Latent Variables

Student (Within)

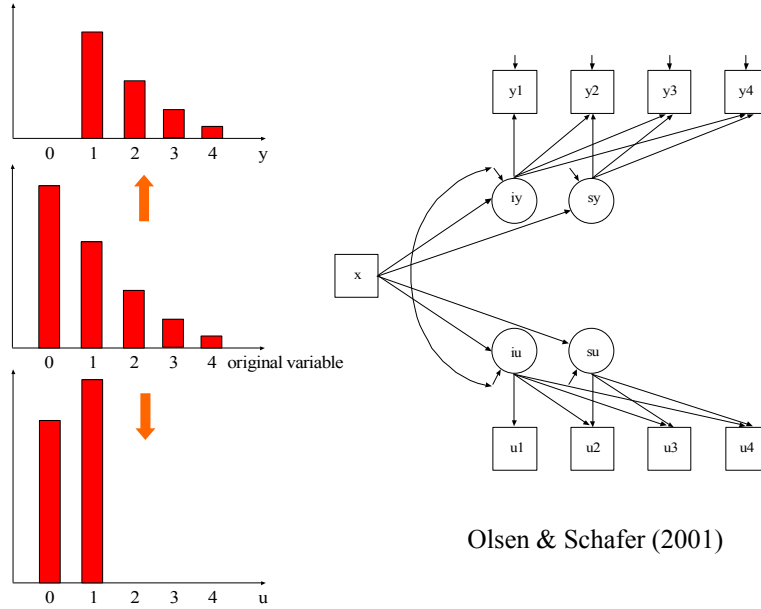


School (Between)

