5. Two-Level Analysis
With Random Intercepts (Difficulties) And Random Loadings (Discrimination)

5.1 Advances In Multiple-Group Analysis: Invariance Across Groups
5.1.1 Hospital Data Example
5.1.2 Hospital As Fixed Mode: Conventional Multiple-Group Factor Analysis
5.1.3 Hospital As Random Mode: Conventional Two-Level Factor Analysis
5.2 New Solution No. 2: Group Is Random Mode Two-Level Factor Analysis With Random Loadings
5.2.1 New Solution No.2: Group Is Random Mode. UG Ex9.19
5.2.2 Monte Carlo Simulations For Groups As Random Mode: Two-Level Random Loadings Modeling
5.3 Two-Level Random Loadings In IRT: The PISA Data
5.4 Testing For Non-Zero Variance Of Random Loadings
5.5 Two-Level Random Loadings: Individual Differences Factor Analysis

6. 3-Level Analysis
6.1 Types of Observed Variables and Random Slopes for 3-Level Analysis
6.2 3-Level Regression
6.3 3-Level Regression: Nurses Data
6.4 3-Level Path Analysis: UG Example 9.21
6.5 3-Level MIMIC Analysis
6.6 3-Level Growth Analysis
6.7 TYPE=THREELEVEL COMPLEX
6.8 3-Level and Cross-Classified Multiple Imputation

7. Cross-Classified Analysis: Introductory
   7.1 Cross-Classified Regression
   7.2 Cross-Classified Regression: UG Example 9.24
   7.3 Cross-Classified Regression: Pupcross Data
   7.4 Cross-Classified Path Analysis: UG Example 9.25

8. Cross-Classified Analysis, More Advanced
   8.1 2-Mode Path Analysis: Random Contexts In Gonzalez Et Al.
   8.2 2-Mode Path Analysis: Monte Carlo Simulation
   8.3 Cross-Classified SEM
   8.4 Monte Carlo Simulation Of Cross-Classified SEM
   8.5 Cross-Classified Models: Types Of Random Effects
   8.6 Random Items, Generalizability Theory
   8.7 Random Item 2-Parameter IRT: TIMMS Example
   8.8 Random Item Rasch IRT Example

9. Advances In Longitudinal Analysis
9.1 BSEM for Aggressive-Disruptive Behavior In The Classroom
9.2 Cross-Classified Analysis Of Longitudinal Data
9.3 Cross-Classified Monte Carlo Simulation
9.4 Cross-Classified Growth Modeling: UG Example 9.27
9.5 Cross-Classified Analysis Of Aggressive-Disruptive Behavior In The Classroom
9.6 Cross-Classified / Multiple Membership Applications
1. Overview Of Day 3

More advanced day, focusing on the cutting-edge features in Version 7 related to multilevel analysis of complex survey data and item response theory (IRT) extensions.

Topics:

- IRT analysis, categorical factor analysis
  - Basic IRT
  - Intermediate IRT
- Multilevel analysis
  - Two-level analysis with random loadings (discriminations)
  - Three-level analysis
  - Cross-classified analysis
- Advanced IRT analysis
  - Group comparisons such as cross-national studies
  - Random items, G-theory
  - Random contexts
  - Longitudinal studies
Mplus Readings Related To Day 3


Let \( u_{ij} \) be a binary item \( j (j = 1, 2, \ldots p) \) for individual \( i (i = 1, 2, \ldots n) \), and express the probability of the outcome \( u_{ij} = 1 \) for this item as a function of \( m \) factors \( \eta_{i1}, \eta_{i2}, \ldots, \eta_{im} \) as follows,

\[
P(u_{ij} = 1 \mid \eta_{i1}, \eta_{i2}, \ldots, \eta_{im}) = F[-\tau_j + \sum_{k=1}^{m} \lambda_{jk} \eta_{ik}],
\]

where with the logistic model and the general argument \( x \), \( F[x] \) represents the logistic function

\[
F[x] = \frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}},
\]

and with the probit model \( F[x] \) represents the standard normal distribution function \( \Phi[x] \).

The model is completed by assuming conditional independence among the items and normality for the factors.
Item Characteristic Curves From Maximum Likelihood IRT Analysis Of Seven Binary Aggression Items Measuring A Single Factor
Information Curve From Maximum Likelihood IRT Analysis Of Seven Binary Aggression Items Measuring A Single Factor
### Mplus Offers Three Estimators For IRT And Factor Analysis Of Categorical Items

<table>
<thead>
<tr>
<th>Criteria for comparison</th>
<th>Weighted least squares</th>
<th>Maximum likelihood</th>
<th>Bayes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large number of factors</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Large number of variables</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Large number of subjects</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Small number of subjects</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Statistical efficiency</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Missing data handling</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Test of LRV structure</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Ordered polytomous variables</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Heywood cases</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Zero cells</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Residual correlations</td>
<td>+</td>
<td>–</td>
<td>±</td>
</tr>
</tbody>
</table>
- High-dimensional analysis using WLSMV, Bayes, and ML two-tier
- Bi-factor EFA
- Modification indices, correlated residuals
- Multiple-group analysis
- Mixtures*
- Complex survey data handling: Stratification, weights
- Multilevel: two-level, three-level, and cross-classified
- Random loadings (discrimination) using Bayesian analysis
- Random item IRT
- Random subjects and contexts

Bayesian estimation of exploratory factor analysis implemented in Mplus version 7 for models with continuous and categorical variables

Asparouhov and Muthén (2012). Comparison of computational methods for high dimensional item factor analysis

Asymptotically the Bayes EFA is the same as the ML solution

Bayes EFA for categorical variable is a full information estimation method without using numerical integration and therefore feasible with any number of factors

New in Mplus Version 7: Improved performance of ML-EFA for categorical variables, in particular high-dimensional EFA models with Montecarlo integration; improved unrotated starting values and standard errors
The first step in the Bayesian estimation is the estimation of the unrotated model as a CFA model using the MCMC method.

- Obtain posterior distribution for the unrotated solution.
- To obtain the posterior distribution of the rotated parameters we simply rotate the generated unrotated parameters in every MCMC iteration, using oblique or orthogonal rotation.
- No priors. Priors could be specified currently only for the unrotated solution.
- If the unrotated estimation takes many iterations to converge, use THIN to reduce the number of rotations.
This MCMC estimation is complicated by identification issues that are similar to label switching in the Bayesian estimation of Mixture models.

There are two types of identification issues in the Bayes EFA estimation.

The first type is identification issues related to the unrotated parameters: loading sign switching.

Solution: constrain the sum of the loadings for each factor to be positive. Implemented in Mplus Version 7 for unrotated EFA and CFA. New in Mplus Version 7, leads to improved convergence in Bayesian SEM estimation.

\[ \sum_{i=1}^{p} \lambda_{ij} > 0 \]
The second type is identification issues related to the rotated parameters: loading sign switching and order of factor switching.

Solution: Align the signs $s_j$ and factor order $\sigma$ to minimize MSE between the current estimates $\lambda$ and the average estimate from the previous MCMC iterations $L$

$$\sum_{i,j} (s_j \lambda_i \sigma(j) - L_{ij})^2$$

Minimize over all sign allocations $s_j$ and factor permutations $\sigma$
Factor scores for the rotated solutions also available. Confidence intervals and posterior distribution plots

Using the optimal rotation in each MCMC iteration we rotate the unrotated factors to obtain the posterior distribution of the rotated factors

With continuous variables Bayes factor is computed to compare EFA with different number of factors. PPP value is computed with continuous or categorical variables
Bayes factors is an easy and quick way to compare models using BIC

\[
BF = \frac{P(H_1)}{P(H_0)} = \frac{\text{Exp}(-0.5BIC_{H_1})}{\text{Exp}(-0.5BIC_{H_0})}
\]

Values of \( BF \) greater than 3 are considered evidence in support of \( H_1 \)

New in Mplus Version 7: BIC is now included for all models with continuous items (single level and no mixtures)

The above method can be used to easily compare nested and non-nested models
Bayes EFA: Simulation Study \((n = 500)\)

Absolute bias, coverage and log-likelihood for EFA model with 7 factors and 35 ordered polytomous variables.

<table>
<thead>
<tr>
<th>Method</th>
<th>(\lambda_{11})</th>
<th>(\lambda_{12})</th>
<th>Log-Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mplus Monte 500</td>
<td>.01(0.97)</td>
<td>.00(0.83)</td>
<td>-28580.3</td>
</tr>
<tr>
<td>Mplus Monte 5000</td>
<td>.01(0.96)</td>
<td>.00(0.87)</td>
<td>-28578.4</td>
</tr>
<tr>
<td>Mplus Bayes</td>
<td>.01(.90)</td>
<td>.00(.96)</td>
<td>-</td>
</tr>
<tr>
<td>Mplus WLSMV</td>
<td>.00(.94)</td>
<td>.00(.89)</td>
<td>-</td>
</tr>
<tr>
<td>IRTPRO MHRM</td>
<td>.00(.54)</td>
<td>.00(.65)</td>
<td>-28665.2</td>
</tr>
</tbody>
</table>
Average standard error, ratio between average standard error and standard deviation for the EFA model with 7 factors and ordered polytomous variables.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\lambda_{11}$</th>
<th>$\lambda_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mplus Monte 500</td>
<td>0.033(1.00)</td>
<td>0.031(0.72)</td>
</tr>
<tr>
<td>Mplus Monte 5000</td>
<td>0.033(0.99)</td>
<td>0.035(0.81)</td>
</tr>
<tr>
<td>Mplus Bayes</td>
<td>0.030(0.97)</td>
<td>0.032(0.98)</td>
</tr>
<tr>
<td>Mplus WLSMV</td>
<td>0.030(0.97)</td>
<td>0.038(0.85)</td>
</tr>
<tr>
<td>IRTPRO MHRM</td>
<td>0.012(0.42)</td>
<td>0.026(0.65)</td>
</tr>
</tbody>
</table>

Bayes EFA is the most accurate full information estimation method for high-dimensional EFA with categorical variables.
Example is based on Mplus User’s Guide example 4.1 generated with 4 factors and 12 indicators.

```
DATA: FILE IS ex4.1.dat;
VARIABLE: NAMES ARE y1-y12;
ANALYSIS: TYPE = EFA 1 5; estimator=bayes;
```

We estimate EFA with 1, 2, 3, 4 or 5 factors.
Bayes factor results: The posterior probability that the number of factors is 4 is: 99.59%. However, this is a power result - there is enough information in the data to support 4 factors and not enough to support 5 factors. Use BITER = (10000)

| POSTERIOR PROBABILITIES FOR ALL MODELS:          |
| 1-FACTOR MODEL    | 0.0000 |
| 2-FACTOR MODEL    | 0.0000 |
| 3-FACTOR MODEL    | 0.0041 |
| 4-FACTOR MODEL    | 0.9959 |
| 5-FACTOR MODEL    | 0.0000 |
EXPLORATORY FACTOR ANALYSIS WITH 4 FACTOR(S):

Number of Free Parameters 66

Bayesian Posterior Predictive Checking using Chi-Square

95% Confidence Interval for the Difference Between the Observed and the Replicated Chi-Square Values

-35.978 41.864

Posterior Predictive P-Value 0.376

Information Criterion

Deviance (DIC) 3231.767
Estimated number of parameters (pD) 67.580
Bayesian (BIC) 3423.991
Bayes factor for the model with 4 factor(s) against the model with 3 factor(s): 243.6367
Log of Bayes factor for the model with 4 factor(s) against the model with 3 factor(s): 5.4957
### Bayes EFA: Results

#### GEOMIN ROTATED LOADINGS (* significant at 5% level)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>0.622*</td>
<td>-0.008</td>
<td>0.028</td>
<td>-0.026</td>
</tr>
<tr>
<td>Y2</td>
<td>0.762*</td>
<td>-0.025</td>
<td>0.035</td>
<td>0.010</td>
</tr>
<tr>
<td>Y3</td>
<td>0.734*</td>
<td>0.055</td>
<td>-0.038</td>
<td>0.021</td>
</tr>
<tr>
<td>Y4</td>
<td>-0.112</td>
<td>0.592*</td>
<td>-0.008</td>
<td>-0.150</td>
</tr>
<tr>
<td>Y5</td>
<td>0.021</td>
<td>0.844*</td>
<td>-0.049</td>
<td>0.007</td>
</tr>
<tr>
<td>Y6</td>
<td>0.053</td>
<td>0.639*</td>
<td>-0.029</td>
<td>-0.007</td>
</tr>
<tr>
<td>Y7</td>
<td>-0.009</td>
<td>0.037</td>
<td>0.788*</td>
<td>0.004</td>
</tr>
<tr>
<td>Y8</td>
<td>0.032</td>
<td>-0.059</td>
<td>0.646*</td>
<td>-0.115</td>
</tr>
<tr>
<td>Y9</td>
<td>0.008</td>
<td>0.040</td>
<td>0.602*</td>
<td>0.095</td>
</tr>
<tr>
<td>Y10</td>
<td>-0.042</td>
<td>0.065</td>
<td>0.005</td>
<td>0.671*</td>
</tr>
<tr>
<td>Y11</td>
<td>0.013</td>
<td>-0.078</td>
<td>-0.027</td>
<td>0.714*</td>
</tr>
<tr>
<td>Y12</td>
<td>0.051</td>
<td>-0.054</td>
<td>0.009</td>
<td>0.710*</td>
</tr>
</tbody>
</table>
4. Bayes Factor Scores Handling

- Version 7 uses improved language for factor scores with Bayesian estimation. The same language as for other estimators.

- SAVEDATA: FILE=fs.dat; SAVE=FS(300); FACTORS=factor names; This command specifies that 300 imputations will be used to estimate the factor scores and that plausible value distributions are available for plotting.

- Posterior mean, median, confidence intervals, standard error, all imputed values, distribution plot for each factor score for each latent variable for any model estimated with the Bayes estimator.

