Lecture 9: Three-Level Models for Clustered Longitudinal Data

- **Decomposing Clustered Longitudinal Variation**
- Three-Level Model Specification
- Time-Varying Predictors in Three-Level Models
- Clustered Longitudinal Model Examples in SAS, SPSS, and STATA

What determines number of ‘levels’?

- The idea of ‘levels’ concerns the model for the variances, NOT the model for the means
- How many dimensions of sampling do you have?
  - Time within Person → 2-Level Model
  - Time within Person within School → 3-Level Model
  - Time within Person within School within City → 4-Level Model
  - Sometimes additional dimensions are crossed instead of nested, and thus don’t require additional levels (crossed random effects)
- Generally, you will need at least one pile of variance per dimension (for 4 levels, that’s at least 3 U’s and an ‘e’)
  - Put in whatever predictors you want for each level, but keep in mind that the usefulness of your predictors will be constrained by how much Y variance comes from that dimension

Two Options to Deal with Grouping

Clustering/Grouping as Fixed Effects (stays 2-level)

- Include (#groups-1) dummy codes for group membership in the model for the means → ***so group is not another "level"***
- Inference about ***specific*** differences between groups via fixed effects for group, but then you ***cannot*** include between-groups predictors
- Snijders & Bosker ch.4, p. 44 recommend if #groups < 10ish

Clustering/Grouping as a Random Effect (is 3-level)

- Estimate a ***variance*** for group differences in model for variances, such that ***group becomes another "level" of the model***
- Makes an inference about ***population*** of groups via random effects variances, for which you ***can*** include between-groups predictors
- Better option if #groups > 10ish

3-Level Model for Clustered Longitudinal Data: Empty Model

3-Level Clustered MLM

$$\text{L1: } y_{tij} = \beta_{0ij} + e_{tij}$$

$$\text{L2: } \beta_{0ij} = \delta_{00j} + U_{0ij}$$

$$\text{L3: } \delta_{00j} = \gamma_{000} + V_{00j}$$

Fixed Effects:

γ_{000} → grand mean intercept

Random Effects:

U_{0ij} → deviation for level 2

V_{00j} → deviation for level 3

Error:

e_{tij} → level-1 residual deviation

$$Y_{tij} = (\gamma_{000}) + (V_{00j} + U_{0ij} + e_{tij})$$

2 ICC's within a 3-Level Empty Model:
Example: Time within Person within Family

ICC for level 2 (and level 3) relative to level 1:

$$ICC_{L2} = \frac{\text{(Level 3) + (Level 2)}}{\text{(Level 3) + (Level 2) + (Level 1)}}$$

→ This ICC expresses similarity of time points from same person (and by definition, from the same family) → of the ***total variation in Y***, how much of it is ***between-persons***?

ICC for level 3 relative to level 2 (ignoring level 1):

$$ICC_{L3} = \frac{\text{(Level 3)}}{\text{(Level 3) + (Level 2)}}$$

→ This ICC expresses similarity of persons from same family (ignoring within-person variation) → of ***that total BP variation in Y***, how much of that is actually ***between-families***?

2 ICCs within a 3-Level Empty Model:
Example: Student within Class within School

ICC for level 2 (and level 3) relative to level 1:

$$ICC_{L2} = \frac{\text{(Level 3) + (Level 2)}}{\text{(Level 3) + (Level 2) + (Level 1)}}$$

→ This ICC expresses similarity of students from same class (and by definition, from the same school) → of the ***total variation in Y***, how much of it is ***between-classes***?