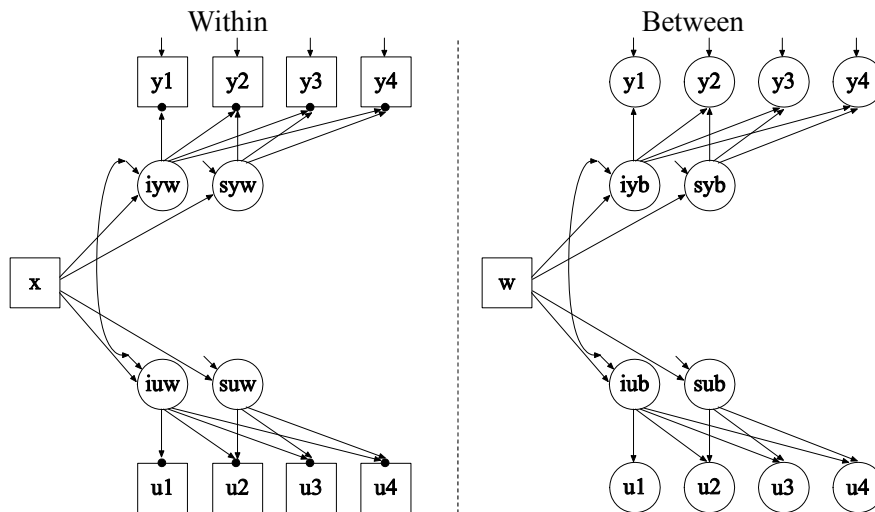


124

Two-Part (Semicontinuous) Growth Modeling



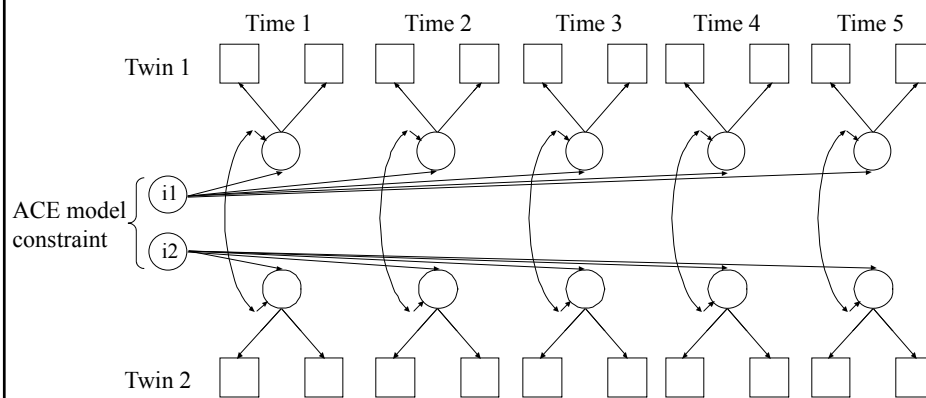
Two-Level, Two-Part Growth Modeling



Multiple Indicator Growth Modeling As Two-Level Analysis

127

Wide Data Format, Single-Level Approach

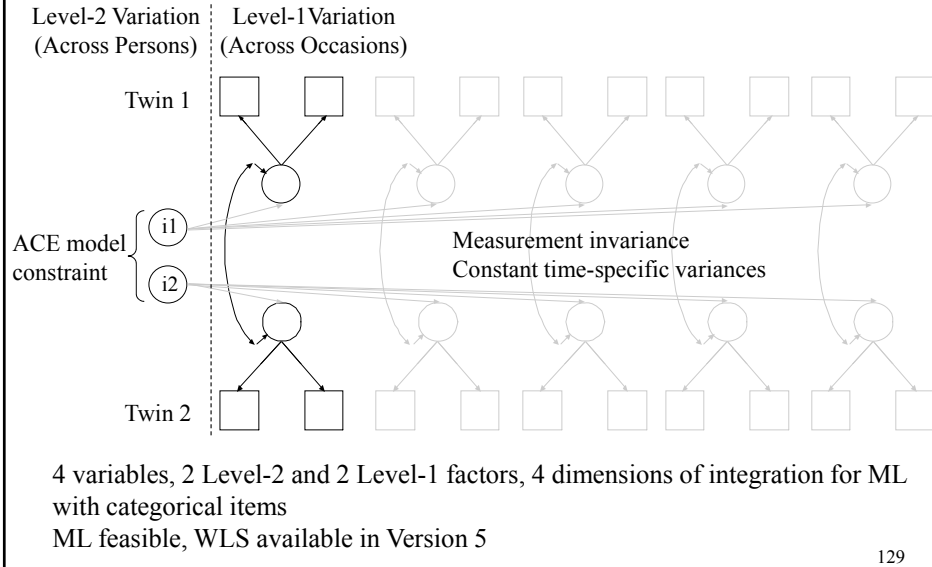


20 variables, 12 factors, 10 dimensions of integration for ML with categorical items

ML very hard, WLS easy

128

Long Format, Two-Level Approach



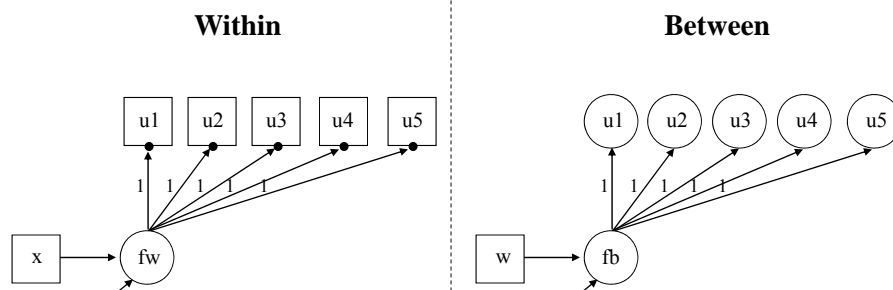
Multilevel Discrete-Time Survival Analysis

Multilevel Discrete-Time Survival Analysis

- Muthén and Masyn (2005) in Journal of Educational and Behavioral Statistics
- Masyn dissertation
- Continuous-time survival:
 - Asparouhov, Masyn and Muthén (2006)
 - Muthén, B., Asparouhov, T., Boye, M., Hackshaw, M. & Naegeli, A. (2009). Applications of continuous-time survival in latent variable models for the analysis of oncology randomized clinical trial data using Mplus. Technical Report.

131

Multilevel Discrete-Time Survival Frailty Modeling



Vermunt (2003)

132

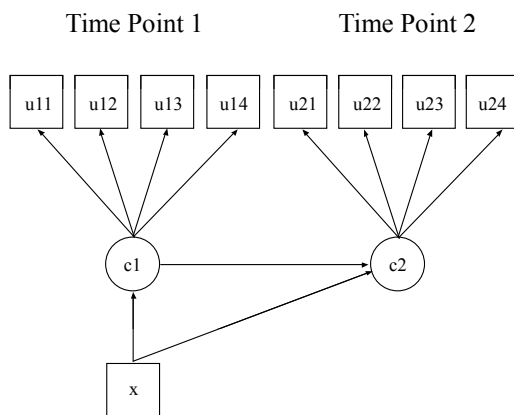
Two-Level Latent Transition Analysis

133

Latent Transition Analysis

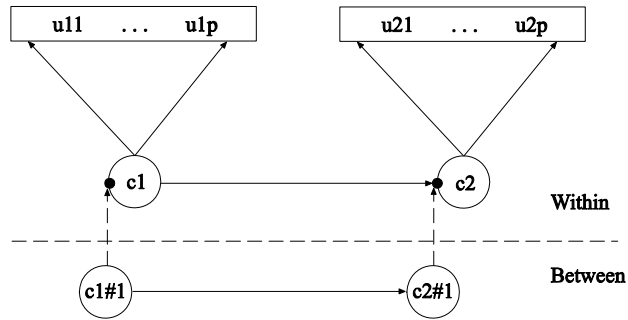
Transition Probabilities

		c2	
		1	2
c1	1	0.8	0.2
	2	0.4	0.6



134

Two-Level Latent Transition Analysis



Asparouhov, T. & Muthen, B. (2008). Multilevel mixture models. In Hancock, G. R., & Samuelsen, K. M. (Eds.). *Advances in latent variable mixture models*, pp 27 - 51. Charlotte, NC: Information Age Publishing, Inc.