- Bayes factor score advantages: more accurate than ML factor scores in small sample size, Bayes factor score more accurate in secondary analysis such as for example computing correlations between factor.
Asparouhov & Muthén (2010). Plausible values for latent variables using Mplus

Factor analysis with 3 indicators and 1 factor. Simulated data with N=45. True factor values are known. Bayes factor score estimates are more accurate. Bayes factor score SE are more accurate

ML factor scores are particularly unreliable when Var(Y) is near 0

<table>
<thead>
<tr>
<th></th>
<th>ML</th>
<th>Bayes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.636</td>
<td>0.563</td>
</tr>
<tr>
<td>Coverage</td>
<td>20%</td>
<td>89%</td>
</tr>
<tr>
<td>Average SE</td>
<td>0.109</td>
<td>0.484</td>
</tr>
</tbody>
</table>
5. Two-Level Analysis
With Random Intercepts (Difficulties)
And Random Loadings (Discrimination)

- Measurement invariance across groups
- Overview and an example of hospital ratings (continuous items)
- Two-level random loadings in IRT using the PISA math data (binary items)
- Testing for non-zero variance of random loadings
- Individual differences factor analysis
5.1 Advances In Multiple-Group Analysis: Invariance Across Groups

- An old dilemma
- Two new solutions
Fixed Versus Random Groups

- **Fixed mode:**
  - Inference to only the groups in the sample
  - Small to medium number of groups

- **Random mode:**
  - Inference to a population of groups from which the current set of groups is a random sample
  - Medium to large number of groups
New solution no. 1, suitable for a small to medium number of groups
- A new BSEM approach where group is a fixed mode:
  Multiple-group BSEM (see Utrecht video, Part 1 handout)
- Approximate invariance allowed

New solution no. 2, suitable for a medium to large number of groups
- A new Bayes approach where group is a random mode
- No limit on the number of groups

- Survey of 67 hospitals, \( n = 7168 \) employee respondents, approximately 100/hospital
- 6 dimensions of an overall ”quality improvement implementation” based on the Malcom Baldrige National Quality Award criteria
- Focus on 10 items measuring a leadership dimension
- Continuous items
Hospital as Fixed Mode:
- Old approach: Conventional multiple-group factor analysis
- New approach: BSEM multiple-group factor analysis

Hospital as Random Mode:
- Old approach: Conventional two-level factor analysis
- New approach: Bayes random loadings two-level factor analysis
  (random factor variances also possible)
5.1.2 Hospital As Fixed Mode: Conventional Multiple-Group Factor Analysis

Regular ML analysis:

VARIABLE:  USEVARIABLES = lead21-lead30! info31-info37!
  ! straqp38-straqp44 hru45-hru52 qm53-qm58 hosp;
MISSING = ALL(-999);
!CLUSTER = hosp;
GROUPING = hosp (101 102 104 105 201 301-306
308 310-314 316-320 322 401-403 405-409 412-416
501-503 505-512 602-609 612-613 701 801 901-908);
ANALYSIS:  ESTIMATOR = ML;
PROCESSORS = 8;
MODEL:  lead BY lead21-lead30; ! specifies measurement invariance
PLOT:  TYPE = PLOT2;
OUTPUT:  TECH1 TECH8 MODINDICES(ALL);
Maximum-likelihood analysis with $\chi^2$ test of model fit and modification indices.

Holding measurement parameters equal across groups/hospitals results in poor fit with many moderate-sized modification indices and none that sticks out as much larger than the others.

Conventional multiple-group factor analysis ”fails”.
5.1.3 Group As Random Mode: Conventional Two-Level Factor Analysis

- Recall random effects ANOVA (individual \( i \) in cluster \( j \)):

\[
y_{ij} = \nu + \eta_j + \varepsilon_{ij} = y_{Bj} + y_{Wj}
\]  

(3)

- Two-level factor analysis (\( r = 1, 2, \ldots, p \) items; 1 factor on each level):

\[
y_{rij} = \nu_r + \lambda_{B} \eta_{Bj} + \varepsilon_{Brij} + \lambda_{W} \eta_{Wij} + \varepsilon_{Wrij}
\]  

(4)

- Alternative expression often used in 2-level IRT:

\[
y_{rij} = \nu_r + \lambda_r \eta_{ij} + \varepsilon_{rij},
\]  

(5)

\[
\eta_{ij} = \eta_{Bj} + \eta_{Wij},
\]  

(6)

so that \( \lambda \) is the same for between and within.
USEVARIABLES = lead21-lead30;
MISSING = ALL (-999);
CLUSTER = hosp;

ANALYSIS: TYPE = TWOLEVEL;
ESTIMATOR = ML;
PROCESSORS = 8;

MODEL: %WITHIN%
leadw BY lead21-lead30* (lam1-lam10);
leadw @1;

%BETWEEN%
leadb BY lead21-lead30* (lam1-lam10);
leadb;

OUTPUT: TECH1 TECH8 MODINDICES(ALL);
Results For Hospital As Random Mode: Conventional Two-Level Factor Analysis

Equality of within- and between-level factor loadings cannot be rejected by $\chi^2$ difference testing.

10% of the total variance in the leadership factor is due to between-hospital variation.

No information about measurement invariance across hospitals.
Consider a single factor $\eta$. For factor indicator $r$ ($r = 1, 2, \ldots p$) for individual $i$ in group (cluster) $j$,

$$y_{rij} = \nu_{rj} + \lambda_{rj} \eta_{ij} + \varepsilon_{ij},$$

(7)

$$\eta_{ij} = \eta_j + \zeta_{ij}, (this \ may \ be \ viewed \ as \ \eta_{Bj} + \eta_{Wij})$$

(8)

$$\nu_{rj} = \nu_r + \delta_{\nu_j},$$

(9)

$$\lambda_{rj} = \lambda_r + \delta_{\lambda_j},$$

(10)

where $\nu_r$ is the mean of the $r^{th}$ intercept and $\lambda_r$ is the mean of the $r^{th}$ factor loading. Because the factor loadings are free, the factor metric is set by fixing $V(\zeta_{ij}) = 1$ (the between-level variance $V(\eta_j)$ is free). Note that the same loading is multiplying both the between- and within-level parts of the factor $\eta$. 
Two-Level Factor Analysis With Random Loadings: 3 Model Versions

\[ y_{rij} = \nu_{rj} + \lambda_{rj} \eta_{ij} + \varepsilon_{ij}, \quad (11) \]
\[ \eta_{ij} = \eta_j + \zeta_{ij}, \text{ (this may be viewed as } \eta_{Bj} + \eta_{Wij}) \quad (12) \]
\[ \nu_{rj} = \nu_r + \delta_{\nu_j}, \quad (13) \]
\[ \lambda_{rj} = \lambda_r + \delta_{\lambda_j}, \quad (14) \]

A first alternative to this model is that \( V(\eta_j) = 0 \) so that the factor with random loadings has only within-level variation. Instead, there can be a separate between-level factor with non-random loadings, measured by the random intercepts of the \( y \) indicators as in regular two-level factor analysis, \( y_{rj} = \lambda_{Br} \eta_{Bj} + \zeta_{rj} \), where \( y_{rj} \) is the between part of \( y_{rij} \).

A second alternative is that the \( \lambda_{Br} \) loadings are equal to the means of the random loadings \( \lambda_r \).
5.2.1 Group Is Random Mode. UG Ex9.19

Part 1: Random factor loadings (decomposition of the factor into within- and between-level parts)

TITLE: this is an example of a two-level MIMIC model with continuous factor indicators, random factor loadings, two covariates on within, and one covariate on between with equal loadings across levels
DATA: FILE = ex9.19.dat;
VARIABLE: NAMES = y1-y4 x1 x2 w clus;
  WITHIN = x1 x2;
  BETWEEN = w;
  CLUSTER = clus;
ANALYSIS: TYPE = TWOLEVEL RANDOM;
  ESTIMATOR = BAYES;
  PROCESSORS = 2;
  BITER = (1000);
MODEL:
  %WITHIN%
  s1-s4 | f BY y1-y4;
  f@1;
  f ON x1 x2;
  %BETWEEN%
  f ON w;
  f; ! defaults: s1-s4; [s1-s4];
PLOT: TYPE = PLOT2;
OUTPUT: TECH1 TECH8;
Part 2: Random factor loadings and a separate between-level factor

MODEL: %WITHIN%
   s1-s4 | f BY y1-y4;
f@1;
f ON x1 x2;
%BETWEEN%
   fb BY y1-y4;
   fb ON w;

f@0; is the between-level default
Part 3: Random factor loadings and a separate between-level factor with loadings equal to the mean of the random loadings

MODEL:

%WITHIN%
  s1-s4  |  f  BY  y1-y4;
  f@1;
  f  ON  x1  x2;

%BETWEEN%
  fb  BY  y1-y4*  (lam1-lam4);
  fb  ON  w;
  [s1-s4*1]  (lam1-lam4);
5.2.2 Monte Carlo Simulations For Groups As Random Mode:
Two-Level Random Loadings Modeling

- The effect of treating random loadings as fixed parameters
  - Continuous variables
  - Categorical variables
- Small number of clusters/groups
The Effect Of Treating Random Loadings As Fixed Parameters With Continuous Variables

Table: Absolute bias and coverage for factor analysis model with random loadings - comparing random intercepts and loadings and v.s. random intercepts and fixed loadings models

<table>
<thead>
<tr>
<th>parameter</th>
<th>Bayes</th>
<th>ML with fixed loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>0.00(0.97)</td>
<td>0.20(0.23)</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.01(0.95)</td>
<td>0.14(0.66)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.01(0.96)</td>
<td>0.00(0.80)</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.02(0.89)</td>
<td>0.00(0.93)</td>
</tr>
</tbody>
</table>

Ignoring the random loadings leads to biased mean and variance parameters and poor coverage. The loading is unbiased but has poor coverage.
Table: Absolute bias and coverage for factor analysis model with categorical data and random loadings - comparing random loadings and intercepts v.s. random intercepts and fixed loadings models

<table>
<thead>
<tr>
<th>parameter</th>
<th>Bayes</th>
<th>WLSMV with fixed loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>0.05(0.96)</td>
<td>0.17(0.63)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.03(0.92)</td>
<td>0.13(0.39)</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.05(0.91)</td>
<td>0.11(0.70)</td>
</tr>
</tbody>
</table>

Ignoring the random loadings leads to biased mean, loading and variance parameters and poor coverage.
Many applications have small number of clusters/groups. How many variables and random effects can we use?

- Independent random effects model - works well even with 50 variables (100 random effects) and 10 clusters
- Weakly informative priors are needed to eliminate biases for cluster level variance parameters
- Correlated random effects model (1-factor model) - works only when ”number of clusters > number of random effects”. More than 10 clusters are needed with 5 variables or more.
- What happens if you ignore the correlation: standard error underestimation, decreased accuracy in cluster specific estimates


Using BSEM with 1-factor model for the random effects and tiny priors $N(1, \sigma)$ for the loadings resolves the problem.

Fox (2010). Bayesian Item Response Modeling. Springer

Program for International Student Assessment (PISA 2003)

9,769 students across 40 countries

8 binary math items
Random Loadings In IRT

- $Y_{ijk}$ - outcome for student i, in country j and item k

$$ P(Y_{ijk} = 1) = \Phi(a_{jk} \theta_{ij} + b_{jk}) $$

$$ a_{jk} \sim N(a_k, \sigma_{a,k}), b_{jk} \sim N(b_k, \sigma_{b,k}) $$

Both discrimination ($a$) and difficulty ($b$) vary across country

- The $\theta$ ability factor is decomposed as

$$ \theta_{ij} = \theta_j + \epsilon_{ij} $$

$$ \theta_j \sim N(0, \nu), \epsilon_{ij} \sim N(0, \nu_j), \sqrt{\nu_j} \sim N(1, \sigma) $$

- The mean and variance of the ability vary across country
- For identification purposes the mean of $\sqrt{\nu_j}$ is fixed to 1, this replaces the traditional identification condition that $\nu_j = 1$
- Model preserves common measurement scale while accommodating measurement non-invariance as long as the variation in the loadings is not big
Three two-level factor models with random loadings
Testing for significance of the random loadings
Two methods for adding cluster specific factor variance in addition to the random loadings
All models can be used with continuous outcomes as well
Model 1 - without cluster specific factor variance, cluster specific discrimination, cluster specific difficulty, cluster specific factor mean

\[ P(Y_{ijk} = 1) = \Phi(a_{jk} \theta_{ij} + b_{jk}) \]

\[ a_{jk} \sim N(a_k, \sigma_{a,k}), \quad b_{jk} \sim N(b_k, \sigma_{b,k}) \]

\[ \theta_{ij} = \theta_j + \varepsilon_{ij} \]

\[ \varepsilon_{ij} \sim N(0, 1) \]

\[ \theta_j \sim N(0, \nu) \]
Note that cluster specific factor variance is confounded with cluster specific factor loadings (it is not straightforward to separate the two). Ignoring cluster specific factor variance should not lead to misfit. It just increases variation in the factor loadings which absorbs the variation in the factor variance.

Model 1 setup in Mplus: the factor $f$ is used on both levels to represent the within $\varepsilon_{ij}$ and the between $\theta_j$ part of the factor.

```plaintext
model:
  %within%
  s1-s8 | f by y1-y8; f@1;

  %between%
  f y1-y8 s1-s8;
```

All between level components are estimated as independent. Dependence can be introduced by adding factor models on the between level or covariances.
## PISA Results - Discrimination (Mean Of Random Loadings) And Difficulty

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.D.</th>
<th>One-Tailed P-Value</th>
<th>95% C.I. Lower 2.5%</th>
<th>95% C.I. Upper 2.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Means</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>0.735</td>
<td>0.037</td>
<td>0.000</td>
<td>0.666</td>
<td>0.820</td>
</tr>
<tr>
<td>S2</td>
<td>1.036</td>
<td>0.058</td>
<td>0.000</td>
<td>0.931</td>
<td>1.158</td>
</tr>
<tr>
<td>S3</td>
<td>0.631</td>
<td>0.026</td>
<td>0.000</td>
<td>0.588</td>
<td>0.687</td>
</tr>
<tr>
<td>S4</td>
<td>0.622</td>
<td>0.031</td>
<td>0.000</td>
<td>0.558</td>
<td>0.678</td>
</tr>
<tr>
<td>S5</td>
<td>0.528</td>
<td>0.032</td>
<td>0.000</td>
<td>0.466</td>
<td>0.593</td>
</tr>
<tr>
<td>S6</td>
<td>0.353</td>
<td>0.035</td>
<td>0.000</td>
<td>0.287</td>
<td>0.421</td>
</tr>
<tr>
<td>S7</td>
<td>0.607</td>
<td>0.032</td>
<td>0.000</td>
<td>0.549</td>
<td>0.677</td>
</tr>
<tr>
<td>S8</td>
<td>0.616</td>
<td>0.034</td>
<td>0.000</td>
<td>0.561</td>
<td>0.686</td>
</tr>
<tr>
<td><strong>Thresholds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1$1</td>
<td>-0.481</td>
<td>0.055</td>
<td>0.000</td>
<td>-0.581</td>
<td>-0.372</td>
</tr>
<tr>
<td>Y2$1</td>
<td>0.371</td>
<td>0.072</td>
<td>0.000</td>
<td>0.238</td>
<td>0.519</td>
</tr>
<tr>
<td>Y3$1</td>
<td>0.059</td>
<td>0.047</td>
<td>0.115</td>
<td>-0.037</td>
<td>0.145</td>
</tr>
<tr>
<td>Y4$1</td>
<td>-0.270</td>
<td>0.049</td>
<td>0.000</td>
<td>-0.368</td>
<td>-0.172</td>
</tr>
<tr>
<td>Y5$1</td>
<td>0.059</td>
<td>0.037</td>
<td>0.056</td>
<td>-0.013</td>
<td>0.125</td>
</tr>
<tr>
<td>Y6$1</td>
<td>-1.496</td>
<td>0.044</td>
<td>0.000</td>
<td>-1.584</td>
<td>-1.409</td>
</tr>
<tr>
<td>Y7$1</td>
<td>-0.691</td>
<td>0.045</td>
<td>0.000</td>
<td>-0.777</td>
<td>-0.605</td>
</tr>
<tr>
<td>Y8$1</td>
<td>-0.862</td>
<td>0.038</td>
<td>0.000</td>
<td>-0.926</td>
<td>-0.786</td>
</tr>
</tbody>
</table>
## PISA Results - Random Variation Across Countries