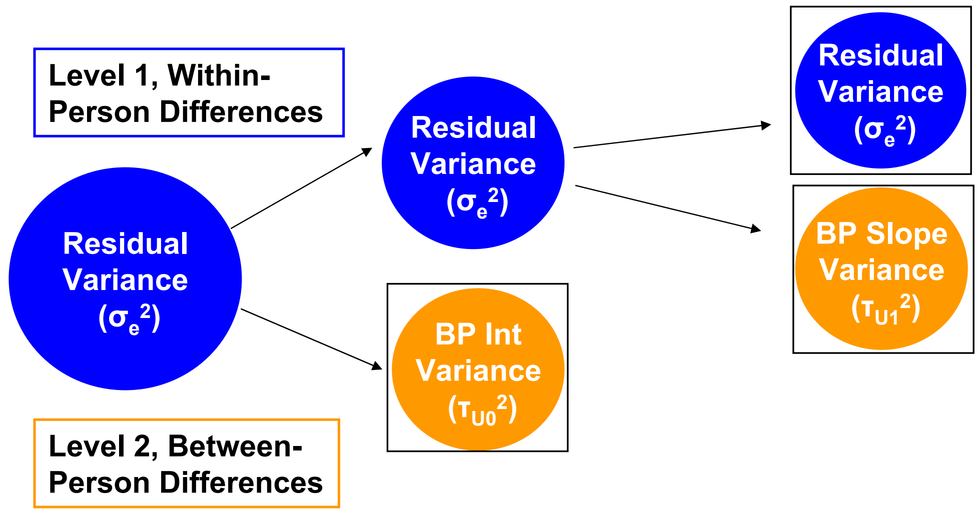
ICC for level 3 relative to level 2 (ignoring level 1):

$$ICC_{L3} = \frac{\text{(Level 3)}}{\text{(Level 3) + (Level 2)}}$$

→ This ICC expresses similarity of classes from same school (ignoring within-class variation) → of ***that total between-group variation in Y***, how much of that is actually ***between-schools***?

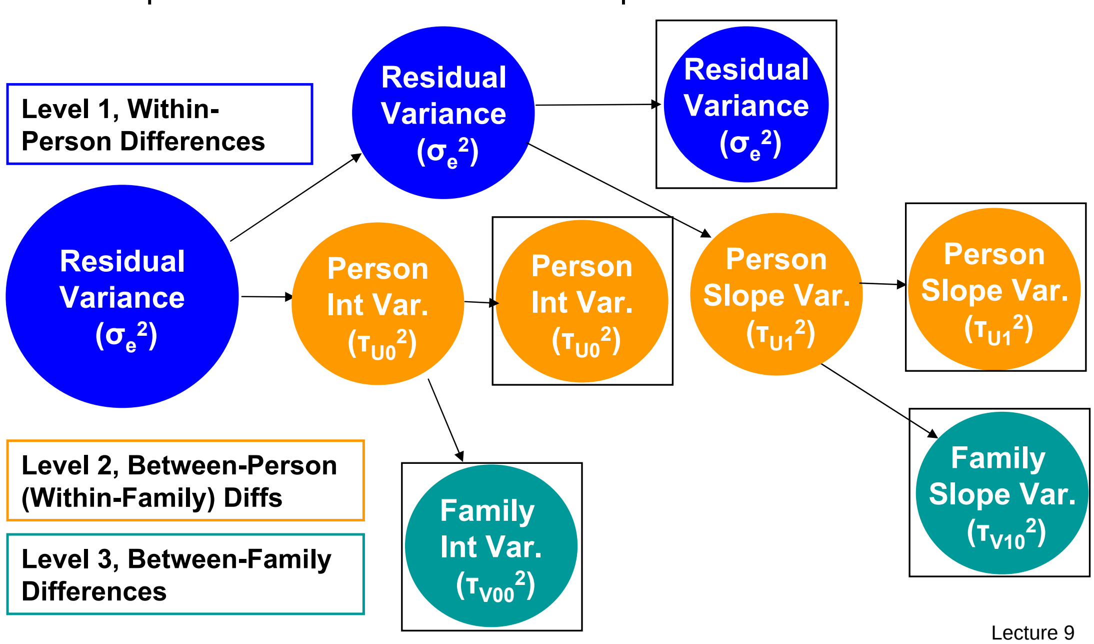
2-Level Models for the Variances

- Summary: Where does the correlation go?
  Into a new variance component (or 'pile of variance')
- **Example 2-level unconditional longitudinal model:**