135

Input For Two-Level LTA

```

CLUSTER = classrm;
USEVAR = stublf bkrulelf bkthinlf-teaself athortlf
         stubls bkrulels bkthinls-teasels athortls;
CATEGORICAL = stublf-athortls;
MISSING = all (999);
CLASSES = cf(2) cs(2);

DEFINE:
    CUT stublf-athortls(1.5);

ANALYSIS:
    TYPE = TWOLEVEL MIXTURE;
    PROCESS = 2;

MODEL:
    %WITHIN%
    %OVERALL%
    cs#1 ON cf#1;
    %BETWEEN%
    OVERALL%
    cs#1 ON cf#1;
    cs#1*1 cf#1*1;
    
```

136

Input For Two-Level LTA (Continued)

```

MODEL cf:
  %BETWEEN%
    %cf#1%
    [stublf$1-athort1f$1] (1-9);
    %cf#2%
    [stublf$1-athort1f$1] (11-19);
MODEL cs:
  %BETWEEN%
    %cs#1%
    [stubls$1-athort1s$1] (1-9);
    %cs#2%
    [stubls$1-athort1s$1] (11-19);
OUTPUT:
  TECH1 TECH8;
PLOT:
  TYPE = PLOT3;
  SERIES = stublf-athort1f(*);
  
```

137

Output Excerpts Two-Level LTA

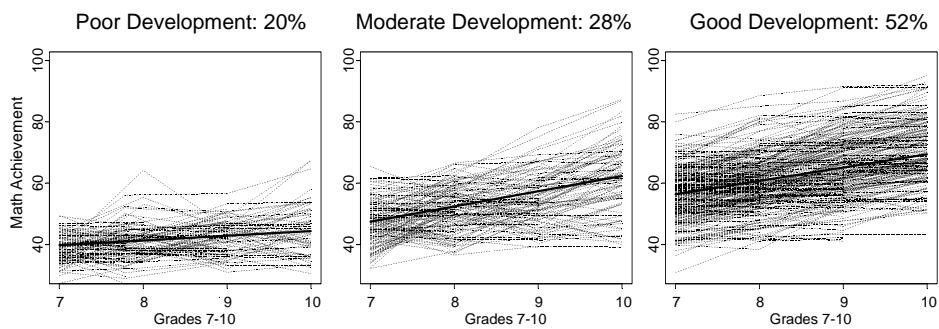
Categorical Latent Variables				
Within Level		Estimates	S.E.	Est./S.E.
CS#1	ON			
	CF#1	3.938	0.407	9.669
Means				
	CF#1	-0.126	0.189	-0.664
	CS#1	-1.514	0.221	-6.838
Between Level				
CS#1	ON			
	CF#1	0.411	0.15	2.735
Variances				
	CF#1	2.062	0.672	3.067
Residual Variances				
	CS#1	0.469	0.237	1.984

138

Multilevel Growth Mixture Modeling

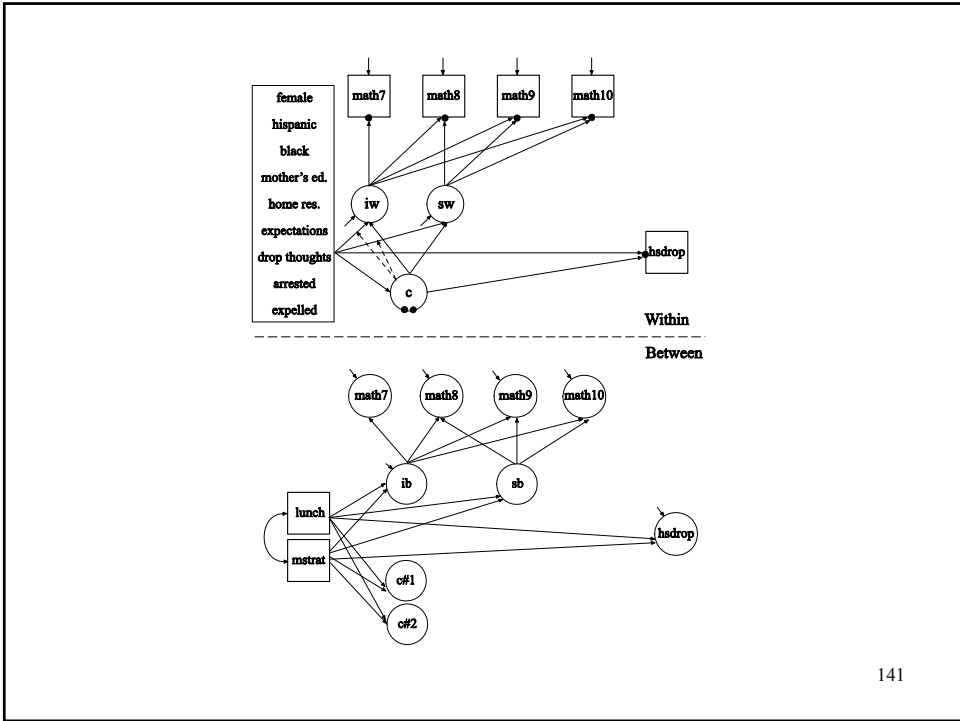
139

Growth Mixture Modeling: LSAY Math Achievement Trajectory Classes And The Prediction Of High School Dropout