<table>
<thead>
<tr>
<th>Variances</th>
<th>Estimate</th>
<th>Posterior S.D.</th>
<th>One-Tailed P-Value</th>
<th>95% C.I. Lower 2.5%</th>
<th>95% C.I. Upper 2.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>0.044</td>
<td>0.017</td>
<td>0.000</td>
<td>0.022</td>
<td>0.087</td>
</tr>
<tr>
<td>Y2</td>
<td>0.063</td>
<td>0.025</td>
<td>0.000</td>
<td>0.030</td>
<td>0.124</td>
</tr>
<tr>
<td>Y3</td>
<td>0.029</td>
<td>0.011</td>
<td>0.000</td>
<td>0.013</td>
<td>0.057</td>
</tr>
<tr>
<td>Y4</td>
<td>0.034</td>
<td>0.012</td>
<td>0.000</td>
<td>0.018</td>
<td>0.066</td>
</tr>
<tr>
<td>Y5</td>
<td>0.011</td>
<td>0.006</td>
<td>0.000</td>
<td>0.003</td>
<td>0.027</td>
</tr>
<tr>
<td>Y6</td>
<td>0.034</td>
<td>0.016</td>
<td>0.000</td>
<td>0.013</td>
<td>0.075</td>
</tr>
<tr>
<td>Y7</td>
<td>0.023</td>
<td>0.011</td>
<td>0.000</td>
<td>0.008</td>
<td>0.050</td>
</tr>
<tr>
<td>Y8</td>
<td>0.007</td>
<td>0.006</td>
<td>0.000</td>
<td>0.001</td>
<td>0.023</td>
</tr>
<tr>
<td>F</td>
<td>0.317</td>
<td>0.085</td>
<td>0.000</td>
<td>0.204</td>
<td>0.528</td>
</tr>
<tr>
<td>S1</td>
<td>0.013</td>
<td>0.010</td>
<td>0.000</td>
<td>0.002</td>
<td>0.037</td>
</tr>
<tr>
<td>S2</td>
<td>0.078</td>
<td>0.040</td>
<td>0.000</td>
<td>0.027</td>
<td>0.181</td>
</tr>
<tr>
<td>S3</td>
<td>0.006</td>
<td>0.005</td>
<td>0.000</td>
<td>0.001</td>
<td>0.022</td>
</tr>
<tr>
<td>S4</td>
<td>0.012</td>
<td>0.008</td>
<td>0.000</td>
<td>0.002</td>
<td>0.031</td>
</tr>
<tr>
<td>S5</td>
<td>0.017</td>
<td>0.010</td>
<td>0.000</td>
<td>0.005</td>
<td>0.044</td>
</tr>
<tr>
<td>S6</td>
<td>0.020</td>
<td>0.012</td>
<td>0.000</td>
<td>0.005</td>
<td>0.048</td>
</tr>
<tr>
<td>S7</td>
<td>0.006</td>
<td>0.007</td>
<td>0.000</td>
<td>0.001</td>
<td>0.027</td>
</tr>
<tr>
<td>S8</td>
<td>0.011</td>
<td>0.008</td>
<td>0.000</td>
<td>0.002</td>
<td>0.033</td>
</tr>
</tbody>
</table>
Factor scores can be obtained for the mean ability parameter using the country specific factor loadings. Highest and lowest 3 countries.

<table>
<thead>
<tr>
<th>Country</th>
<th>Estimate and confidence limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIN</td>
<td>0.749 ( 0.384 , 0.954 )</td>
</tr>
<tr>
<td>KOR</td>
<td>0.672 ( 0.360 , 0.863 )</td>
</tr>
<tr>
<td>MAC</td>
<td>0.616 ( 0.267 , 1.041 )</td>
</tr>
<tr>
<td>BRA</td>
<td>-0.917 ( -1.166 , -0.701 )</td>
</tr>
<tr>
<td>IDN</td>
<td>-1.114 ( -1.477 , -0.912 )</td>
</tr>
<tr>
<td>TUN</td>
<td>-1.156 ( -1.533 , -0.971 )</td>
</tr>
</tbody>
</table>
Country-Specific Distribution
For The Mean Ability Parameter For FIN
Random loadings have small variances, however even small variance of 0.01 implies a range for the loading of $4\times SD=0.4$, i.e., substantial variation in the loadings across countries.

How can we test significance for the variance components? If variance is not near zero the confidence intervals are reliable. However, when the variance is near 0 the confidence interval does not provide evidence for statistical significance.

Example: $\text{Var}(S2)=0.078$ with confidence interval $[0.027,0.181]$ is significant but $\text{Var}(S7)=0.006$ with confidence interval $[0.001,0.027]$ is not clear. Caution: if the number of clusters on the between level is small all these estimates will be sensitive to the prior.
Verhagen & Fox (2012) Bayesian Tests of Measurement Invariance

Test the null hypothesis $\sigma = 0$ using Bayesian methodology

Substitute null hypothesis $\sigma < 0.001$

Estimate the model with $\sigma$ prior $\text{IG}(1,0.005)$ with mode 0.0025 (If we push the variances to zero with the prior, would the data provide any resistance?)

\[
BF = \frac{P(H_0)}{P(H_1)} = \frac{P(\sigma < 0.001|\text{data})}{P(\sigma < 0.001)} = \frac{P(\sigma < 0.001|\text{data})}{0.7%}
\]

$BF > 3$ indicates loading has 0 variance, i.e., loading invariance
Other cutoff values are possible such as 0.0001 or 0.01

Implemented in Mplus in Tech16

Estimation should be done in two steps. First estimate a model with non-informative priors. Second in a second run estimate the model with IG(1,0.005) variance prior to test the significance

How well does this work? The problem of testing for zero variance components is difficult. ML T-test or LRT doesn’t provide good solution because it is a borderline testing

New method which is not studied well but there is no alternative particularly for the case of random loadings. The random loading model can not be estimated with ML due to too many dimensions of numerical integration
Testing For Non-Zero Variance Of Random Loadings

- Simulation: Simple factor analysis model with 5 indicators, N=2000, variance of factor is free, first loading fixed to 1. Simulate data with Var(f)=0.0000001. Using different BITER commands with different number of min iterations
  - BITER=100000; rejects the non-zero variance hypothesis 51% of the time
  - BITER=100000(5000); rejects the non-zero variance hypothesis 95% of the time
  - BITER=100000(10000); rejects the non-zero variance hypothesis 100% of the time
- Conclusion: The variance component test needs good number of iterations due to estimation of tail probabilities
- Power: if we generate data with Var(f)=0.05, the power to detect significantly non-zero variance component is 50% comparable to ML T-test of 44%
Add IG(1,0.005) prior for the variances we want to test

MODEL:

\[
\begin{align*}
\text{%WITHIN}\% & \\
& s1-s8 \mid f \text{ BY } y1-y8; \\
& f@1; \\
& \text{%BETWEEN}\% \\
& f; \\
& y1-y8 (v1-v8); \\
& s1-s8 (v9-v16); \\
\end{align*}
\]

MODEL PRIORS:

v1-v16 \sim IG(1, 0.005);

OUTPUT:

TECH1 TECH16;
Bayes factor greater than 3 in any column indicate non-significance (at the corresponding level). For example, Bayes factor greater than 3 in the second column indicates variance is less than 0.001.

Bayes factor=10 in column 3 means that a model with variance smaller than 0.001 is 10 times more likely than a model with non-zero variance.

The small variance prior that is used applies to a particular variance threshold hypothesis. For example, if you want to test the hypothesis $v < 0.001$, use the prior $v \sim IG(1,0.005)$, and look for the results in the second column. If you want to test the hypothesis $v < 0.01$, use the prior $v \sim IG(1,0.05)$, and look for the results in the third column.

Parameters 9-16 variances of the difficulty parameters
Parameters 26-33 variances of the discrimination parameters
Testing significance of variance components

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BF for &lt;0.0001</th>
<th>BF for &lt;0.001</th>
<th>BF for &lt;0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter 9</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Parameter 10</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Parameter 11</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0115</td>
</tr>
<tr>
<td>Parameter 12</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0041</td>
</tr>
<tr>
<td>Parameter 13</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.5431</td>
</tr>
<tr>
<td>Parameter 14</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0324</td>
</tr>
<tr>
<td>Parameter 15</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0417</td>
</tr>
<tr>
<td>Parameter 16</td>
<td>0.0000</td>
<td>0.7093</td>
<td>1.3432</td>
</tr>
<tr>
<td>Parameter 26</td>
<td>0.0000</td>
<td>2.4226</td>
<td>1.5264</td>
</tr>
<tr>
<td>Parameter 27</td>
<td>0.0000</td>
<td>0.0982</td>
<td>0.7791</td>
</tr>
<tr>
<td>Parameter 28</td>
<td>0.0000</td>
<td>4.2996</td>
<td>1.5951</td>
</tr>
<tr>
<td>Parameter 29</td>
<td>0.0000</td>
<td>1.0476</td>
<td>1.4318</td>
</tr>
<tr>
<td>Parameter 30</td>
<td>0.0000</td>
<td>0.4911</td>
<td>1.2564</td>
</tr>
<tr>
<td>Parameter 31</td>
<td>0.0000</td>
<td>0.3929</td>
<td>0.9429</td>
</tr>
<tr>
<td>Parameter 32</td>
<td>0.0000</td>
<td>2.4117</td>
<td>1.5428</td>
</tr>
<tr>
<td>Parameter 33</td>
<td>0.0000</td>
<td>1.1895</td>
<td>1.4523</td>
</tr>
</tbody>
</table>
Estimate a model with fixed and random loadings. Loading 3 is now a fixed parameter rather than random.

MODEL:

%WITHIN%
f @ 1;
s1-s2 | f BY y1-y2;
f BY y3*1;
s4-s8 | f BY y4-y8;

%BETWEEN%
f;
y1-y8;
s1-s8;
Model 2 - Between level factor has different (non-random) loadings

\[ P(Y_{ijk} = 1) = \Phi(a_{jk} \theta_{ij} + c_k \theta_j + b_{jk}) \]

\[ a_{jk} \sim N(a_k, \sigma_{a,k}), \quad b_{jk} \sim N(b_k, \sigma_{b,k}) \]

\[ \theta_{ij} \sim N(0, 1) \]

\[ \theta_j \sim N(0, 1) \]

Model 2 doesn’t have the interpretation that \( \theta_j \) is the between part of the \( \theta_{ij} \) since the loadings are different
Model 3 - Between level factor has loadings equal to the mean of the random loadings

\[ P(Y_{ijk} = 1) = \Phi(a_{jk} \theta_{ij} + a_k \theta_j + b_{jk}) \]

\[ a_{jk} \sim N(a_k, \sigma_{a,k}), b_{jk} \sim N(b_k, \sigma_{b,k}) \]

\[ \theta_{ij} \sim N(0, 1) \]

\[ \theta_j \sim N(0, \nu) \]

Model 3 has the interpretation that \( \theta_j \) is approximately the between part of the \( \theta_{ij} \)

Model 3 is nested within Model 2 and can be tested by testing the proportionality of between and within loadings
Model 3 setup. The within factor $f$ now represents only $\theta_{ij}$, $fb$ represents $\theta_j$.

```
model:

%within%
sl-s8 | f by y1-y8;
f@1;

%between%
y1-y8 sl-s8;
[s1-s8*1] (p1-p8);
fb by y1-y8*1 (p1-p8);
fb;
```
Replace $\text{Var}(\theta_{ij}) = 1$ with $\text{Var}(\theta_{ij}) = 0.51 + (0.7 + \sigma_j)^2$ where $\sigma_j$ is a zero mean cluster level random effect. The constant 0.51 is needed to avoid variances fixed to 0 which cause poor mixing. This approach can be used for any variance component on the within level.

```
model:

%within%
s1-s8 | f by y1-y8;
sigma | e by f;
e@1;
f@0.51;

%between%
y1-y8 s1-s8;
[s1-s8*1] (p1-p8);
fb by y1-y8*1 (p1-p8);
fb sigma;
[sigma@0.7];
```
Variability in the loadings is confounded with variability in the factor variance

A model is needed that can naturally separate the across-country variation in the factor loadings and the across-country variation in the factor variance

From a practical perspective we want to have as much variation in the factor variance and as little as possible in the factor loadings to pursue the concept of measurement invariance or approximate measurement invariance.
Replace $\text{Var}(\theta_{ij}) = 1$ with $\text{Var}(\theta_{ij}) = (1 + \sigma_j)^2$ where $\sigma_j$ is a zero mean cluster level random effect. This model is equivalent to having $\text{Var}(\theta_{ij}) = 1$ and the discrimination parameters as

$$ a_{jk} = (1 + \sigma_j)(a_k + \varepsilon_{jk}) $$

Because $\sigma_j$ and $\varepsilon_{jk}$ are generally small, the product $\sigma_j \cdot \varepsilon_{jk}$ is of smaller magnitude so it is ignored

$$ a_{jk} \approx a_k + \varepsilon_{jk} + a_k \sigma_j $$

$\sigma_j$ can be interpreted as **between level latent factor for the random loadings** with loadings $a_k$ equal to the means of the random loadings
Factor analysis estimation tends to absorb most of the correlation between the indicators within the factor model and to minimize the residual variances.

Thus the model will try to explain as much as possible the variation between the correlation matrices across individual as a variation in the factor variance rather than as a variation in the factor loadings.

Thus this model is ideal for evaluating and separating the loading non-invariance and the factor variance non-invariance.

Testing $\text{Var}(\varepsilon_{jk}) = 0$ is essentially a test for measurement invariance. Testing $\text{Var}(\sigma_j) = 0$ is essentially a test for factor variance invariance across the cluster.
Method 2 setup. Optimal in terms of mixing and convergence.

MODEL:

\[
\text{%WITHIN%}
\begin{align*}
    & s1-s8 \mid f \text{ BY } y1-y8; \\
    & f@1; \\
\end{align*}
\text{%BETWEEN%}
\begin{align*}
    & y1-y8 \text{ s1-s8; } \\
    & [s1-s8*1] (p1-p8); \\
    & fb \text{ BY } y1-y8*1 (p1-p8); \\
    & sigma \text{ BY } s1-s8*1 (p1-p8); \\
    & fb \text{ sigma; }
\end{align*}
\]

- An example of the growing amount of EMA data
- 84 borderline personality disorder (BPD) patients. The mood factor for each individual is measured with 21 self-rated continuous items. Each individual is measured several times a day for 4 weeks for total of about 100 assessments
- Factor analysis is done as a two-level model where cluster=individual, many assessments per cluster
This data set is perfect to check if a measurement instrument is interpreted the same way by different individuals. Some individuals' responses may be more correlated for some items, i.e., the factor analysis should be different for different individuals.

Example: suppose that one individual answers item 1 and 2 always the same way and a second individual doesn’t. We need separate factor analysis models for the two individuals, individually specific factor loadings.