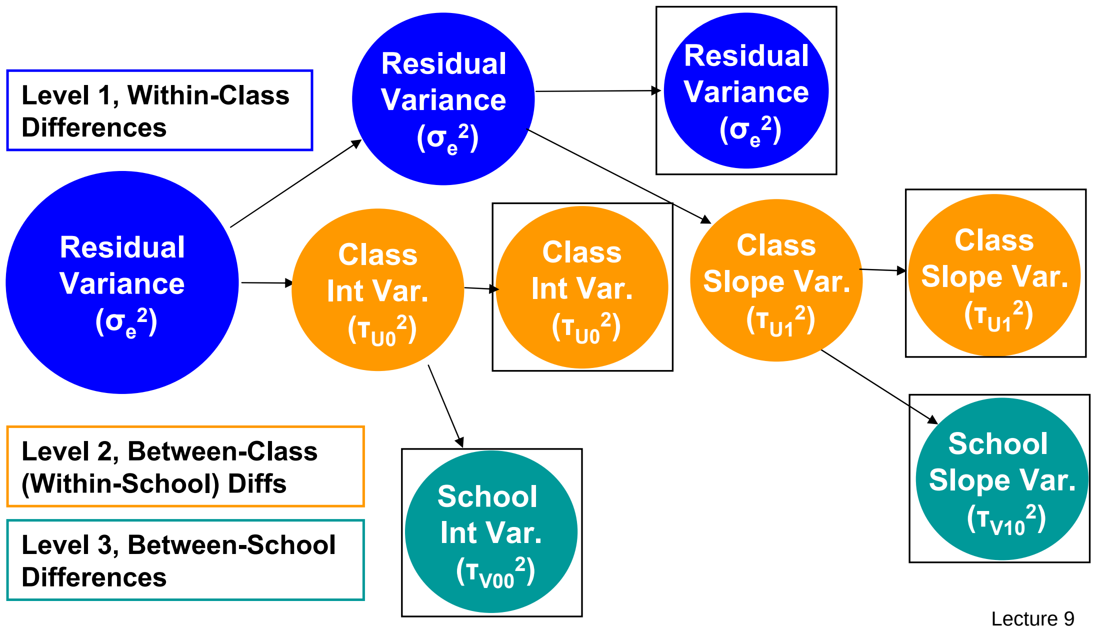
3-Level Models for the Variances

- Example **3-level** model of time within person within families:



3-Level Models for the Variances

- Example **3-level** model of student within class within schools:



Effects of Nesting on Effective *N*

- Design Effect expresses how much effective sample size needs to be adjusted due to nesting

$$\text{ICC} = \frac{\text{Between Variance}}{\text{Between Variance} + \text{Within Variance}}$$

Design Effect = 1 + ((n – 1) * ICC)

$\boldsymbol{N}_{\textbf{effective}}$ = N * n / design effect

where n = average cluster size, N = # clusters

- As ICC increases and cluster size increases, effective sample size goes down
- For 3+ levels, multiple design effects should be examined!

Design Effects in 3-Level Nesting

Design Effect = 1 + (($n_{lower\ level}$ – 1) * ICC)

$N_{\text{effective}} = N_{higher\ level} * n_{lower\ level}$ / design effect

- 10 kids from each of 30 classrooms
- 30 classrooms sampled across 5 schools (6 each)
  - Level 2 ICC (kids in classrooms and schools) = .10
  - Level 3 ICC (classrooms in schools) = .05
- **Level 1 Design effect** = 1 + (9 * .10) = **1.90**
  - $\boldsymbol{N}_{\textbf{effective}}$ = (30 * 10) / (1 + (9 * .10)) = **159** (not 300)
- **Level 2 Design Effect** = 1 + (5 * .05) = **1.25**
  - $\boldsymbol{N}_{\textbf{effective}}$ = (5 * 6) / (1 + (5 * .05)) = **24** (not 30)

Does a non-significant ICC mean you can ignore any levels of grouping?

- The effective sample size depends on BOTH the ICC and the number of people per group: As ICC goes UP and group size goes UP, the effective sample size goes DOWN
  - So there is NO VALUE OF ICC that is uniformly “safe” to ignore, not even 0, because…
- …Unconditional and conditional (after predictors) ICCs may differ
  - Reducing the residual variance often results in an increase in the random intercept variance, which then increases the conditional ICC
- So just do a multilevel analysis…
  - Even if “that’s not your question”… you still have to care that your data are clustered and model that dependency appropriately because of:
    - Effect of clustering on level-2 fixed effect SE’s
    - Potential for contextual effects of level-1 and level-2 predictors at level 3

Lecture 9: Three-Level Models for Clustered Longitudinal Data

- Decomposing Clustered Longitudinal Variation
- **Three-Level Model Specification**
- Time-Varying Predictors in Three-Level Models
- Clustered Longitudinal Model Examples in SAS, SPSS, and STATA