Dropout:	69%	8%	1%
----------	-----	----	----

140



141

Input For A Multilevel Growth Mixture Model For LSAY Math Achievement

```

TITLE:      multilevel growth mixture model for LSAY math
            achievement

DATA:      FILE = lsayfull_Dropout.dat;

VARIABLE:  NAMES = female mothed homerer math7 math8 math9 math10
            expel arrest hisp black hsdrop expect lunch mstrat
            drophth7 schcode;
            !lunch = % of students eligible for full lunch
            !assistance (9th)
            !mstrat = ratio of students to full time math
            !teachers (9th)
            MISSING = ALL (9999);
            CATEGORICAL = hsdrop;
            CLASSES = c (3);
            CLUSTER = schcode;
            WITHIN = female mothed homerer expect drophth7 expel
            arrest hisp black;
            BETWEEN = lunch mstrat;

```

142

Input For A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

```
DEFINE:      lunch = lunch/100;  
            mstrat = mstrat/1000;  
  
ANALYSIS:   TYPE = MIXTURE TWOLEVEL;  
            ALGORITHM = INTEGRATION;  
  
OUTPUT:     SAMPSTAT STANDARDIZED TECH1 TECH8;  
  
PLOT:       TYPE = PLOT3;  
            SERIES = math7-math10 (s);
```

143

Input For A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

```
MODEL:      %WITHIN%  
            %OVERALL%  
            iw sw | math7@0 math8@1 math9@2 math10@3;  
            iw sw c hsdrop ON female hisp black mothed homeres  
            expect droptht7 expel arrest;  
  
            %c#1%  
            math7-math10*20;  
            iw*13 sw*3;  
  
            %c#2%  
            math7-math10*30;  
            iw*8 sw*3;  
            iw sw ON female hisp black mothed homeres expect  
            droptht7 expel arrest;
```

144

Input For A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

```
%c#3%  
math7-math10*10;  
iw*34 sw*2;  
iw sw ON female hisp black mothed homeres expect  
droptht7 expel arrest;  
  
%BETWEEN%  
%OVERALL%  
ib sb | math7@0 math8@1 math9@2 math10@3;  
sb@0;  
ib*1: hsdrop*1: ib WITH hsdrop;  
math7-math10@0;  
ib sb c hsdrop ON lunch mstrat;
```

145

Input For A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

```
%c#1%  
[ib*40 sb*1];  
[hsdrop$1*-.3];  
  
%c#2%  
[ib*40 sb*5];  
[hsdrop$1*.9];  
  
%c#3%  
[ib*45 sb*3];  
[hsdrop$1*1.2];
```

146

Output Excerpts A Multilevel Growth Mixture Model For LSAY Math Achievement

Summary of Data

Number of patterns	13
Number of y patterns	13
Number of u patterns	1
Number of clusters	44
Size (s)	Cluster ID with Size s
12	304
13	305
38	112
39	109
40	138
42	120
43	307
44	303
45	143 146

147

Output Excerpts A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

46	101				
48	144	106			
51	102	308			
52	136	118	133	111	
53	140	142	108	131	122 124
54	301	117	127	137	126
55	103	141	123		
56	110				
57	147				
58	121	105	145	135	
59	119				
73	104				
89	302				
94	309				
118	115				

148

Output Excerpts A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

THE STANDARD ERRORS OF THE MODEL PARAMETER ESTIMATES MAY NOT BE TRUSTWORTHY FOR SOME PARAMETERS DUE TO A NON-POSITIVE DEFINITE FIRST-ORDER DERIVATIVE PRODUCT MATRIX. THIS MAY BE DUE TO THE STARTING VALUES BUT MAY ALSO BE AN INDICATION OF MODEL NONIDENTIFICATION. THE CONDITION NUMBER IS -0.333D-16. PROBLEM INVOLVING PARAMETER 54.

THE NONIDENTIFICATION IS MOST LIKELY DUE TO HAVING MORE PARAMETERS THAN THE NUMBER OF CLUSTERS. REDUCE THE NUMBER OF PARAMETERS.

THE MODEL ESTIMATION TERMINATED NORMALLY

149

Output Excerpts A Multilevel Growth Mixture Model For LSAY Math Achievement (Continued)

Tests Of Model Fit

Loglikelihood

H0 Value	-26245.590
H0 Scaling Correction Factor for MLR	1.210

Information Criteria

Number of Free Parameters	124
Akaike (AIC)	52739.179
Bayesian (BIC)	53453.372
Sample-Size Adjusted BIC ($n^* = (n + 2) / 24$)	53059.399

FINAL CLASS COUNTS AND PROPORTIONS OF TOTAL SAMPLE SIZE BASED ON ESTIMATED POSTERIOR PROBABILITIES

Latent Classes

1	664.86988	0.28365
2	452.46710	0.19303
3	1226.66302	0.52332

CLASSIFICATION QUALITY

Entropy	0.625
---------	-------

150

**Output Excerpts A Multilevel Growth Mixture
Model For LSAY Math Achievement (Continued)**

Parameter	Estimates	S.E.	Two-Tailed	
			Est./S.E.	P-Value
Between Level				
Latent Class 1				
ib				
math7	1.000	0.000	999.000	999.000
math8	1.000	0.000	999.000	999.000
math9	1.000	0.000	999.000	999.000
math10	1.000	0.000	999.000	999.000
sb				
math7	0.000	0.000	999.000	999.000
math8	1.000	0.000	999.000	999.000
math9	2.000	0.000	999.000	999.000
math10	3.000	0.000	999.000	999.000
ib				
lunch	-2.154	1.298	-1.659	0.097
mstrat	-14.197	2.928	-4.849	0.000