If the within level correlation matrix varies across cluster that means that the loadings are individually specific.

Should in general factors loadings be individually specific? This analysis can NOT be done in cross-sectional studies, only longitudinal studies with multiple assessments.
Large across-time variance of the mood factor is considered a core feature of BPD that distinguishes this disorder from other disorders like depressive disorders.

The individual-specific factor variance is the most important feature in this study.

The individual-specific factor variance is confounded with individual-specific factor loadings.

How to separate the two? Answer: **Factor Model for the Random Factor Loadings** as in the PISA data.
Let $Y_{pij}$ be item $p$, for individual $i$, at assessment $j$. Let $X_i$ be an individual covariate. The model is given by

\[
Y_{pij} = \mu_p + \zeta_{pi} + s_{pi} \eta_{ij} + \epsilon_{pij}
\]

\[
\eta_{ij} = \eta_{i} + \beta_1 X_i + \xi_{ij}
\]

\[
s_{pi} = \lambda_p + \lambda_p \sigma_{i} + \epsilon_{pi}
\]

\[
\sigma_{i} = \beta_2 X_i + \zeta_{i}
\]

$\beta_1$ and $\beta_2$ represent the effect of the covariate $X$ on the mean and the variance of the mood factor.

IDFA has individually specific: item mean, item loading, factor mean, factor variance.
Many different ways to set up this model in Mplus. The setup below gives the best mixing/convergence performance.

MODEL:

\%WITHIN\%

s1-s21 | f BY jittery-scornful;
f@1;

\%BETWEEN\%

f ON x; f;
s1-s21 jittery-scornful;
[s1-s21*1] (lambda1-lambda21);
sigma BY s1-s21*1 (lambda1-lambda21);
sigma ON x; sigma;
Individual Differences Factor Analysis Results

All variance components are significant. Percent Loading Invariance = the percentage of the variation of the loadings that is explained by factor variance variation.

<table>
<thead>
<tr>
<th>item</th>
<th>Res Var</th>
<th>Mean</th>
<th>Var of Mean</th>
<th>Loading</th>
<th>Var of Loading</th>
<th>Percent Loading Invariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>0.444</td>
<td>1.505</td>
<td>0.287</td>
<td>0.261</td>
<td>0.045</td>
<td>0.29</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.628</td>
<td>1.524</td>
<td>0.482</td>
<td>0.377</td>
<td>0.080</td>
<td>0.32</td>
</tr>
<tr>
<td>Item 3</td>
<td>0.331</td>
<td>1.209</td>
<td>0.057</td>
<td>0.556</td>
<td>0.025</td>
<td>0.77</td>
</tr>
<tr>
<td>Item 4</td>
<td>0.343</td>
<td>1.301</td>
<td>0.097</td>
<td>0.553</td>
<td>0.030</td>
<td>0.73</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.304</td>
<td>1.094</td>
<td>0.017</td>
<td>0.483</td>
<td>0.053</td>
<td>0.54</td>
</tr>
</tbody>
</table>
Clear evidence that measurement items are not interpreted the same way by different individuals and thus individual-specific adjustments are needed to the measurement model to properly evaluate the underlying factors: IDFA model

IDFA model clearly separates factor variance variation from the factor loadings variation

6. 3-Level Analysis

Continuous outcomes: ML and Bayesian estimation

Categorical outcomes: Bayesian estimation (Bayes uses probit)

Count and nominal outcomes: Not yet available
Each $Y$ variable is decomposed as

$$Y_{ijk} = Y_{1ijk} + Y_{2jk} + Y_{3k},$$

where $Y_{1ijk}$, $Y_{2jk}$, and $Y_{3k}$ are components of $Y_{ijk}$ on levels 1, 2, and 3. Here, $Y_{2jk}$, and $Y_{3k}$ may be seen as random intercepts on respective levels, and $Y_{1ijk}$ as a residual.

- Some variables may not have variation over all levels. To avoid variances that are near zero which cause convergence problems specify/restrict the variation level.
- `WITHIN=Y`, has variation on level 1, so $Y_{2jk}$ and $Y_{3k}$ are not in the model.
- `WITHIN=(level2) Y`, has variation on level 1 and level 2.
- `WITHIN=(level3) Y`, has variation on level 1 and level 3.
- `BETWEEN= Y`, has variation on level 2 and level 3.
- `BETWEEN=(level2) Y`, has variation on level 2.
- `BETWEEN=(level3) Y`, has variation on level 3.
Type 1: Defined on the level 1
%WITHIN%
s | y ON x;
The random slope $s$ has variance on level 2 and level 3

Type 2: Defined on the level 2
%BETWEEN level2%
s | y ON x;
The random slope $s$ has variance on level 3 only

The dependent variable can be an observed $Y$ or a factor. The covariate $X$ should be specified as WITHIN= for type 1 or BETWEEN=(level2) for type 2, i.e., no variation beyond the level it is used at
6.2 3-Level Regression

\[ Level 1 : y_{ijk} = \beta_{0jk} + \beta_{1jk} x_{ijk} + \epsilon_{ijk}, \quad (15) \]

\[ Level 2a : \beta_{0jk} = \gamma_{00k} + \gamma_{01k} w_{jk} + \zeta_{0jk}, \quad (16) \]

\[ Level 2b : \beta_{1jk} = \gamma_{10k} + \gamma_{11k} w_{jk} + \zeta_{1jk}, \quad (17) \]

\[ Level 3a : \gamma_{00k} = \kappa_{000} + \kappa_{001} z_{k} + \delta_{00k}, \quad (18) \]

\[ Level 3b : \gamma_{01k} = \kappa_{010} + \kappa_{011} z_{k} + \delta_{01k}, \quad (19) \]

\[ Level 3c : \gamma_{10k} = \kappa_{100} + \kappa_{101} z_{k} + \delta_{10k}, \quad (20) \]

\[ Level 3d : \gamma_{11k} = \kappa_{110} + \kappa_{111} z_{k} + \delta_{11k}, \quad (21) \]

where

- \( x, w, \) and \( z \) are covariates on the different levels
- \( \beta \) are level 2 random effects
- \( \gamma \) are level 3 random effects
- \( \kappa \) are fixed effects
- \( \epsilon, \zeta \) and \( \delta \) are residuals on the different levels
3-Level Regression Example: UG Example 9.20

Within

| x | s1 | y |

Between 2

| w | s2 | y |
|   | s12 | s1 |

Between 3

| z | y |
|   | s1 |
|   | s2 |
|   | s12 |
### TITLE:
this is an example of a three-level regression with a continuous dependent variable

### DATA:
FILE = ex9.20.dat;

### VARIABLE:
NAMES = y x w z level2 level3;

### CLUSTER:
level3 level2;

### WITHIN:
x;

### BETWEEN:
(level2) w (level3) z;

### ANALYSIS:
TYPE = THREELEVEL RANDOM;

### MODEL:

%WITHIN%
s1 | y ON x;

%BETWEEN level2%
s2 | y ON w;
s12 | s1 ON w;
y WITH s1;

%BETWEEN level3%
y ON z;
s1 ON z;
s2 ON z;
s12 ON z;
y WITH s1 s2 s12;
s1 WITH s2 s12;
s2 WITH s12;

### OUTPUT:
TECH1 TECH8;
6.3 3-Level Regression: Nurses Data

Source: Hox (2010). Multilevel Analysis. Hypothetical data discussed in Section 2.4.3

- Study of stress in hospitals
- Reports from nurses working in wards nested within hospitals
- In each of 25 hospitals, 4 wards are selected and randomly assigned to experimental or control conditions (cluster-randomized trial)
- 10 nurses from each ward are given a test that measures job-related stress
- Covariates are age, experience, gender, type of ward (0=general care, 1=special care), hospital size (0=small, 1=medium, 2=large)
- Research question: Is the experimental effect different in different hospitals? - Random slope varying on level 3
3-Level Regression Example: Nurses Data

- **Within**
  - age
  - gender
  - exper
  - stress

- **Between Ward**
  - expcon
  - ward
  - stress

- **Between Hospital**
  - hosp size
  - stress
  - s
TITLE: Nurses data from Hox (2010)
DATA: FILE = nurses.dat;
VARIABLE: NAMES = hospital ward wardid nurse age gender experience stress wardtype hospsize expcon zage zgender zexperience zstress zwardtyi zhospsize zexpcon cexpcon chospsize;
CLUSTER = hospital wardid;
WITHIN = age gender experience;
BETWEEN = (hospital) hospsize (wardid) expcon wardtype;
USEVARIABLES = stress expcon age gender experience wardtype hospsize;
CENTERING = GRANDMEAN(expcon hospsize);
ANALYSIS: TYPE = THREELEVEL RANDOM;
ESTIMATOR = MLR;
MODEL:

%WITHIN%
stress ON age gender experience;
%BETWEEN wardid%
s | stress ON expcon;
stress ON wardtype;
%BETWEEN hospital%
s stress ON hospsize;
s; s WITH stress;

OUTPUT: TECH1 TECH8;
SAVEDATA: SAVE = FSCORES;
FILE = fs.dat;
PLOT: TYPE = PLOT2 PLOT3;
<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>Two-Tailed P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WITHIN Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>0.022</td>
<td>0.002</td>
<td>11.911</td>
<td>0.000</td>
</tr>
<tr>
<td>gender</td>
<td>-0.455</td>
<td>0.032</td>
<td>-14.413</td>
<td>0.000</td>
</tr>
<tr>
<td>experience</td>
<td>-0.062</td>
<td>0.004</td>
<td>-15.279</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress</td>
<td>0.217</td>
<td>0.011</td>
<td>20.096</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>BETWEEN wardid Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wardtype</td>
<td>0.053</td>
<td>0.076</td>
<td>0.695</td>
<td>0.487</td>
</tr>
<tr>
<td></td>
<td>Estimates</td>
<td>S.E.</td>
<td>Est./S.E.</td>
<td>Two-Tailed P-Value</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------</td>
<td>------</td>
<td>-----------</td>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress</td>
<td>0.109</td>
<td>0.033</td>
<td>3.298</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>BETWEEN hospital Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hosphsize</td>
<td>0.998</td>
<td>0.191</td>
<td>5.217</td>
<td>0.000</td>
</tr>
<tr>
<td>stress ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hosphsize</td>
<td>-0.041</td>
<td>0.152</td>
<td>-0.270</td>
<td>0.787</td>
</tr>
<tr>
<td>s WITH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress</td>
<td>-0.036</td>
<td>0.058</td>
<td>-0.615</td>
<td>0.538</td>
</tr>
<tr>
<td></td>
<td>Estimates</td>
<td>S.E.</td>
<td>Est./S.E.</td>
<td>Two-Tailed P-Value</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>-------</td>
<td>-----------</td>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Intercepts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress</td>
<td>5.753</td>
<td>0.102</td>
<td>56.171</td>
<td>0.000</td>
</tr>
<tr>
<td>s</td>
<td>-0.699</td>
<td>0.111</td>
<td>-6.295</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stress</td>
<td>0.143</td>
<td>0.051</td>
<td>2.813</td>
<td>0.005</td>
</tr>
<tr>
<td>s</td>
<td>0.178</td>
<td>0.087</td>
<td>2.060</td>
<td>0.039</td>
</tr>
</tbody>
</table>
6.4 3-Level Path Analysis: UG Example 9.21

Within

Between 2

Between 3
### 3-Level Path Analysis: UG Ex 9.21 Input

**TITLE:** this an example of a three-level path analysis with a continuous and a categorical dependent variable

**DATA:** FILE = ex9.21.dat;

**VARIABLE:** NAMES = u y2 y3 x w z level2 level3;

- CATEGORICAL = u;
- CLUSTER = level3 level2;
- WITHIN = x;
- BETWEEN = y2 (level2) w (level3) z y3;

**ANALYSIS:** TYPE = THREELEVEL;
- ESTIMATOR = BAYES;
- PROCESSORS = 2;
- BITERATIONS = (1000);

**MODEL:**

- **%WITHIN%
  - u ON y x;
  - y ON x;
- **%BETWEEN level2%
  - u ON w y y2;
  - y ON w;
  - y2 ON w;
  - y WITH y2;
- **%BETWEEN level3%
  - u ON y y2;
  - y ON z;
  - y2 ON z;
  - y3 ON y y2;
  - y WITH y2;
  - u WITH y3;

**OUTPUT:** TECH1 TECH8;
6.5 3-Level MIMIC Analysis

Within

Between 2

Bengt Muthén & Tihomir Asparouhov  Mplus Modeling  95/ 186
3-Level MIMIC Analysis, Continued
TITLE: this is an example of a three-level MIMIC model with continuous factor indicators, two covariates on within, one covariate on between level 2, one covariate on between level 3 with random slopes on both within and between level 2

DATA: FILE = ex9.22.dat;
VARIABLE: NAMES = y1-y6 x1 x2 w z level2 level3;
CLUSTER = level3 level2;
WITHIN = x1 x2;
BETWEEN = (level2) w (level3) z;

ANALYSIS: TYPE = THREELEVEL RANDOM;
MODEL:
%WITHIN%
fw1 BY y1-y3;
fw2 BY y4-y6;
fw1 ON x1;
s | fw2 ON x2;
%BETWEEN level2%
fb2 BY y1-y6;
sf2 | fb2 ON w;
ss | s ON w;
fb2 WITH s;
%BETWEEN level3%
fb3 BY y1-y6;
fb3 ON z;
s ON z;
sf2 ON z;
ss ON z;
fb3 WITH s sf2 ss;
s WITH sf2 ss;
sf2 WITH ss;

OUTPUT: TECH1 TECH8;
3-Level MIMIC Analysis, Monte Carlo Input: 
5 Students (14 Parameters) In 30 Classrooms (13 Parameters) 
In 50 Schools (28 Parameters)

MONTECARLO:

 NAMES = y1-y6 x1 x2 w z;
 NOBSERVATIONS = 7500;
 NREPS = 500;
 CSIZES = 50[30(5)];
 NCSIZE = 1[1];
 !SAVE = ex9.22.dat;
 WITHIN = x1 x2;
 BETWEEN = (level2) w (level3) z;

ANALYSIS:

 TYPE = THREELEVEL RANDOM;
 ESTIMATOR = MLR;
REPLICATION 499:
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 -0.239D-16. PROBLEM INVOLVING PARAMETER 51.

THE NONIDENTIFICATION IS MOST LIKELY DUE TO HAVING MORE PARAMETERS THAN THE
NUMBER OF LEVEL 3 CLUSTERS. REDUCE THE NUMBER OF PARAMETERS.

REPLICATION 500:
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 -0.190D-16. PROBLEM INVOLVING PARAMETER 52.