3-Level Models for Clustered Longitudinal Data: Empty Model

3-Level Clustered MLM

L1: $y_{tij} = \beta_{0ij} + e_{tij}$

L2: $\beta_{0ij} = \delta_{00j} + U_{0ij}$

L3: $\delta_{00j} = \gamma_{000} + V_{00j}$

Fixed Effects:

γ_{000} → grand mean intercept

Random Effects:

U_{0ij} → deviation for level 2

V_{00j} → deviation for level 3

Error:

e_{tij} → level-1 residual deviation

$$y_{tij} = (\gamma_{000}) + (V_{00j} + U_{0ij} + e_{tij})$$

Adding 8 Possible Effects of Predictors

X → L1, Z → L2, Q? → L3

L1: $y_{tij} = \beta_{0ij} + \beta_{1ij} X_{1tij} + e_{tij}$

L2: $\beta_{0ij} = \delta_{00j} + \delta_{01j} Z_{1ij} + U_{0ij}$

$\beta_{1ij} = \delta_{10j} + \delta_{11j} Z_{1ij} + U_{1ij}$

L3: $\delta_{00j} = \gamma_{000} + \gamma_{001} Q_{1j} + V_{00j}$ → intercept, int*Q

$\delta_{01j} = \gamma_{010} + \gamma_{011} Q_{1j} + V_{01j}$ → intercept*Z, Q*Z

$\delta_{10j} = \gamma_{100} + \gamma_{101} Q_{1j} + V_{10j}$ → intercept*X, Q*X

$\delta_{11j} = \gamma_{110} + \gamma_{111} Q_{1j} + [V_{11j}]$ → Z*X, Q*Z*X

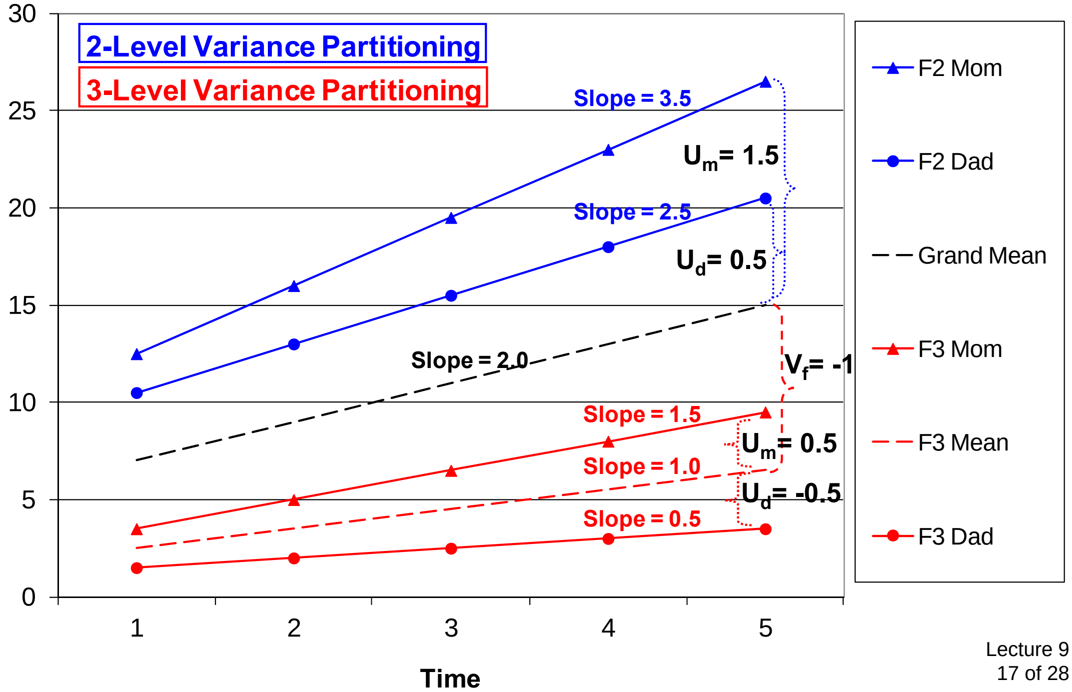
Means side = int + Q + Z + QZ + X + QX + ZX + QZX

Variances side = res(1) + int(3) + int(2) + X(3) + X(2) + [XZ(3)]