151

**Output Excerpts A Multilevel Growth Mixture
Model For LSAY Math Achievement (Continued)**

Parameter	Estimates	S.E.	Two-Tailed	
			Est./S.E.	P-Value
sb ON				
lunch	-0.610	0.624	-0.978	0.328
mstrat	0.716	1.065	0.673	0.501
hsdrop ON				
lunch	1.226	0.543	2.256	0.024
mstrat	-0.269	1.455	-0.185	0.854
ib WITH				
hsdrop	-0.413	0.329	-1.256	0.209
Intercepts				
math7	0.000	0.000	999.000	999.000
math8	0.000	0.000	999.000	999.000
math9	0.000	0.000	999.000	999.000
math10	0.000	0.000	999.000	999.000
ib	40.636	1.672	24.310	0.000
sb	0.951	0.906	1.050	0.294

152

**Output Excerpts A Multilevel Growth Mixture
Model For LSAY Math Achievement (Continued)**

Parameter	Estimates	S.E.	Two-Tailed	
			Est./S.E.	P-Value
Thresholds				
hsdrop\$1	-0.350	0.490	-0.716	0.474
Residual Variances				
hsdrop	0.548	0.217	2.521	0.012
math7	0.000	0.000	999.000	999.000
math8	0.000	0.000	999.000	999.000
math9	0.000	0.000	999.000	999.000
math10	0.000	0.000	999.000	999.000
ib	3.469	1.012	3.429	0.001
sb	0.000	0.000	999.000	999.000

153

**Output Excerpts A Multilevel Growth Mixture
Model For LSAY Math Achievement (Continued)**

Parameter	Estimates	S.E.	Two-Tailed	
			Est./S.E.	P-Value
Within Level				
Categorical Latent Variables				
c#1 ON				
female	-0.801	0.261	-3.067	0.002
hisp	-0.066	0.616	-0.107	0.914
black	0.927	0.386	2.400	0.016
mothed	-0.015	0.107	-0.144	0.886
homeres	-0.058	0.080	-0.720	0.472
expect	-0.251	0.073	-3.425	0.001
droptht7	1.715	0.413	4.151	0.000
expel	0.606	0.389	1.559	0.119
arrest	1.076	0.378	2.845	0.004

154

**Output Excerpts A Multilevel Growth Mixture
Model For LSAY Math Achievement (Continued)**

Parameter	Estimates	S.E.	Two-Tailed	
			Est./S.E.	P-Value
c#2 ON				
female	-1.507	0.639	-2.359	0.018
hispanic	1.093	0.549	1.990	0.047
black	-1.078	0.715	-1.509	0.131
mothered	-0.229	0.138	-1.660	0.097
homeres	0.121	0.102	1.188	0.235
expect	-0.801	0.261	-3.067	0.002
droptht7	-0.066	0.616	-0.107	0.914
expel	0.927	0.386	2.400	0.016
arrest	-0.015	0.107	-0.144	0.886
Intercepts				
c#1	0.526	0.552	0.953	0.341
c#2	-0.518	0.659	-0.786	0.432

155

**Output Excerpts A Multilevel Growth Mixture
Model For LSAY Math Achievement (Continued)**

Parameter	Estimates	S.E.	Two-Tailed	
			Est./S.E.	P-Value
Between Level				
c#1 ON				
lunch	1.946	0.851	2.285	0.022
mstrat	-2.408	2.634	-0.914	0.361
c#2 ON				
lunch	0.248	1.450	0.171	0.864
mstrat	-0.810	2.382	-0.340	0.734

156

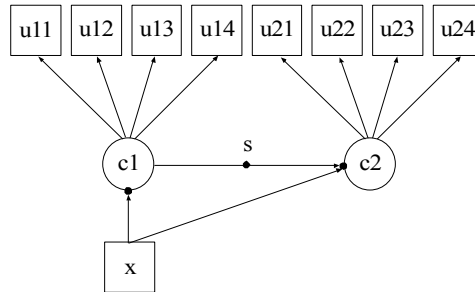
Between-Level Latent Classes

157

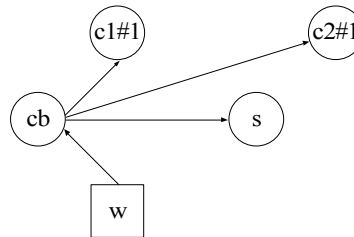
Latent Transition Analysis

158

UG Ex. 10.3 : Two-Level LTA With A Covariate And A Between-Level Categorical Latent Variable



Within



Between

159

Input For Two-Level LTA

```

TITLE:      this is an example of a two-level LTA with a covariate
            and a between-level categorical latent variable
DATA:      FILE = ex8.dat;
VARIABLE:  NAMES ARE u11-u14 u21-u24 x w dumb dum1 dum2 clus;
            USEVARIABLES = u11-w;
            CATEGORICAL = u11-u14 u21-u24;
            CLASSES = cb(2) c1(2) c2(2);
            WITHIN = x;
            BETWEEN = cb w;
            CLUSTER = clus;
ANALYSIS:  TYPE = TWOLEVEL MIXTURE;
            PROCESSORS = 2;
MODEL:
            %WITHIN%
            %OVERALL%
            c2#1 ON c1#1 x;
            c1#1 ON x;
            %BETWEEN%
            %OVERALL%
            c1#1 ON cb#1;
    
```