THE NONIDENTIFICATION IS MOST LIKELY DUE TO HAVING MORE PARAMETERS THAN THE
NUMBER OF LEVEL 3 CLUSTERS. REDUCE THE NUMBER OF PARAMETERS.
<table>
<thead>
<tr>
<th>Population</th>
<th>ESTIMATES</th>
<th>S. E.</th>
<th>M. S. E.</th>
<th>95%</th>
<th>% Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between LEVEL2 Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FB2 BY**

| Y1  | 1.000 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 1.000 | 0.000 |
| Y2  | 1.000 | 0.9980 | 0.0236 | 0.0237 | 0.0006 | 0.952 | 1.000 |
| Y3  | 1.000 | 0.9999 | 0.0237 | 0.0239 | 0.0006 | 0.940 | 1.000 |
| Y4  | 1.000 | 0.9987 | 0.0271 | 0.0272 | 0.0007 | 0.936 | 1.000 |
| Y5  | 1.000 | 1.0005 | 0.0265 | 0.0270 | 0.0007 | 0.948 | 1.000 |
| Y6  | 1.000 | 0.9987 | 0.0277 | 0.0269 | 0.0008 | 0.944 | 1.000 |

**FB2 WITH**

| S   | 0.000 | 0.0001 | 0.0238 | 0.0222 | 0.0006 | 0.940 | 0.060 |

**Residual Variances**

| Y1  | 0.500 | 0.5009 | 0.0343 | 0.0338 | 0.0012 | 0.940 | 1.000 |
| Y2  | 0.500 | 0.4988 | 0.0345 | 0.0338 | 0.0012 | 0.928 | 1.000 |
| Y3  | 0.500 | 0.5004 | 0.0347 | 0.0336 | 0.0012 | 0.936 | 1.000 |
| Y4  | 0.500 | 0.4995 | 0.0333 | 0.0339 | 0.0011 | 0.950 | 1.000 |
| Y5  | 0.500 | 0.4988 | 0.0337 | 0.0337 | 0.0011 | 0.946 | 1.000 |
| Y6  | 0.500 | 0.5002 | 0.0350 | 0.0339 | 0.0012 | 0.932 | 1.000 |
| FB2 | 0.500 | 0.5021 | 0.0327 | 0.0321 | 0.0011 | 0.934 | 1.000 |
| S   | 0.600 | 0.6018 | 0.0384 | 0.0374 | 0.0015 | 0.938 | 1.000 |
## 3-Level MIMIC Analysis, Monte Carlo Output, Continued

### Between LEVEL3 Level

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FB3</td>
<td>BY</td>
<td>Y1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.000</td>
<td>1.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.000</td>
<td>1.0112</td>
<td>0.1396</td>
<td>0.1372</td>
<td>0.0196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.000</td>
<td>1.0091</td>
<td>0.1608</td>
<td>0.1403</td>
<td>0.0259</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.000</td>
<td>1.0063</td>
<td>0.1491</td>
<td>0.1398</td>
<td>0.0222</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.000</td>
<td>1.0094</td>
<td>0.1532</td>
<td>0.1420</td>
<td>0.0235</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.000</td>
<td>1.0155</td>
<td>0.1585</td>
<td>0.1418</td>
<td>0.0253</td>
</tr>
</tbody>
</table>

| FB3    | ON  | Z   |     |     |     |     |
|        |     | 0.500 | 0.5053 | 0.1055 | 0.0932 | 0.0111 | 0.9060 | 1.0000 |

| S      | ON  | Z   |     |     |     |     |
|        |     | 0.300 | 0.2947 | 0.0859 | 0.0791 | 0.0074 | 0.9120 | 0.9400 |

| SF2    | ON  | Z   |     |     |     |     |
|        |     | 0.200 | 0.1988 | 0.0834 | 0.0794 | 0.0069 | 0.9220 | 0.7040 |

| SS     | ON  | Z   |     |     |     |     |
|        |     | 0.300 | 0.3016 | 0.0863 | 0.0790 | 0.0074 | 0.9180 | 0.9380 |

| FB3    | WITH | S   |     |     |     |     |
|        |      | 0.000 | 0.0018 | 0.0501 | 0.0466 | 0.0025 | 0.9400 | 0.0600 |

| S      | WITH | SF2 |     |     |     |     |
|        |      | 0.000 | 0.0050 | 0.0499 | 0.0462 | 0.0025 | 0.9440 | 0.0560 |

| SF2    | WITH | SS  |     |     |     |     |
|        |      | 0.000 | 0.0008 | 0.0487 | 0.0466 | 0.0024 | 0.9320 | 0.0680 |

| S      | WITH | SS  |     |     |     |     |
|        |      | 0.000 | -0.0025 | 0.0448 | 0.0438 | 0.0020 | 0.9440 | 0.0560 |

<p>| SF2    | WITH | SS  |     |     |     |     |
|        |      | 0.000 | -0.0008 | 0.0471 | 0.0440 | 0.0022 | 0.9400 | 0.0600 |</p>
<table>
<thead>
<tr>
<th></th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
<th>Y5</th>
<th>Y6</th>
<th>S</th>
<th>SF2</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercepts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>0.500</td>
<td>0.4945</td>
<td>0.0995</td>
<td>0.1031</td>
<td>0.0099</td>
<td>0.966</td>
<td>0.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y2</td>
<td>0.500</td>
<td>0.4924</td>
<td>0.1035</td>
<td>0.1031</td>
<td>0.0108</td>
<td>0.932</td>
<td>0.992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y3</td>
<td>0.500</td>
<td>0.4920</td>
<td>0.1051</td>
<td>0.1029</td>
<td>0.0111</td>
<td>0.942</td>
<td>0.998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y4</td>
<td>0.500</td>
<td>0.4967</td>
<td>0.1059</td>
<td>0.1034</td>
<td>0.0112</td>
<td>0.940</td>
<td>0.998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y5</td>
<td>0.500</td>
<td>0.4974</td>
<td>0.0996</td>
<td>0.1029</td>
<td>0.0099</td>
<td>0.946</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y6</td>
<td>0.500</td>
<td>0.4975</td>
<td>0.1011</td>
<td>0.1033</td>
<td>0.0102</td>
<td>0.950</td>
<td>0.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.200</td>
<td>0.1977</td>
<td>0.0837</td>
<td>0.0809</td>
<td>0.0070</td>
<td>0.926</td>
<td>0.664</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF2</td>
<td>1.000</td>
<td>1.0051</td>
<td>0.0867</td>
<td>0.0814</td>
<td>0.0075</td>
<td>0.934</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>0.500</td>
<td>0.5042</td>
<td>0.0853</td>
<td>0.0808</td>
<td>0.0073</td>
<td>0.944</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Residual Variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>0.200</td>
<td>0.1906</td>
<td>0.0556</td>
<td>0.0506</td>
<td>0.0032</td>
<td>0.872</td>
<td>0.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y2</td>
<td>0.200</td>
<td>0.1893</td>
<td>0.0554</td>
<td>0.0499</td>
<td>0.0032</td>
<td>0.884</td>
<td>0.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y3</td>
<td>0.200</td>
<td>0.1922</td>
<td>0.0545</td>
<td>0.0504</td>
<td>0.0030</td>
<td>0.892</td>
<td>0.994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y4</td>
<td>0.200</td>
<td>0.1928</td>
<td>0.0597</td>
<td>0.0502</td>
<td>0.0036</td>
<td>0.868</td>
<td>0.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y5</td>
<td>0.200</td>
<td>0.1911</td>
<td>0.0550</td>
<td>0.0507</td>
<td>0.0031</td>
<td>0.872</td>
<td>0.998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y6</td>
<td>0.200</td>
<td>0.1907</td>
<td>0.0517</td>
<td>0.0504</td>
<td>0.0028</td>
<td>0.906</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB3</td>
<td>0.300</td>
<td>0.2899</td>
<td>0.0901</td>
<td>0.0842</td>
<td>0.0082</td>
<td>0.892</td>
<td>0.992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.300</td>
<td>0.2885</td>
<td>0.0639</td>
<td>0.0622</td>
<td>0.0042</td>
<td>0.906</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF2</td>
<td>0.300</td>
<td>0.2905</td>
<td>0.0656</td>
<td>0.0619</td>
<td>0.0044</td>
<td>0.888</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>0.300</td>
<td>0.2850</td>
<td>0.0673</td>
<td>0.0622</td>
<td>0.0047</td>
<td>0.870</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.6 3-Level Growth Analysis
### TITLE:
this is an example of a three-level growth model with a continuous outcome and one covariate on each of the three levels

### DATA:
FILE = ex9.23.dat;

### VARIABLE:
NAMES = y1-y4 x w z level2 level3;

### CLUSTER:
level3 level2;

### WITHIN:
x;

### BETWEEN:
(level2) w (level3) z;

### ANALYSIS:
TYPE = THREELEVEL;

### MODEL:

**Within Level (level1):**

```plaintext
iw sw | y1@0 y2@1 y3@2 y4@3;
iw sw ON x;
iw sw ON x;
```

**Between Level 2 (level2):**

```plaintext
ib2 sb2 | y1@0 y2@1 y3@2 y4@3;
ib2 sb2 ON w;
```

**Between Level 3 (level3):**

```plaintext
ib3 sb3 | y1@0 y2@1 y3@2 y4@3;
ib3 sb3 ON z;
```

### OUTPUT:
TECH1 TECH8;

Available with ESTIMATOR=MLR when all dependent variables are continuous.

Cluster sampling: CLUSTER=cluster4 cluster3 cluster2; For example, cluster=district school classroom;

cluster4 nested above cluster3 nested above cluster2

cluster4 provides information about cluster sampling of level 3 units, cluster3 is modeled as level 3, cluster2 is modeled as level 2

cluster4 affects only the standard errors and not the point estimates, adjusts the standard error upwards for non-independence of level 3 units
Other sampling features: Stratification (nested above cluster 4, 5 levels total), finite population sampling and weights

Three weight variables for unequal probability of selection

weight=w1; bweight=w2; b2weight=w3;

\[ w_3 = \frac{1}{P(\text{level 3 unit is selected})} \]

\[ w_2 = \frac{1}{P(\text{level 2 unit is selected} \mid \text{the level 3 unit is selected})} \]

\[ w_1 = \frac{1}{P(\text{level 1 unit is selected} \mid \text{the level 2 unit is selected})} \]

Weights are scaled to sample size at the corresponding level

Other scaling methods possible:
New Multiple Imputation Methods

- Multiple imputations for three-level and cross-classified data
- Continuous and categorical variables
- H0 imputations. Estimate a three-level or cross-classified model with the Bayes estimator. Not available as H1 imputation where the imputation model is setup as unrestricted model.
- The imputation model can be an unrestricted model or a restricted model. Restricted models will be easier to estimate especially when the number of clustering units is not large
- In the input file simply add the DATA IMPUTATION command
variable:
  names are y1-y10 c1 c2;
  cluster=c2 c1;
  missing=all(999);

data:    file=3imp.dat;

analysis:  type = threelevel; estimator=bayes;

data imputation:
  ndatasets = 10;
  save = 3levImp*.dat;
  impute = y1-y10;

model:
  %within%
  y1-y10*1;
  e1 by y1-y10*1; e1@1;

  %between c1%
  y1-y10*.5;
  e2 by y1-y10*1; e2@1;

  %between c2%
  y1-y10*.3;
  e3 by y1-y10*1; e3@1;
Regression analysis
Path analysis (both subject and context are random modes)
SEM
Random items (both subject and item are random modes)
Longitudinal analysis (both subject and time are random modes)
Students are cross-classified by school and neighbourhood at level 2. An example with 33 students:

<table>
<thead>
<tr>
<th>Neighbourhood 1</th>
<th>School 1</th>
<th>School 2</th>
<th>School 3</th>
<th>School 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XXXX</td>
<td>XX</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Neighbourhood 2</td>
<td>X</td>
<td>XXXXX</td>
<td>XXXX</td>
<td>XX</td>
</tr>
<tr>
<td>Neighbourhood 3</td>
<td>XX</td>
<td>XX</td>
<td>XXXXX</td>
<td>XXXXXXX</td>
</tr>
</tbody>
</table>

Cross-Classified Data

- $Y_{pijk}$ is the $p$–th observation for person $i$ belonging to level 2 cluster $j$ and level 3 cluster $k$.
- Level 2 clusters are not nested within level 3 clusters
- Examples:
  - Natural Nesting: Students performance scores are nested within students and teachers. Students are nested within schools and neighborhoods.
  - Design Nesting: Studies where observations are nested within persons and treatments/situations.
  - Complex Sampling: Observations are nested within sampling units and another variable unrelated to the sampling.
  - Generalizability theory: Items are considered a random sample from a population of items.
Why do we need to model both sets of clustering?

- Discover the true predictor/explanatory effect stemming from the clusters
- Ignoring clustering leads to incorrect standard errors
- Modeling with fixed effects leads to too many parameters and less accurate model
7.1 Cross-Classified Regression

Consider an outcome $y_{ijk}$ for individual $i$ nested within the cross-classification of level 2a with index $j$ and level 2b with index $k$. For example, level 2a is the school an individual goes to and level 2b is the neighborhood the individual lives in. This is not a three-level structure because a school an individual goes to need not be in the neighborhood the individual lives in. Following is a simple model,

$$y_{ijk} = \beta_0 + \beta_1 x_{ijk} + \beta_{2a} j + \beta_{2b} k + \epsilon_{ijk}, \quad (22)$$

$$\beta_{2a} j = \gamma_a w_{2a} j + \zeta_{2a} j, \quad (23)$$

$$\beta_{2b} k = \gamma_b z_{2b} k + \zeta_{2b} k, \quad (24)$$

where

- $x$, $w_{2a}$, and $z_{2b}$ are covariates on the different levels
- $\beta_0$, $\beta_1$, $\gamma_a$ and $\gamma_b$ are fixed effect coefficients on the different levels
- $\epsilon$, $\beta_{2a} j$ and $\beta_{2b} k$ are random effects on the different levels
7.2 Cross-Classified Regression: UG Example 9.24

Within

Level 2a

Level 2b
TITLE:  this is an example of a two-level regression for a continuous dependent variable using cross-classified data
DATA:  FILE = ex9.24.dat;
VARIABLE:  NAMES = y x1 x2 w z level2a level2b;
CLUSTER = level2b level2a;
WITHIN = x1 x2;
BETWEEN = (level2a) w (level2b) z;
ANALYSIS:  TYPE = CROSSCLASSIFIED RANDOM;
ESTIMATOR = BAYES;
PROCESSORS = 2;
BITERATIONS = (2000);
MODEL:  %WITHIN%
y ON x1;
s | y ON x2;
%BETWEEN level2a%
y ON w;
s ON w;
y WITH s;
%BETWEEN level2b%
y ON z;
s ON Z;
y WITH s;
OUTPUT:  TECH1 TECH8;
7.3 Cross-Classified Regression: Pupcross Data


- 1000 pupils, attending 100 different primary schools, going on to 30 secondary schools
- Outcome: Achievement measured in secondary school
- \( x \) covariate: pupil gender (0=male, 1=female), pupil ses
- \( w_{2a} \) covariate: pdenom (0=public, 1=denom); primary school denomination
- \( z_{2b} \) covariate: sdenom (0=public, 1=denom); secondary school denomination
Cross-Classified Modeling Of Pupcross Data

Within

Pschool (2a)