Meaning of Random Effects Example: Time in Persons in Families

- <u>Random effects associated with **level-1** predictors:</u>
  - → Does the effect of X_{tij} vary over **persons**? (each person needs own)
    - → This indicates a random slope for X_{tij} across persons (level 2)
    - → Specified on random statement for level 2
  - → Does the effect of X_{tij} vary over **families**? (each family needs own)
    - → This indicates a random slope for X_{tij} across families (level 3)
    - → Specified on random statement for level 3
- I recommend working your way UP for assessing random effects of level-1 predictors:
  - Is random slope for X_{tij} significant at level 2?
    If yes, how much of that slope variance is due to level 3?
  - However – it is possible to have effects of level-1 predictors as random at level 3 only (but unlikely)

Random Slopes of Level-1 X at Level 2 AND Level 3?



Meaning of Random Effects Example: Time in Persons in Families

- <u>Random effects can be associated with **level-2** predictors:</u>
  - → Does the effect of Z_{ij} vary over **families**? (each family needs own)
    - → This indicates a random slope for Z_{ij} across family (level 3)
    - → Specified on random statement for level 3

- <u>Not all slopes (effects of X_{tij}, Z_{ij}) need to be random!!</u>
  - More random effects → higher probability of SAS/SPSS/STATA vomit instead of results
    - "not positive definite" **G** matrix of random effects variances/covariances

ICCs for Slope Variance

- Once random slopes are included, one can compute ICCs for the slope variance, too:
  - Add random slope for linear time at level 2 and level 3?
  - ICC for proportion of slope variance that is at level 3:

$$ICC_{Slope} = \frac{\text{(L3 slope var)}}{\text{(L3 slope var) + (L2 slope var)}}$$

- Be careful when the model for the variances is 'uneven' across level 2 and level 3 – for example:
  - Random quadratic slope variance at level 2, but only random linear slope variance at level 3? The random linear slope variance does not have the same meaning at each level, so I wouldn't report its ICC.

Random Level 2: intercept, linear, quad	L2: Linear is when time=0
Random Level 3: intercept, linear	L3: Linear is all the time

Lecture 9: Three-Level Models for Clustered Longitudinal Data

- Decomposing Clustered Longitudinal Variation
- Three-Level Model Specification
- **Time-Varying Predictors in Three-Level Models**
- Clustered Longitudinal Model Examples in SAS, SPSS, and STATA

The HORROR of Time-Varying Predictors in 3-Level Models

- Remember separating between- and within-person effects? Now there are 3 sets of effects to worry about!
- Example: Effect of Stress on Wellbeing, measured over time within person within families:
  - **Level 1** (Time): During **Times** of more stress, people function less well than in times of less stress
  - **Level 2** (Person): **People** in the family who have more stress function less well than people in the family who have less stress
  - **Level 3** (Family): **Families** who have more stress function less well than families who have less stress
- Each of these effects may be of different magnitudes!
- (Convergence of level-2 predictors across level 3 also!)

Test Separate Effects Per Level Using PERSON-MC:

- **Level 1 (time):** *Time-varying stress relative to person mean*
  - → $WPstress_{tij} = Stress_{tij} – PersonMeanStress_{ij}$
  - → Within-Person effect ≠ 0?
  - → **Total** within-person effect of having more stress ***than usual*** ≠ 0?
- **Level 2 (person):** *Person mean stress relative to family*
  - → $WFstress_{ij} = PersonMeanStress_{ij} – FamilyMeanStress_{j}$
  - → Within-Family effect ≠ 0?
  - → **Total** effect of having more stress ***than other family members*** ≠ 0?
- **Level 3 (family):** *Family mean stress relative to all families (constant)*
  - → $BFstress_{j} = FamilyMeanStress_{j} – C$
  - → Between-Family effect ≠ 0?
  - → **Total** effect of having more stress ***than other families*** ≠ 0?