160

Input For Two-Level LTA (Continued)

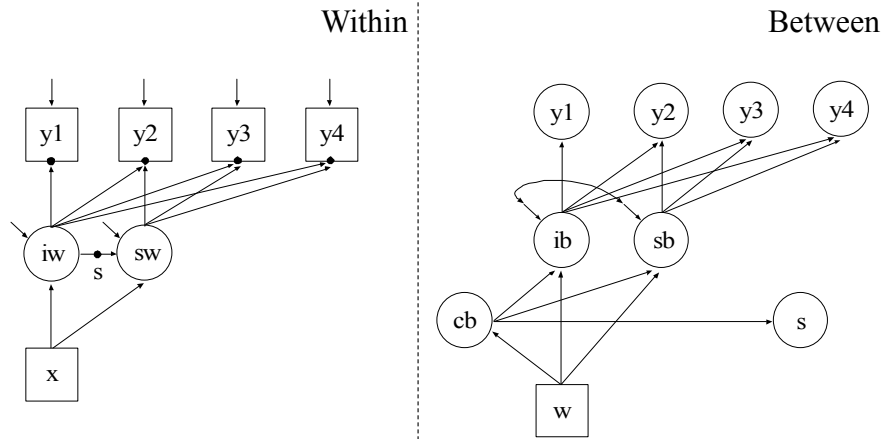
```
      c2#1 ON cb#1;  
      cb#1 ON w;  
MODEL cb:  
      %WITHIN%  
      %cb#1%  
      c2#1 ON c1#1;  
MODEL c1:  
      %BETWEEN%  
      %c1#1%  
      [u11$1-u14$1] (1-4);  
      %c1#2%  
      [u11$1-u14$1] (5-8);  
MODEL c2:  
      %BETWEEN%  
      %c2#1%  
      [u21$1-u24$1] (1-4);  
      %c2#2%  
      [u21$1-u24$1] (5-8);  
OUTPUT:  TECH1 TECH8;
```

161

Growth Modeling

162

UG Ex. 10.8: Two-Level Growth Model For A Continuous Outcome (Three-Level Analysis) With A Between-Level Categorical Latent Variable



163

Input For Two-Level Growth Model

```

TITLE:      this is an example of a two-level growth model for a
             continuous outcome (three-level analysis) with a
             between-level categorical latent variable

DATA:      FILE = ex5.dat;

VARIABLE:  NAMES ARE y1-y4 x w dummy clus;
           USEVARIABLES = y1-w;
           CLASSES = cb(2);
           WITHIN = x;
           BETWEEN = cb w;
           CLUSTER = clus;

ANALYSIS:  TYPE = TWOLEVEL MIXTURE RANDOM;
           PROCESSORS = 2;

MODEL:

%WITHIN%
%OVERALL%
iw sw | y1@0 y2@1 y3@2 y4@3;
y1-y4 (1);
iw sw ON x;
s | sw ON iw;

```

164

Input For Two-Level Growth Model (Continued)

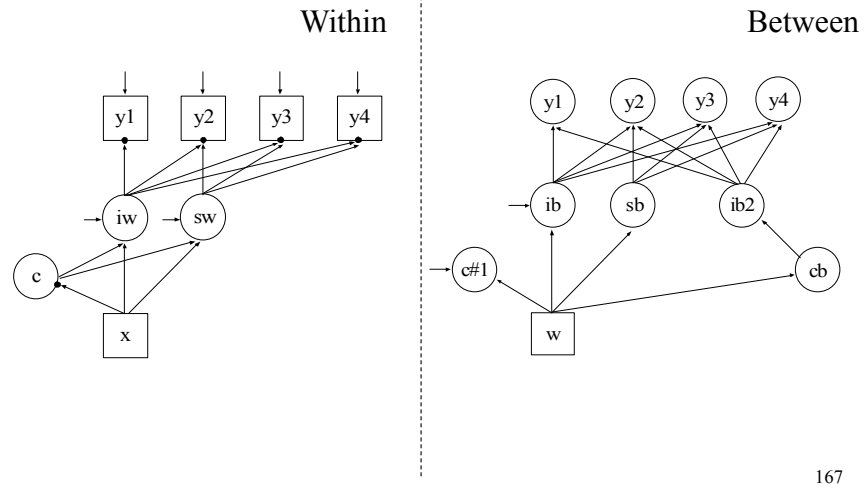
```
%BETWEEN%  
%OVERALL%  
ib sb | y1@0 y2@1 y3@2 y4@3;  
y1-y4@0;  
ib sb ON w;  
cb#1 ON w;  
s@0;  
%cb#1%  
[ib sb s];  
%cb#2%  
[ib sb s];  
OUTPUT: TECH1 TECH8;
```

165

Growth Mixture Modeling

166

UG Ex.10.10: Two-Level GMM (Three-Level Analysis) For A Continuous Outcome With A Between-Level Categorical Latent Variable