Sschool (2b)
TITLE: Pupcross: No covariates
DATA: FILE = pupcross.dat;
VARIABLE: NAMES = pupil pschool sschool achieve pupsex pupses pdenom sdenom;
  USEVARIABLES = achieve;
  CLUSTER = pschool sschool;
ANALYSIS: ESTIMATOR = BAYES;
  TYPE = CROSSCLASSIFIED;
  PROCESSORS = 2;
  FBITER = 5000;
MODEL: %WITHIN%
  achieve;
  %BETWEEN pschool%
  achieve;
  %BETWEEN sschool%
  achieve;
OUTPUT: TECH1 TECH8;
<table>
<thead>
<tr>
<th>Size (s)</th>
<th>Cluster ID with Size s</th>
<th>Size (s)</th>
<th>Cluster ID with Size s</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>9</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>12</td>
<td>43</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>24</td>
<td>23 5 22</td>
</tr>
<tr>
<td>24</td>
<td>15</td>
<td>15</td>
<td>47 23 5 22</td>
</tr>
<tr>
<td>26</td>
<td>3 17</td>
<td>16</td>
<td>30 9</td>
</tr>
<tr>
<td>27</td>
<td>1 30</td>
<td>17</td>
<td>7 26 38</td>
</tr>
<tr>
<td>28</td>
<td>23</td>
<td>18</td>
<td>1 3 6 45 14 28</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
<td>19</td>
<td>29 17 49 35 21 20</td>
</tr>
<tr>
<td>31</td>
<td>26 25</td>
<td>20</td>
<td>16 2</td>
</tr>
<tr>
<td>32</td>
<td>2</td>
<td>21</td>
<td>40 32 46 11 19 13 4 39</td>
</tr>
<tr>
<td>33</td>
<td>8 13</td>
<td>22</td>
<td>34 27</td>
</tr>
<tr>
<td>34</td>
<td>4 18</td>
<td>23</td>
<td>15 18</td>
</tr>
<tr>
<td>35</td>
<td>29</td>
<td>24</td>
<td>25 44 37</td>
</tr>
<tr>
<td>37</td>
<td>27 11</td>
<td>25</td>
<td>36 31 10</td>
</tr>
<tr>
<td>39</td>
<td>22 19</td>
<td>26</td>
<td>8</td>
</tr>
<tr>
<td>41</td>
<td>16</td>
<td>27</td>
<td>42</td>
</tr>
<tr>
<td>42</td>
<td>21 7</td>
<td>29</td>
<td>48 12</td>
</tr>
<tr>
<td>45</td>
<td>14</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>46</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>Posterior S.D.</td>
<td>One-Tailed P-Value</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>WITHIN level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.513</td>
<td>0.024</td>
<td>0.000</td>
</tr>
<tr>
<td>BETWEEN school level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.075</td>
<td>0.028</td>
<td>0.000</td>
</tr>
<tr>
<td>BETWEEN school level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>6.341</td>
<td>0.084</td>
<td>0.000</td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.183</td>
<td>0.046</td>
<td>0.000</td>
</tr>
</tbody>
</table>
TITLE: Pupcross: Adding pupil gender and ses
DATA: FILE = pupcross.dat;
VARIABLE: NAMES = pupil pschool sschool achieve pupsex pupses pdenom sdenom;
USEVARIABLES = achieve pupsex pupses;
CLUSTER = pschool sschool;
WITHIN = pupsex pupses;
ANALYSIS: ESTIMATOR = BAYES;
TYPE = CROSSCLASSIFIED;
PROCESSORS = 2;
FBITER = 5000;
MODEL: %WITHIN%
achieve ON pupsex pupses;
%BETWEEN pschool%
achieve;
%BETWEEN school%
achieve;
OUTPUT: TECH1 TECH8;
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Posterior S.D.</th>
<th>One-Tailed P-Value</th>
<th>95% C.I. Lower 2.%</th>
<th>95% C.I. Upper 2.5%</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WITHIN level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pupsex</td>
<td>0.262</td>
<td>0.046</td>
<td>0.000</td>
<td>0.171</td>
<td>0.353</td>
<td>*</td>
</tr>
<tr>
<td>pupses</td>
<td>0.114</td>
<td>0.016</td>
<td>0.000</td>
<td>0.081</td>
<td>0.145</td>
<td>*</td>
</tr>
<tr>
<td>Residual variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.477</td>
<td>0.022</td>
<td>0.000</td>
<td>0.434</td>
<td>0.523</td>
<td>*</td>
</tr>
<tr>
<td><strong>BETWEEN sschool level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.073</td>
<td>0.028</td>
<td>0.000</td>
<td>0.038</td>
<td>0.145</td>
<td>*</td>
</tr>
<tr>
<td><strong>BETWEEN pschool level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>5.757</td>
<td>0.109</td>
<td>0.000</td>
<td>5.539</td>
<td>5.975</td>
<td>*</td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.183</td>
<td>0.046</td>
<td>0.000</td>
<td>0.116</td>
<td>0.297</td>
<td>*</td>
</tr>
</tbody>
</table>
TITLE: Pupil gender and ses with random ses slope for primary schools

VARIABLE: NAMES = pupil pschool sschool achieve pupsex pupses pdenom sdenom;
USEVARIABLES = achieve pupsex pupses;
CLUSTER = pschool sschool;
WITHIN = pupsex pupses;

ANALYSIS: ESTIMATOR = BAYES;
TYPE = CROSSCLASSIFIED RANDOM;
PROCESSORS = 2; FBITER = 5000;

MODEL: %WITHIN%
achieve ON pupsex;
s | achieve ON pupses;
%BETWEEN PSCHOOL%
achieve;
s;
%BETWEEN SSCHOOL%
achieve;
s@0;

OUTPUT: TECH1 TECH8;
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Posterior S.D.</th>
<th>One-Tailed P-Value</th>
<th>95% C.I. Lower 2.%</th>
<th>95% C.I. Upper 2.5%</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WITHIN level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pupsex</td>
<td>0.253</td>
<td>0.045</td>
<td>0.000</td>
<td>0.163</td>
<td>0.339</td>
<td>*</td>
</tr>
<tr>
<td><strong>Residual variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.465</td>
<td>0.022</td>
<td>0.000</td>
<td>0.424</td>
<td>0.510</td>
<td>*</td>
</tr>
<tr>
<td><strong>BETWEEN sschool level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.071</td>
<td>0.027</td>
<td>0.000</td>
<td>0.038</td>
<td>0.140</td>
<td>*</td>
</tr>
<tr>
<td>s</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td><strong>BETWEEN pschool level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>5.758</td>
<td>0.105</td>
<td>0.000</td>
<td>5.557</td>
<td>5.964</td>
<td>*</td>
</tr>
<tr>
<td>s</td>
<td>0.116</td>
<td>0.019</td>
<td>0.000</td>
<td>0.077</td>
<td>0.153</td>
<td>*</td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.110</td>
<td>0.045</td>
<td>0.000</td>
<td>0.042</td>
<td>0.216</td>
<td>*</td>
</tr>
<tr>
<td>s</td>
<td>0.006</td>
<td>0.002</td>
<td>0.000</td>
<td>0.002</td>
<td>0.011</td>
<td>*</td>
</tr>
</tbody>
</table>
TITLE: Pupil gender and ses plus pschool pdenom

VARIABLE: NAMES = pupil pschool sschool achieve pupsex pupses pdenom sdenom;
USEVARIABLES = achieve pupsex pupses pdemon; !sdenom;
CLUSTER = pschool sschool;
WITHIN = pupsex pupses;
BETWEEN = (pschool) pdenom; ! (sschool) sdenom;

ANALYSIS: ESTIMATOR = BAYES;
TYPE = CROSSCLASSIFIED;
PROCESSORS = 2; FBITER = 5000;

MODEL: %WITHIN%
achieve ON pupsex pupses;
%BETWEEN PSCHOOL%
achieve ON pdenom;
%BETWEEN SSCHOOL%
achieve; ! ON sdenom;

OUTPUT: TECH1 TECH8;
PLOT: TYPE = PLOT3;
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Posterior S.D.</th>
<th>One-Tailed P-Value</th>
<th>95% C.I. Lower 2%</th>
<th>95% C.I. Upper 2.5%</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WITHIN level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pupsex</td>
<td>0.261</td>
<td>0.047</td>
<td>0.000</td>
<td>0.168</td>
<td>0.351</td>
<td>*</td>
</tr>
<tr>
<td>pupses</td>
<td>0.113</td>
<td>0.016</td>
<td>0.000</td>
<td>0.080</td>
<td>0.143</td>
<td>*</td>
</tr>
<tr>
<td>Residual variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.477</td>
<td>0.023</td>
<td>0.000</td>
<td>0.436</td>
<td>0.522</td>
<td>*</td>
</tr>
<tr>
<td><strong>BETWEEN sschool level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.073</td>
<td>0.028</td>
<td>0.000</td>
<td>0.038</td>
<td>0.145</td>
<td>*</td>
</tr>
<tr>
<td><strong>BETWEEN pschool level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pdenom</td>
<td>0.207</td>
<td>0.131</td>
<td>0.058</td>
<td>-0.053</td>
<td>0.465</td>
<td></td>
</tr>
<tr>
<td>Intercepts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>5.643</td>
<td>0.136</td>
<td>0.000</td>
<td>5.375</td>
<td>5.912</td>
<td>*</td>
</tr>
<tr>
<td>Residual variances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>achieve</td>
<td>0.175</td>
<td>0.045</td>
<td>0.000</td>
<td>0.112</td>
<td>0.288</td>
<td>*</td>
</tr>
</tbody>
</table>
7.4 Cross-Classified Path Analysis: UG Example 9.25

Within

Level 2a

Level 2b
TITLE: this is an example of a two-level path analysis with continuous dependent variables using cross-classified data

DATA: FILE = ex9.25.dat;

VARIABLE: NAMES = y1 y2 x w z level2a level2b;
CLUSTER = level2b level2a;

WITHIN = x;
BETWEEN = (level2a) w (level2b) z;

ANALYSIS: TYPE = CROSSCLASSIFIED;
ESTIMATOR = BAYES;
PROCESSORS = 2;

MODEL:

%WITHIN%
y2 ON y1 x;
y1 ON x;

%BETWEEN level2a%
y1-y2 ON w;
y1 WITH y2;

%BETWEEN level2b%
y1-y2 ON z;
y1 WITH y2;

OUTPUT: TECH1 TECH8;
Advanced topics:

- 2-mode path analysis
- Cross-classified SEM
- Random item IRT

- A population of situations that might elicit negative emotional responses
- 11 situations (e.g. blamed for someone else’s failure after a sports match, a fellow student fails to return your notes the day before an exam, you hear that a friend is spreading gossip about you) viewed as randomly drawn from a population of situations
- 4 binary responses: Frustration, antagonistic action, irritation, anger
- n=679 high school students
- Level 2 cluster variables are situations and students
- 1 observation for each pair of clustering units
Research questions: Which of the relationships below are significant? Are the relationships the same on the situation level as on the subject level?

Figure 3. Graphical representation of the research questions. $a$, $b$, $c$, and $d$ are effect parameters.
VARIABLE: NAMES = frust antag irrit anger student situation;
CLUSTER = situation student;
CATEGORICAL = frust antag irrit anger;
DATA: FILE = gonzalez.dat;
ANALYSIS: TYPE = CROSSCLASSIFIED;
ESTIMATOR = BAYES;
BITERATIONS = (10000);
MODEL: %WITHIN%
  irrit anger ON frust antag;
  irrit WITH anger;
  frust WITH antag;
%BETWEEN student%
  irrit ON frust (1);
  anger ON frust (2);
  irrit ON antag (3);
  anger ON antag (4);
  irrit; anger; irrit WITH anger;
  frust; antag; frust WITH antag;
%BETWEEN situation%
irrit ON frust (1);
anger ON frust (2);
irrit ON antag (3);
anger ON antag (4);
irrit; anger; irrit WITH anger;
frust; antag; frust WITH antag;

OUTPUT: TECH8 TECH9 STDY;
PLOT: TYPE = PLOT2;
8.2 2-Mode Path Analysis: Monte Carlo Simulation Using The Gonzalez Model

M is the number of cluster units for both between levels, $\beta$ is the common slope, $\psi$ is the within-level correlation, $\tau$ is the binary outcome threshold. Table gives bias (coverage).

<table>
<thead>
<tr>
<th>Para</th>
<th>M=10</th>
<th>M=20</th>
<th>M=30</th>
<th>M=50</th>
<th>M=100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.13(0.92)</td>
<td>0.05(0.89)</td>
<td>0.00(0.97)</td>
<td>0.01(0.92)</td>
<td>0.01(0.94)</td>
</tr>
<tr>
<td>$\psi_{2,11}$</td>
<td>0.11(1.00)</td>
<td>0.06(0.96)</td>
<td>0.01(0.98)</td>
<td>0.00(0.89)</td>
<td>0.02(0.95)</td>
</tr>
<tr>
<td>$\psi_{2,12}$</td>
<td>0.15(0.97)</td>
<td>0.06(0.92)</td>
<td>0.05(0.97)</td>
<td>0.03(0.87)</td>
<td>0.01(0.96)</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>0.12(0.93)</td>
<td>0.01(0.93)</td>
<td>0.00(0.90)</td>
<td>0.03(0.86)</td>
<td>0.00(0.91)</td>
</tr>
</tbody>
</table>

Small biases for $M = 10$. Due to parameter equalities information is combined from both clustering levels. Adding unconstrained level 1 model: tetrachoric correlation matrix.
8.3 Cross-Classified SEM

- General SEM model: 2-way ANOVA. $Y_{pijk}$ is the $p$–th variable for individual $i$ in cluster $j$ and cross cluster $k$

$$Y_{pijk} = Y_{1pijk} + Y_{2pj} + Y_{3pk}$$

- 3 sets of structural equations - one on each level

$$Y_{1ijk} = \nu + \Lambda_1 \eta_{ijk} + \varepsilon_{ijk}$$
$$\eta_{ijk} = \alpha + B_1 \eta_{ijk} + \Gamma_1 x_{ijk} + \xi_{ijk}$$

$$Y_{2j} = \Lambda_2 \eta_j + \varepsilon_j$$
$$\eta_j = B_2 \eta_j + \Gamma_2 x_j + \xi_j$$

$$Y_{3k} = \Lambda_3 \eta_k + \varepsilon_k$$
$$\eta_k = B_3 \eta_k + \Gamma_3 x_k + \xi_k$$
The regression coefficients on level 1 can be a random effects from each of the two clustering levels: combines cross-classified SEM and cross classified HLM.

Bayesian MCMC estimation: used as a frequentist estimator.

Easily extends to categorical variables.