Test 3-Level Convergence of Effects Using Grand-MC:

- **Level 1 (time):** *Time-varying stress (relative to sample constant)*
  - → $\text{TVstress}_{tij} = \text{Stress}_{tij} - C$
  - → Within-Person effect ≠ 0?
  - → **Total** within-person effect of having more stress ***than usual*** ≠ 0?
- **Level 2 (person):** *Person mean stress (relative to sample constant)*
  - → $\text{BPstress}_{ij} = \text{PersonMeanStress}_{ij} - C$
  - → Within-Person and Within-Family effects ≠ ?
  - → **Contextual** effect of having more stress ***than other family members*** ≠ 0?
- **Level 3 (family):** *Family mean stress relative to all families (constant)*
  - → $\text{BFstress}_{j} = \text{FamilyMeanStress}_{j} - C$
  - → Within-Family and Between-Family effects ≠ ?
  - → **Contextual** effect of having more stress ***than other families*** ≠ 0?

Test Separate Effects Per Level Using GROUP-MC:

- **Level 1 (student):** *Student attitude relative to class mean*
  - → $\text{WCatt}_{ijk} = \text{Att}_{ijk} - \text{ClassMeanAtt}_{jk}$
  - → Between-Student/Within-Class effect ≠ 0?
  - → **Total** within-class effect of having better attitude ***than my class*** ≠ 0?
- **Level 2 (class):** *Class mean attitude relative to school*
  - → $\text{WSatt}_{ijk} = \text{ClassMeanAtt}_{jk} - \text{SchoolMeanAtt}_{k}$
  - → Between-Class/Within-School effect ≠ 0?
  - → **Total** within-school effect of having better attitude ***than other classes in the same school*** ≠ 0?
- **Level 3 (school):** *School mean attitude relative to all schools (constant)*
  - → $\text{BSatt}_{k} = \text{SchoolMeanAtt}_{k} - C$
  - → Between-School effect ≠ 0?
  - → **Total** effect of having better attitude ***than other schools*** ≠ 0?

Test 3-Level Convergence of Effects Using Grand-MC:

- **Level 1 (student):** *Student attitude (relative to sample constant)*
  → $WCatt_{ijk}$ = att_{ijk} – C
  → Between-Student, Within-Class effect ≠ 0?
  → **Total** within-class effect of having better attitude ***than my class*** ≠ 0?
- **Level 2 (class):** *Class mean attitude (relative to sample constant)*
  → $BCatt_{jk}$ = $ClassMeanAtt_{jk}$ – C
  → Within-Class and Within-School effects ≠ ?
  → **Contextual** effect of better attitude ***than other classes in school*** ≠ 0?
- **Level 3 (school):** *School mean attitude relative to all schools (constant)*
  → $BSatt_k$ = $SchoolMeanAtt_k$ – C
  → Within-School and Between-School effects ≠ ?
  → **Contextual** effect of having better attitude ***than other schools*** ≠ 0?

Implications of Removing Effects Under Each Centering Method

- **Person/Group-MC**: Removing terms means the effect at that level does not exist (= 0)
  - First remove L3 effect? Assume L3 BF effect = 0
    - *L1 effect = WP effect, L2 effect = WF effect*
  - Then remove L2 effect? Assume L2 WF effect = 0
    - *L1 effect = WP effect*
- **Grand-MC**: Removing terms means the effect at that level is equivalent to the effect at the level beneath it
  - First remove L3 effect? Assume L3 BF effect = L2 WF effect
    - *L1 effect = WP effect, L2 effect = 'smushed' WF and BF effects*
  - Then remove L2 effect? Assume L2 **BP** effect = L1 effect
    - *L1 effect = 'smushed' WP, WF, and BF effects*

Summary: 3-Level Models

- Estimating 3-level models requires no new concepts, but everything is just at an order of complexity higher:
  - Proportioning variance over 3 levels instead of 2 → 2+ ICCs
  - Random slope variance will come from term directly beneath:
    - Level-2 random slope comes from level-1 residual
    - Level-3 random slope comes from level-2 random slope (or residual)
  - Level-1 effects can be random over level 2, level 3, or both
    - ICCs can be computed for level-1 slopes that are random over both level-2 and level-3 (assuming the L2 and L3 models match)
    - Convergence of level-1 effects should be tested over levels 2 AND 3
  - Level-2 effects can be random over level 3
    - Convergence of level-2 effects should be tested over level 3
  - Level-3 effects cannot be random; no convergence testing needed
  - Phew….

Lecture 9: Three-Level Models for Clustered Longitudinal Data

- Decomposing Clustered Longitudinal Variation
- Three-Level Model Specification
- Time-Varying Predictors in Three-Level Models
- **Clustered Longitudinal Model Examples in SAS, SPSS, and STATA**