167

Input For Two-Level GMM (Three-Level Analysis)

```

TITLE:      this is an example of a two-level GMM (three-level
            analysis) for a continuous outcome with a between-
            level categorical latent variable
DATA:       FILE = ex6.dat;
VARIABLE:   NAMES ARE y1-y4 x w dummyb dummy clus;
            USEVARIABLES = y1-w;
            CLASSES = cb(2) c(2);
            WITHIN = x;
            BETWEEN = cb w;
            CLUSTER = clus;
ANALYSIS:   TYPE = TWOLEVEL MIXTURE;
            PROCESSORS = 2;
MODEL:      %WITHIN%
            %OVERALL%
            iw sw | y1@0 y2@1 y3@2 y4@3;
            iw sw ON x;
            c#1 ON x;
            %BETWEEN%
            %OVERALL%
    
```

168

Input For Two-Level GMM (Continued)

```
ib sb | y1@0 y2@1 y3@2 y4@3;  
ib2 | y1-y4@1;  
y1-y4@0;  
ib sb ON w;  
c#1 ON w;  
sb@0; c#1;  
ib2@0;  
cb#1 ON w;  
MODEL c:  
%BETWEEN%  
%c#1%  
[ib sb];  
%c#2%  
[ib sb];  
MODEL cb:  
%BETWEEN%  
%cb#1%  
[ib2@0];  
%cb#2%  
[ib2];  
OUTPUT: TECH1 TECH8;
```

169

Multilevel Growth Mixture Modeling Example

Baltimore randomized field trial (Muthén, B. & Asparouhov, 2008, pp. 155-157):

- 362 boys in Baltimore public schools observed four times in grades 1-3 in 27 classrooms
- Outcome is aggressive-disruptive behavior
- Classroom-based good behavior intervention in Fall of 1st grade
- Student-level trajectory classes and classroom-level classes

170

Multilevel Growth Mixture Modeling Example (Continued)

Findings:

- 3 student-level trajectory classes, 2 classroom-level classes
- In classrooms with a low level of aggressive behavior, students who are in the two highest trajectory classes benefit from the intervention
- In classrooms with a high level of aggression, only students who are in the lowest trajectory class benefit from the intervention

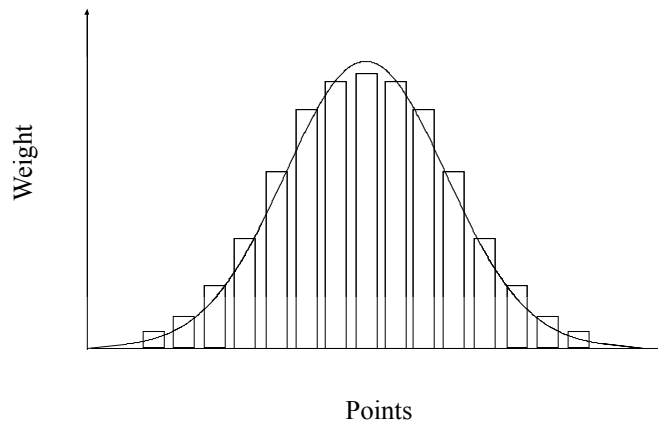
Source: Muthén, B. & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), *Longitudinal Data Analysis*, pp. 143-165. Boca Raton: Chapman & Hall/CRC Press.

171

Technical Aspects Of Multilevel Modeling

172

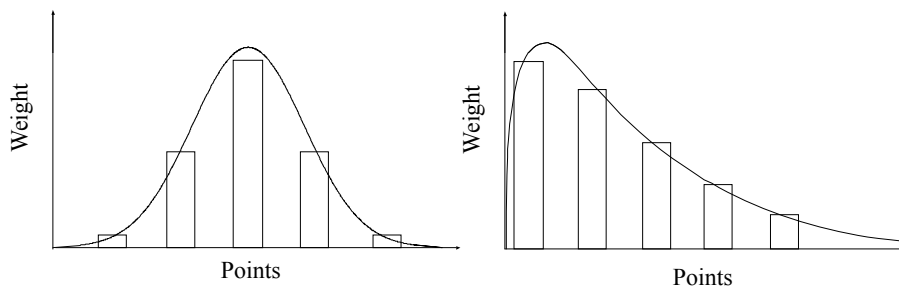
Numerical Integration With A Normal Latent Variable Distribution



Fixed weights and points

173

Non-Parametric Estimation Of The Random Effect Distribution



Estimated weights and points
(class probabilities and class means)

174

Numerical Integration

Numerical integration is needed with maximum likelihood estimation when the posterior distribution for the latent variables does not have a closed form expression. This occurs for models with categorical outcomes that are influenced by continuous latent variables, for models with interactions involving continuous latent variables, and for certain models with random slopes such as multilevel mixture models.

When the posterior distribution does not have a closed form, it is necessary to integrate over the density of the latent variables multiplied by the conditional distribution of the outcomes given the latent variables. Numerical integration approximates this integration by using a weighted sum over a set of integration points (quadrature nodes) representing values of the latent variable.

175

Numerical Integration (Continued)

Numerical integration is computationally heavy and thereby time-consuming because the integration must be done at each iteration, both when computing the function value and when computing the derivative values. The computational burden increases as a function of the number of integration points, increases linearly as a function of the number of observations, and increases exponentially as a function of the dimension of integration, that is, the number of latent variables for which numerical integration is needed.