ML estimation possible only when one of the two level of clustering has small number of units.
1 factor at the individual level and 1 factor at each of the clustering levels, 5 indicator variables on the individual level

\[ y_{pijk} = \mu_p + \lambda_{1,p}f_{1,ijk} + \lambda_{2,p}f_{2,j} + \lambda_{3,p}f_{3,k} + \varepsilon_{2,pj} + \varepsilon_{3, pk} + \varepsilon_{1, pijk} \]

- M level 2 clusters. M level 3 clusters. 1 unit within each cluster intersection. More than 1 unit is possible. Zero units possible: sparse tables
- Monte Carlo simulation: Estimation takes less than 1 min per replication
Cross-Classified Model Example 1: Factor Model Results

Table: Absolute bias and coverage for cross-classified factor analysis model

<table>
<thead>
<tr>
<th>Param</th>
<th>M=10</th>
<th>M=20</th>
<th>M=30</th>
<th>M=50</th>
<th>M=100</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\lambda_{1,1})</td>
<td>0.07(0.92)</td>
<td>0.03(0.89)</td>
<td>0.01(0.95)</td>
<td>0.00(0.97)</td>
<td>0.00(0.91)</td>
</tr>
<tr>
<td>(\theta_{1,1})</td>
<td>0.05(0.96)</td>
<td>0.00(0.97)</td>
<td>0.00(0.95)</td>
<td>0.00(0.99)</td>
<td>0.00(0.94)</td>
</tr>
<tr>
<td>(\lambda_{2,p})</td>
<td>0.21(0.97)</td>
<td>0.11(0.94)</td>
<td>0.10(0.93)</td>
<td>0.06(0.94)</td>
<td>0.00(0.92)</td>
</tr>
<tr>
<td>(\theta_{2,p})</td>
<td>0.24(0.99)</td>
<td>0.10(0.95)</td>
<td>0.04(0.92)</td>
<td>0.05(0.94)</td>
<td>0.02(0.96)</td>
</tr>
<tr>
<td>(\lambda_{3,p})</td>
<td>0.45(0.99)</td>
<td>0.10(0.97)</td>
<td>0.03(0.99)</td>
<td>0.01(0.92)</td>
<td>0.03(0.97)</td>
</tr>
<tr>
<td>(\theta_{3,p})</td>
<td>0.75(1.00)</td>
<td>0.25(0.98)</td>
<td>0.15(0.97)</td>
<td>0.12(0.98)</td>
<td>0.05(0.92)</td>
</tr>
<tr>
<td>(\mu_p)</td>
<td>0.01(0.99)</td>
<td>0.04(0.98)</td>
<td>0.01(0.97)</td>
<td>0.05(0.99)</td>
<td>0.00(0.97)</td>
</tr>
</tbody>
</table>

Perfect coverage. Level 1 parameters estimated very well. Biases when the number of clusters is small \(M = 10\). Weakly informative priors can reduce the bias for small number of clusters.
Type 1: Random slope.

%WITHIN%

s | y ON x;

$s$ has variance on both crossed levels. Dependent variable can be within-level factor. Covariate $x$ should be on the WITHIN = list.

Type 2: Random loading.

%WITHIN%

s | f BY y;

$s$ has variance on both crossed levels. $f$ is a within-level factor. The dependent variable can be a within-level factor.

Type 3: Crossed random loading.

%BETWEEN level2a%

s | f BY y;

$s$ has variance on crossed level 2b and is defined on crossed level 2a. $f$ is a level 2a factor, $s$ is a level 2b factor. This is a way to use the interaction term $s \cdot f$. 
8.6 Random Items, Generalizability Theory

- Items are random samples from a population of items.
- The same or different items may be administered to individuals.
- Suited for computer generated items and adaptive testing.
- 2-parameter IRT model

\[ P(Y_{ij} = 1) = \Phi(a_j \theta_i + b_j) \]

- \( a_j \sim N(a, \sigma_a) \), \( b_j \sim N(b, \sigma_b) \): random discrimination and difficulty parameters
- The ability parameter is \( \theta_i \sim N(0, 1) \)
- Cross-classified model. Nested within items and individuals. 1 or 0 observation in each cross-classified cell.
- Interaction of two latent variables: \( a_j \) and \( \theta_i \): Type 3 crossed random loading
- The model has only 4 parameters - much more parsimonious than regular IRT models.
VARIABLE:
  NAMES = u item individual;
  CLUSTER = item individual;
  CATEGORICAL = u;

ANALYSIS:
  TYPE = CROSS RANDOM;
  ESTIMATOR = BAYES;

MODEL:
  %WITHIN%

  %BETWEEN individual%
  s | f BY u;
  f @ 1 u @ 0;
  %BETWEEN item%
  u s;
8.7 Random Item 2-Parameter IRT: TIMMS Example

- 8 test items, 478 students

<table>
<thead>
<tr>
<th>parameter</th>
<th>estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>average discrimination $a$</td>
<td>0.752</td>
<td>0.094</td>
</tr>
<tr>
<td>average difficulty $b$</td>
<td>0.118</td>
<td>0.376</td>
</tr>
<tr>
<td>variation of discrimination $a$</td>
<td>0.050</td>
<td>0.046</td>
</tr>
<tr>
<td>variation of difficulty $b$</td>
<td>1.030</td>
<td>0.760</td>
</tr>
</tbody>
</table>

- 8 items means that there are only 8 clusters on the %between item% level and therefore the variance estimates at that level are affected by their priors. If the number of clusters is less than 10 or 20 there is prior dependence in the variance parameters.
Using factor scores estimation we can estimate item specific parameter and SE using posterior mean and posterior standard deviation.

<table>
<thead>
<tr>
<th>item</th>
<th>discrimination</th>
<th>SE</th>
<th>difficulty</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>0.797</td>
<td>0.11</td>
<td>-1.018</td>
<td>0.103</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.613</td>
<td>0.106</td>
<td>-0.468</td>
<td>0.074</td>
</tr>
<tr>
<td>Item 3</td>
<td>0.905</td>
<td>0.148</td>
<td>-1.012</td>
<td>0.097</td>
</tr>
<tr>
<td>Item 4</td>
<td>0.798</td>
<td>0.118</td>
<td>-1.312</td>
<td>0.106</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.538</td>
<td>0.099</td>
<td>0.644</td>
<td>0.064</td>
</tr>
<tr>
<td>Item 6</td>
<td>0.808</td>
<td>0.135</td>
<td>0.023</td>
<td>0.077</td>
</tr>
<tr>
<td>Item 7</td>
<td>0.915</td>
<td>0.157</td>
<td>0.929</td>
<td>0.09</td>
</tr>
<tr>
<td>Item 8</td>
<td>0.689</td>
<td>0.105</td>
<td>1.381</td>
<td>0.108</td>
</tr>
</tbody>
</table>
### Table: Random 2-parameter IRT item specific parameters

<table>
<thead>
<tr>
<th>Item</th>
<th>Bayes random discrimination</th>
<th>Bayes random SE</th>
<th>ML fixed discrimination</th>
<th>ML fixed SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>0.797</td>
<td>0.110</td>
<td>0.850</td>
<td>0.155</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.613</td>
<td>0.106</td>
<td>0.579</td>
<td>0.102</td>
</tr>
<tr>
<td>Item 3</td>
<td>0.905</td>
<td>0.148</td>
<td>0.959</td>
<td>0.170</td>
</tr>
<tr>
<td>Item 4</td>
<td>0.798</td>
<td>0.118</td>
<td>0.858</td>
<td>0.172</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.538</td>
<td>0.099</td>
<td>0.487</td>
<td>0.096</td>
</tr>
<tr>
<td>Item 6</td>
<td>0.808</td>
<td>0.135</td>
<td>0.749</td>
<td>0.119</td>
</tr>
<tr>
<td>Item 7</td>
<td>0.915</td>
<td>0.157</td>
<td>0.929</td>
<td>0.159</td>
</tr>
<tr>
<td>Item 8</td>
<td>0.689</td>
<td>0.105</td>
<td>0.662</td>
<td>0.134</td>
</tr>
</tbody>
</table>

- Bayes random estimates are shrunk towards the mean and have smaller standard errors: shrinkage estimate
One can add a predictor for a person’s ability. For example adding gender as a predictor yields an estimate of 0.283 (0.120), saying that males have a significantly higher math mean.

Predictors for discrimination and difficulty random effects, for example, geometry indicator.

More parsimonious model can yield more accurate ability estimates.
De Boeck (2008) Random item IRT models
24 verbal aggression items, 316 persons

\[ P(Y_{ij} = 1) = \Phi(\theta_i + b_j) \]

\[ b_j \sim N(b, \sigma) \]

\[ \theta_i \sim N(0, \tau) \]

**Table: Random Rasch IRT - variance decomposition**

<table>
<thead>
<tr>
<th>parameter</th>
<th>person</th>
<th>item</th>
<th>error</th>
</tr>
</thead>
<tbody>
<tr>
<td>estimates(SE)</td>
<td>1.89(0.19)</td>
<td>1.46(0.53)</td>
<td>2.892</td>
</tr>
<tr>
<td>variance explained</td>
<td>30%</td>
<td>23%</td>
<td>46%</td>
</tr>
</tbody>
</table>
MODEL:

%WITHIN%

%BETWEEN person%
y;

%BETWEEN item%
y;
An old dilemma
Two new solutions
Single-level analysis with $p \times T = 2 \times 5 = 10$ variables, $T = 5$ factors.

- ML hard and impossible as $T$ increases (numerical integration)
- WLSMV possible but hard when $p \times T$ increases and biased unless attrition is MCAR or multiple imputation is done first
- Bayes possible
- Searching for partial measurement invariance is cumbersome
Two-level analysis with $p = 2$ variables, 1 within-factor, 2-between factors, assuming full measurement invariance across time.

- ML feasible
- WLSMV feasible (2-level WLSMV)
- Bayes feasible
Both old approaches have problems
  - Wide, single-level approach easily gets significant non-invariance and needs many modifications
  - Long, two-level approach has to assume invariance

New solution no. 1, suitable for small to medium number of time points
  - A new wide, single-level approach where time is a fixed mode

New solution no. 2, suitable for medium to large number of time points
  - A new long, two-level approach where time is a random mode
  - No limit on the number of time points
New Solution No. 1: Wide Format, Single-Level Approach

Single-level analysis with $p \times T = 2 \times 5 = 10$ variables, $T = 5$ factors.

- Bayes ("BSEM") using approximate measurement invariance, still identifying factor mean and variance differences across time.
New solution no. 2, time is a random mode
A new long, two-level approach
Best of both worlds: Keeping the limited number of variables of the two-level approach without having to assume invariance
Two-level analysis with \( p = 2 \) variables.

- Bayes twolevel random approach with random measurement parameters and random factor means and variances using Type=Crossclassified: Clusters are time and person
Randomized field experiment in Baltimore public schools with a classroom-based intervention aimed at reducing aggressive-disruptive behavior among elementary school students (Ialongo et al., 1999).

This analysis:

- Cohort 1
- 9 binary items at 8 time points, Grade 1 - Grade 7
- \( n = 1174 \)
Aggressive-Disruptive Behavior In The Classroom: ML Versus BSEM

- Traditional ML analysis
  - 8 dimensions of integration
  - Computing time: 25:44 with Integration = Montecarlo(5000)
  - Increasing the number of time points makes ML impossible

- BSEM analysis with approximate measurement invariance across time
  - 156 parameters
  - Computing time: 4:01
  - Increasing the number of time points has relatively less impact
USEVARIABLES = stub1f-tease7s;
CATEGORICAL = stub1f-tease7s;
MISSING = ALL (999);

DEFINE: CUT stub1f-tease7s (1.5);

ANALYSIS: ESTIMATOR = BAYES;
PROCESSORS = 2;

MODEL: f1f by stub1f-tease1f* (lam11-lam19);
f1s by stub1s-tease1s* (lam21-lam29);
f2s by stub2s-tease2s* (lam31-lam39);
f3s by stub3s-tease3s* (lam41-lam49);
f4s by stub4s-tease4s* (lam51-lam59);
f5s by stub5s-tease5s* (lam61-lam69);
f6s by stub6s-tease6s* (lam71-lam79);
f7s by stub7s-tease7s* (lam81-lam89);
f1f@1;
[stub1f$1-tease1f$1] (tau11-tau19);
[stub1s$1-tease1s$1] (tau21-tau29);
[stub2s$1-tease2s$1] (tau31-tau39);
[stub3s$1-tease3s$1] (tau41-tau49);
[stub4s$1-tease4s$1] (tau51-tau59);
[stub5s$1-tease5s$1] (tau61-tau69);
[stub6s$1-tease6s$1] (tau71-tau79);
[stub7s$1-tease7s$1] (tau81-tau89);
[f1f-f7s@0];

MODEL
PRIORS: DO(1,9) DIFF(lam1#-lam8#) ∼ N(0,.01);
DO(1,9) DIFF(tau1#-tau8#) ∼ N(0,.01);
OUTPUT: TECH1 TECH8;
### Estimates For Aggressive-Disruptive Behavior

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.D.</th>
<th>P-Value</th>
<th>Lower 2.5%</th>
<th>Upper 2.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Means</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>S</td>
<td>0.238</td>
<td>0.068</td>
<td>0.000</td>
<td>0.108</td>
<td>0.366</td>
</tr>
<tr>
<td>Q</td>
<td>-0.022</td>
<td>0.011</td>
<td>0.023</td>
<td>-0.043</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Variances</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>9.258</td>
<td>2.076</td>
<td>0.000</td>
<td>6.766</td>
<td>14.259</td>
</tr>
<tr>
<td>S</td>
<td>0.258</td>
<td>0.068</td>
<td>0.000</td>
<td>0.169</td>
<td>0.411</td>
</tr>
<tr>
<td>Q</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Posterior</td>
<td>One-Tailed</td>
<td>95% C.I.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
<td>------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>S.D.</td>
<td>P-Value</td>
<td>Lower 2.5%</td>
<td>Upper 2.5%</td>
</tr>
<tr>
<td>STUB1F</td>
<td>0.428</td>
<td>0.048</td>
<td>0.000</td>
<td>0.338</td>
<td>0.522</td>
</tr>
<tr>
<td>BKRULE1F</td>
<td>0.587</td>
<td>0.068</td>
<td>0.000</td>
<td>0.463</td>
<td>0.716</td>
</tr>
<tr>
<td>HARMO1F</td>
<td>0.832</td>
<td>0.082</td>
<td>0.000</td>
<td>0.677</td>
<td>0.985</td>
</tr>
<tr>
<td>BKTHIN1F</td>
<td>0.671</td>
<td>0.067</td>
<td>0.000</td>
<td>0.546</td>
<td>0.795</td>
</tr>
<tr>
<td>YELL1F</td>
<td>0.508</td>
<td>0.055</td>
<td>0.000</td>
<td>0.405</td>
<td>0.609</td>
</tr>
<tr>
<td>TAKEP1F</td>
<td>0.717</td>
<td>0.072</td>
<td>0.000</td>
<td>0.570</td>
<td>0.839</td>
</tr>
<tr>
<td>FIGHT1F</td>
<td>0.480</td>
<td>0.052</td>
<td>0.000</td>
<td>0.385</td>
<td>0.579</td>
</tr>
<tr>
<td>LIES1F</td>
<td>0.488</td>
<td>0.054</td>
<td>0.000</td>
<td>0.386</td>
<td>0.589</td>
</tr>
<tr>
<td>TEASE1F</td>
<td>0.503</td>
<td>0.055</td>
<td>0.000</td>
<td>0.404</td>
<td>0.608</td>
</tr>
</tbody>
</table>

...  