176

Practical Aspects Of Numerical Integration

- Types of numerical integration available in Mplus with or without adaptive quadrature
 - Standard (rectangular, trapezoid) – default with 15 integration points per dimension
 - Gauss-Hermite
 - Monte Carlo
- Computational burden for latent variables that need numerical integration
 - One or two latent variables Light
 - Three to five latent variables Heavy
 - Over five latent variables Very heavy

177

Practical Aspects Of Numerical Integration (Continued)

- Suggestions for using numerical integration
 - Start with a model with a small number of random effects and add more one at a time
 - Start with an analysis with TECH8 and MITERATIONS=1 to obtain information from the screen printing on the dimensions of integration and the time required for one iteration and with TECH1 to check model specifications
 - With more than 3 dimensions, reduce the number of integration points to 5 or 10 or use Monte Carlo integration with the default of 500 integration points
 - If the TECH8 output shows large negative values in the column labeled ABS CHANGE, increase the number of integration points to improve the precision of the numerical integration and resolve convergence problems
 - Explore using a random subsample

178

Technical Aspects Of Numerical Integration

Maximum likelihood estimation using the EM algorithm computes in each iteration the posterior distribution for normally distributed latent variables f ,

$$[f|y] = [f][y|f] / [y], \quad (97)$$

where the marginal density for $[y]$ is expressed by integration

$$[y] = \int [f][y|f] df. \quad (98)$$

- Numerical integration is not needed for normally distributed y - the posterior distribution is normal

179

Technical Aspects Of Numerical Integration (Continued)

- Numerical integration needed for:
 - Categorical outcomes u influenced by continuous latent variables f , because $[u]$ has no closed form
 - Latent variable interactions $f \times x, f \times y, f_1 \times f_2$, where $[y]$ has no closed form, for example

$$[y] = \int [f_1, f_2][y|f_1, f_2, f_1 f_2] df_1 df_2 \quad (99)$$

- Random slopes, e.g. with two-level mixture modeling

Numerical integration approximates the integral by a sum

$$[y] = \int [f][y|f] df = \sum_{k=1}^K w_k [y|f_k] \quad (100)$$

180

References

Longitudinal Data

- Asparouhov, T. & Muthén, B. (2008). Multilevel mixture models. In Hancock, G. R., & Samuelsen, K. M. (Eds.), *Advances in latent variable mixture models*, pp. 27-51. Charlotte, NC: Information Age Publishing, Inc.
- Choi, K.C. (2002). Latent variable regression in a three-level hierarchical modeling framework: A fully Bayesian approach. Doctoral dissertation, University of California, Los Angeles.
- Fitzmaurice, Laird, & Ware (2004). *Applied longitudinal analysis*. Wiley & Sons.
- Fitzmaurice, Davidian, Verbeke, & Molenberghs (2009). *Longitudinal Data Analysis*. Chapman & Hall/CRC Press.
- Khoo, S.T. & Muthén, B. (2000). Longitudinal data on families: Growth modeling alternatives. *Multivariate applications in substance use research*, J. Rose, L. Chassin, C. Presson & J. Sherman (eds.), Hillsdale, N.J.: Erlbaum, pp. 43-78. (#79)
- Masyn, K. E. (2003). Discrete-time survival mixture analysis for single and recurrent events using latent variables. Doctoral dissertation, University of California, Los Angeles.

181

References (Continued)

- Muthén, B. (1997). Latent variable modeling with longitudinal and multilevel data. In A. Raftery (ed.) *Sociological Methodology* (pp. 453-480). Boston: Blackwell Publishers.
- Muthén, B. (1997). Latent variable growth modeling with multilevel data. In M. Berkane (ed.), *Latent variable modeling with application to causality* (149-161), New York: Springer Verlag.
- Muthén, B. & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), *Longitudinal Data Analysis*, pp. 143-165. Boca Raton: Chapman & Hall/CRC Press.
- Muthén, B. & Asparouhov, T. (2011). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. In J. Hox & J.K. Roberts (eds), *The Handbook of Advanced Multilevel Analysis*, pp 15-40. New York: Taylor and Francis.
- Muthén, B. & Masyn, K. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics*, 30, 27-58.
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Second edition. Newbury Park, CA: Sage Publications.

182

References (Continued)

- Snijders, T. & Bosker, R. (1999). Multilevel analysis. An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.
- Seltzer, M., Choi, K., Thum, Y.M. (2002). Examining relationships between where students start and how rapidly they progress: Implications for conducting analyses that help illuminate the distribution of achievement within schools. CSE Technical Report 560. CRESST, University of California, Los Angeles.

Numerical Integration

- Aitkin, M. A general maximum likelihood analysis of variance components in generalized linear models. Biometrics, 1999, 55, 117-128.
- Bock, R.D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. Psychometrika, 46, 443-459.

183

References (Continued)

- Schilling, S. & Bock, R.D. (2005). High-dimensional maximum marginal likelihood item factor analysis by adaptive quadrature. Psychometrika, 70(3), p533-555.
- Skrondal, A. & Rabe-Hesketh, S. (2004). Generalized latent variable modeling. Multilevel, longitudinal, and structural equation models. London: Chapman Hall.

184