<table>
<thead>
<tr>
<th></th>
<th>Posterior</th>
<th>One-Tailed</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>S.D.</td>
<td>P-Value</td>
</tr>
<tr>
<td>STUB7S</td>
<td>0.360</td>
<td>0.049</td>
<td>0.000</td>
</tr>
<tr>
<td>BKRULE7S</td>
<td>0.512</td>
<td>0.068</td>
<td>0.000</td>
</tr>
<tr>
<td>HARMO7S</td>
<td>0.555</td>
<td>0.074</td>
<td>0.000</td>
</tr>
<tr>
<td>BKTHIN7S</td>
<td>0.459</td>
<td>0.063</td>
<td>0.000</td>
</tr>
<tr>
<td>YELL7S</td>
<td>0.525</td>
<td>0.062</td>
<td>0.000</td>
</tr>
<tr>
<td>TAKEP7S</td>
<td>0.500</td>
<td>0.069</td>
<td>0.000</td>
</tr>
<tr>
<td>FIGHT7S</td>
<td>0.515</td>
<td>0.067</td>
<td>0.000</td>
</tr>
<tr>
<td>LIES7S</td>
<td>0.520</td>
<td>0.070</td>
<td>0.000</td>
</tr>
<tr>
<td>TEASE7S</td>
<td>0.495</td>
<td>0.064</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Displaying Non-Invariant Items: Time Points With Significant Differences Compared To The Mean \((V = 0.01)\)

<table>
<thead>
<tr>
<th>Item</th>
<th>Loading</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>stub</td>
<td>3</td>
<td>1, 2, 3, 6, 8</td>
</tr>
<tr>
<td>bkrule</td>
<td>-</td>
<td>5, 8</td>
</tr>
<tr>
<td>harmo</td>
<td>1, 8</td>
<td>2, 8</td>
</tr>
<tr>
<td>bkthin</td>
<td>1, 2, 3, 7, 8</td>
<td>2, 8</td>
</tr>
<tr>
<td>yell</td>
<td>2, 3, 6</td>
<td>-</td>
</tr>
<tr>
<td>takep</td>
<td>1, 2, 5</td>
<td>1, 2, 5</td>
</tr>
<tr>
<td>fight</td>
<td>1, 5</td>
<td>1, 4</td>
</tr>
<tr>
<td>lies</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>tease</td>
<td>-</td>
<td>1, 4, 8</td>
</tr>
</tbody>
</table>
9.2 Cross-Classified Analysis Of Longitudinal Data

- Observations nested within time and subject
- A large number of time points can be handled via Bayesian analysis
- A relatively small number of subjects is needed
Intensive Longitudinal Data

- Time intensive data: More longitudinal data are collected where very frequent observations are made using new tools for data collection. Walls & Schafer (2006)

- Typically multivariate models are developed but if the number of time points is large these models will fail due to too many variables and parameters involved.

- Factor analysis models will be unstable over time. Is it lack of measurement invariance or insufficient model?

- Random loading and intercept models can take care of measurement and intercept invariance. A problem becomes an advantage.

- Random loading and intercept models produce more accurate estimates for the loadings and factors by borrowing information over time.

- Random loading and intercept models produce more parsimonious model.
TITLE: this is an example of longitudinal modeling using a cross-classified data approach where observations are nested within the cross-classification of time and subjects

MONTECARLO:
NAMES = y1-y3;
NOBSERVATIONS = 7500;
NREPS = 1;
CSIZES = 75[100(1)];! 75 subjects, 100 time points
NCSIZE = 1[1];
WITHIN = (level2a) y1-y3;
SAVE = ex9.27.dat;

ANALYSIS:
TYPE = CROSS RANDOM;
ESTIMATOR = BAYES;
PROCESSORS = 2;
MODEL
POPULATION:

%WITHIN%
  s1-s3 | f by y1-y3;
  f@1;
  y1-y3*1.2; [y1-y3@0];
%BETWEEN level2a% ! across time variation
  s1-s3*0.1;
  [s1-s3*1.3];
  y1-y3*.5;
  [y1-y3@0];
%BETWEEN level2b% ! across subjects variation
  f*1; [f*.5];
  s1-s3@0; [s1-s3@0];
TITLE: this is an example of a multiple indicator growth model with random intercepts and factor loadings using cross-classified data

DATA: FILE = ex9.27.dat;

VARIABLE: NAMES = y1-y3 time subject;
USEVARIABLES = y1-y3 timescor;
CLUSTER = subject time;

DEFINE: timescor = (time-1)/100;

ANALYSIS: TYPE = CROSSCLASSIFIED RANDOM;
ESTIMATOR = BAYES;
PROCESSORS = 2;
BITERATIONS = (1000);

MODEL: %WITHIN%
s1-s3 | f BY y1-y3;
f@1;
s | f ON timescor; !slope growth factor s
y1-y3; [y1-y3@0];

%BETWEEN time%
s1-s3; [s1-s3]; ! random loadings
y1-y3; [y1-y3@0]; ! random intercepts
s0; [s0];

%BETWEEN subject%
s; [s]; ! slope growth factor s

OUTPUT: TECH1 TECH8;
Teacher-rated measurement instrument capturing aggressive-disruptive behavior among a sample of U.S. students in Baltimore public schools (Ialongo et al., 1999).

The instrument consists of 9 items scored as 0 (almost never) through 6 (almost always)

A total of 1174 students are observed in 41 classrooms from Fall of Grade 1 through Grade 6 for a total of 8 time points

The multilevel (classroom) nature of the data is ignored in the current analyses

The item distribution is very skewed with a high percentage in the Almost Never category. The items are therefore dichotomized into Almost Never versus the other categories combined

We analyze the data on the original scale as continuous variables and also the dichotomized scale as categorical
For each student a 1-factor analysis model is estimated with the 9 items at each time point.

Let $Y_{pit}$ be the $p$–th item for individual $i$ at time $t$.

We use cross-classified SEM. Observations are nested within individual and time.

Although this example uses only 8 time points the models can be used with any number of time points.
Model 1: Two-level factor model with intercept non-invariance across time

\[ Y_{pit} = \mu_p + \zeta_{pt} + \xi_{pi} + \lambda_p \eta_{it} + \varepsilon_{pit} \]

- \( \mu_p, \lambda_p \) are model parameters, \( \varepsilon_{pit} \sim N(0, \theta_w,p) \) is the residual
- \( \zeta_{pt} \sim N(0, \sigma_p) \) is a random effect to accommodate intercept non-invariance across time
- To correlate the factors \( \eta_{it} \) within individual \( i \)

\[ \eta_{it} = \eta_{b,i} + \eta_{w,it} \]

- \( \eta_{b,i} \sim N(0, \psi) \) and \( \eta_{w,it} \sim N(0, 1) \). The variance is fixed to 1 to identify the scale in the model
- \( \xi_{pi} \sim N(0, \theta_b,p) \) is a between level residual in the between level factor model
- Without the random effect \( \zeta_{pt} \) this is just a standard two-level factor model
MODEL:

%WITHIN%

f BY y1-y9*1 (11-19);

f@1;

%BETWEEN t1%

y1-y9;

%BETWEEN id%

y1-y9;

fb BY y1-y9*1 (11-19);
Model 2: Adding latent growth model for the factor

\[ \eta_{it} = \alpha_i + \beta_i \cdot t + \eta_{w,it} \]

- \( \alpha_i \sim N(0, \nu_{\alpha}) \) is the intercept and \( \beta_i \sim N(\beta, \nu_{\beta}) \) is the slope. For identification purposes again \( \eta_{w,it} \sim N(0, 1) \)
- The model looks for developmental trajectory across time for the aggressive-disruptive behavior factor
MODEL: s = beta, fb = alpha

%WITHIN%
f BY y1-y9*1 (11-19);
f@1;
s | f ON time;

%BETWEEN t1%
y1-y9;
s@0; [s@0];

%BETWEEN id%
y1-y9;
fb BY y1-y9*1 (11-19);
s*1; [s*0];
Model 3: Adding measurement non-invariance

Replace the fixed loadings $\lambda_p$ with random loadings
$\lambda_{pt} \sim N(\lambda_p, w_p)$

The random loadings accommodate measurement non-invariance across time

All models can be estimated for continuous and categorical scale data
MODEL:

%WITHIN%

s1-s9 | f BY y1-y9;
f@1;
s | f ON time;

%BETWEEN t1%

y1-y9;
f@0; [f@0];
s@0; [s@0];
s1-s9*1; [s1-s9*1];

%BETWEEN id%

y1-y9;
f*1; [f@0];
s*1; [s*0];
s1-s9@0; [s1-s9@0];
Aggressive-Disruptive Behavior Example Continued: Model 3 Results For Continuous Analysis

<table>
<thead>
<tr>
<th></th>
<th>Within Level</th>
<th>Between ID Level</th>
<th>Between T1 Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Residual Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>1.073</td>
<td>0.022</td>
<td>0.000</td>
</tr>
<tr>
<td>Y9</td>
<td>0.630</td>
<td>0.014</td>
<td>0.000</td>
</tr>
<tr>
<td>F</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Between ID Level Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>0.146</td>
<td>0.016</td>
<td>0.000</td>
</tr>
<tr>
<td>Y9</td>
<td>0.052</td>
<td>0.009</td>
<td>0.000</td>
</tr>
<tr>
<td>F</td>
<td>1.316</td>
<td>0.080</td>
<td>0.000</td>
</tr>
<tr>
<td>S</td>
<td>0.026</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Between T1 Level Means</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>1.632</td>
<td>0.120</td>
<td>0.000</td>
</tr>
<tr>
<td>Y9</td>
<td>1.232</td>
<td>0.096</td>
<td>0.000</td>
</tr>
<tr>
<td>S1</td>
<td>0.679</td>
<td>0.023</td>
<td>0.000</td>
</tr>
<tr>
<td>S9</td>
<td>0.705</td>
<td>0.043</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Between T1 Level Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>0.080</td>
<td>0.138</td>
<td>0.000</td>
</tr>
<tr>
<td>Y9</td>
<td>0.047</td>
<td>0.109</td>
<td>0.000</td>
</tr>
<tr>
<td>S1</td>
<td>0.002</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>S9</td>
<td>0.010</td>
<td>0.079</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Model 4: Adding measurement non-invariance also across individuals

Replace the loadings $\lambda_{pt}$ with random loadings

$$\lambda_{pit} = \lambda_{pi} + \lambda_{pt}$$

where $\lambda_{pt} \sim N(\lambda_p, w_p)$ and $\lambda_{pi} \sim N(0, w_i)$

The random loadings accommodate measurement non-invariance across time and individual

Model 4: Adding factor variance non-invariance across time. Can be done either by adding (a) introducing a factor model for the random loadings or (b) introducing a random loadings for the residual of the factor.

We choose (b). $\text{Var}(f) = 0.51 + (0.7 + \sigma_t)^2$ where $\sigma_t$ is a mean zero random effect
Aggressive-Disruptive Behavior Example Continued: Model 4 Setup

model:
%within%
s1-s9 | f by y1-y9;
s | f on time; f@0.51;
ss | e by f; e@1;

%between T1%
y1-y9;
s@0; [s@0];
s1-s9*1; [s1-s9*1];
ss*0.6; [ss@0.7];

%between ID%
y1-y9;
f*1; [f@0];
s*1; [s*0];
s1-s9*1; [s1-s9@0];
ss@0; [ss@0];
### Aggressive-Disruptive Behavior Example Continued: Results For Categorical Analysis

#### Between ID Level

<table>
<thead>
<tr>
<th></th>
<th>Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>0.113 0.039 0.000 0.045 0.195</td>
</tr>
<tr>
<td>Y9</td>
<td>0.194 0.049 0.000 0.108 0.299</td>
</tr>
<tr>
<td>S1</td>
<td>0.082 0.029 0.000 0.041 0.152</td>
</tr>
<tr>
<td>S9</td>
<td>0.030 0.025 0.000 0.001 0.089</td>
</tr>
<tr>
<td>F</td>
<td>1.789 0.241 0.000 1.353 2.300</td>
</tr>
</tbody>
</table>

#### Between T1 Level

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.821 0.094 0.000 0.647 1.018</td>
</tr>
<tr>
<td>S9</td>
<td>1.049 0.125 0.000 0.809 1.299</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1$1</td>
<td>-0.864 0.156 0.000 -1.183 -0.558</td>
</tr>
<tr>
<td>Y2$1</td>
<td>-0.800 0.120 0.000 -1.041 -0.568</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>0.120 0.231 0.000 0.035 0.666</td>
</tr>
<tr>
<td>Y9</td>
<td>0.073 0.158 0.000 0.019 0.416</td>
</tr>
<tr>
<td>S1</td>
<td>0.023 0.055 0.000 0.005 0.137</td>
</tr>
<tr>
<td>S9</td>
<td>0.050 0.110 0.000 0.011 0.302</td>
</tr>
</tbody>
</table>
Other extensions of the above model are possible, for example the growth trend can have time specific random effects: \( f \) and \( s \) can be free over time.

The more clusters there are on a particular level the more elaborate the model can be on that level. However, the more elaborate the model on a particular level is, the slower the convergence.

The main factor \( f \) can have a random effect on each of the levels, however the residuals \( Y_i \) should be uncorrelated on that level. If they are correlated through another factor model such as, \( fb \) by \( y_1 - y_9 \), then \( f \) would be confounded with that factor \( fb \) and the model will be poorly identified.

On each level the most general model would be (if there are no random slopes) the unconstrained variance covariance for the dependent variables \( Y_i \). Any model that is a restriction of that model is in principle identified.
Unlike ML and WLS multivariate modeling, for the time intensive Bayes cross-classified SEM, the more time points there are the more stable and easy to estimate the model is.

Bayesian methods solve problems not feasible with ML or WLS.

Time intensive data naturally fits in the cross-classified modeling framework.


Longitudinal growth model for student self-esteem
Each student has 4 observations: 2 in middle school in wave 1 and 2, and 2 in high school in wave 3 and 4
Students have multiple membership: Membership in middle school and in high school with a random effect from both

$Y_{tsmh}$ is observation at time $t$ for student $s$ in middle school $m$ and high school $h$
The model is

\[ Y_{tsmh} = \beta_1 + \beta_2 T2 + \beta_3 T3 + \beta_4 T4 + \delta_s + \delta_m \mu_t + \delta_h \lambda_t + \epsilon_{tsmh} \]

where T2, T3, T4 are dummy variables for wave 2, 3, 4

- \( \delta_s, \delta_m \) and \( \delta_h \) are zero mean random effect contributions from student, middle school and high school

- \( \mu_t = (1, \mu_2, \mu_3, \mu_4) \)

- \( \lambda_t = (0, 0, 1, \lambda_4) \), i.e., no contribution from the high school in wave 1 and 2 because the student is still in middle school

- \( \epsilon_{tsmh} \) is the residual

- Very simple to setup in Mplus
MODEL:

%WITHIN%
fs BY y1-y4@1;
[y1-y4];

%BETWEEN mschool%
fm BY y1@1 y2-y4;
y1-y4@0; [y1-y4@0];

%BETWEEN hschool%
fh BY y1@0 y2@0 y3@1 y4;
y1-y4@0; [y1-y4@0];