Three-Level (and Crossed) Random Effects Models

- Topics:
  - **Decomposing variation across three levels in clustered longitudinal data**
  - Unconditional (time only) model specification
  - Conditional (other predictors) model specification
  - Other kinds of three-level and crossed designs

What determines the number of levels?

- **Answer: the model for the outcome variance ONLY**
- How many dimensions of sampling in the outcome?
  - Time within person → 2-level model
  - Time within person within family → 3-level model
  - Time within person within family within country → 4-level model
  - Sampling dimensions may also be crossed instead of nested, or may be modeled with fixed effects if the # units is small

- Need at least one pile of variance per dimension (for 3 levels, that's 2 sets of random effects and a residual)
  - Include 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 exists in its relevant sampling dimension

Empty Means, 3-Level Random Intercept Model

Notation: t = level-1 time, i = level-2 person, j = level-3 group

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

Residual = time-specific deviation from **person's** predicted outcome

Level 2: $\beta_{0ij} = \delta_{00j} + U_{0ij}$

Person Random Intercept = person-specific deviation from **group's** predicted outcome

Level 3: $\delta_{00j} = \gamma_{000} + V_{00j}$

Fixed Intercept =grand mean (because no predictors yet)

Group Random Intercept = group-specific deviation from fixed intercept

3 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

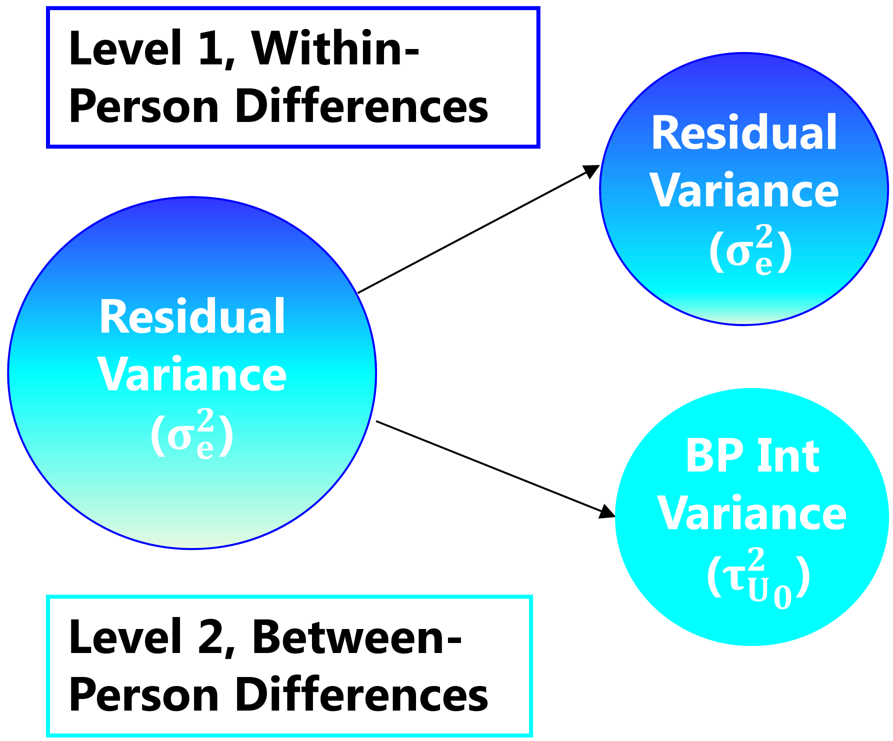
- Level-1 Variance of e_{tij} → σ^2_e
- Level-2 Variance of U_{0ij} → $\tau^2_{U_0}$
- Level-3 Variance of V_{00j} → $\tau^2_{V_{00}}$

Composite equation:

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

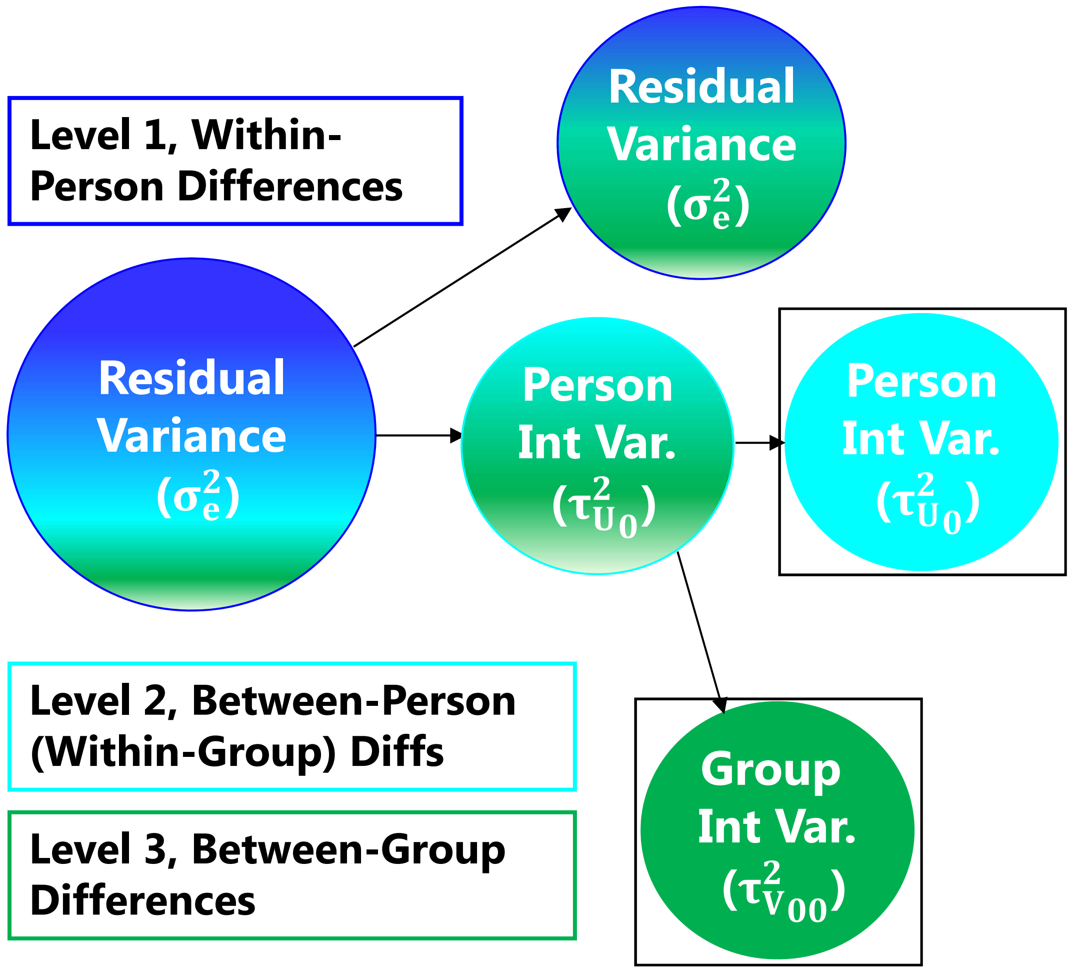
2-Level Random Intercept Model

- Where does each kind of person dependency go? Into a new random effects variance component (or "pile" of variance):
- Let's start with an empty means, random intercept 2-level model for time within person:



3-Level Random Intercept Model

- Now let's see what happens in an empty means, random intercept 3-level model of time within person within groups:



ICCs in a 3-Level Random Intercept Model Example: Time within Person within Group

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

- $\mathrm{ICC_{L2}} = \frac{\mathrm{Between-Person}}{\mathrm{Total}} = \frac{\mathrm{L3+L2}}{\mathrm{L3+L2+L1}} = \frac{\tau^2_{V_{00}} + \tau^2_{U_0}}{\tau^2_{V_{00}} + \tau^2_{U_0} + \sigma^2_e}$

  → This ICC expresses similarity of occasions from same person (and by definition, from the same group) → of the ***total variation in Y***, how much of it is ***between persons, or not due to time***?

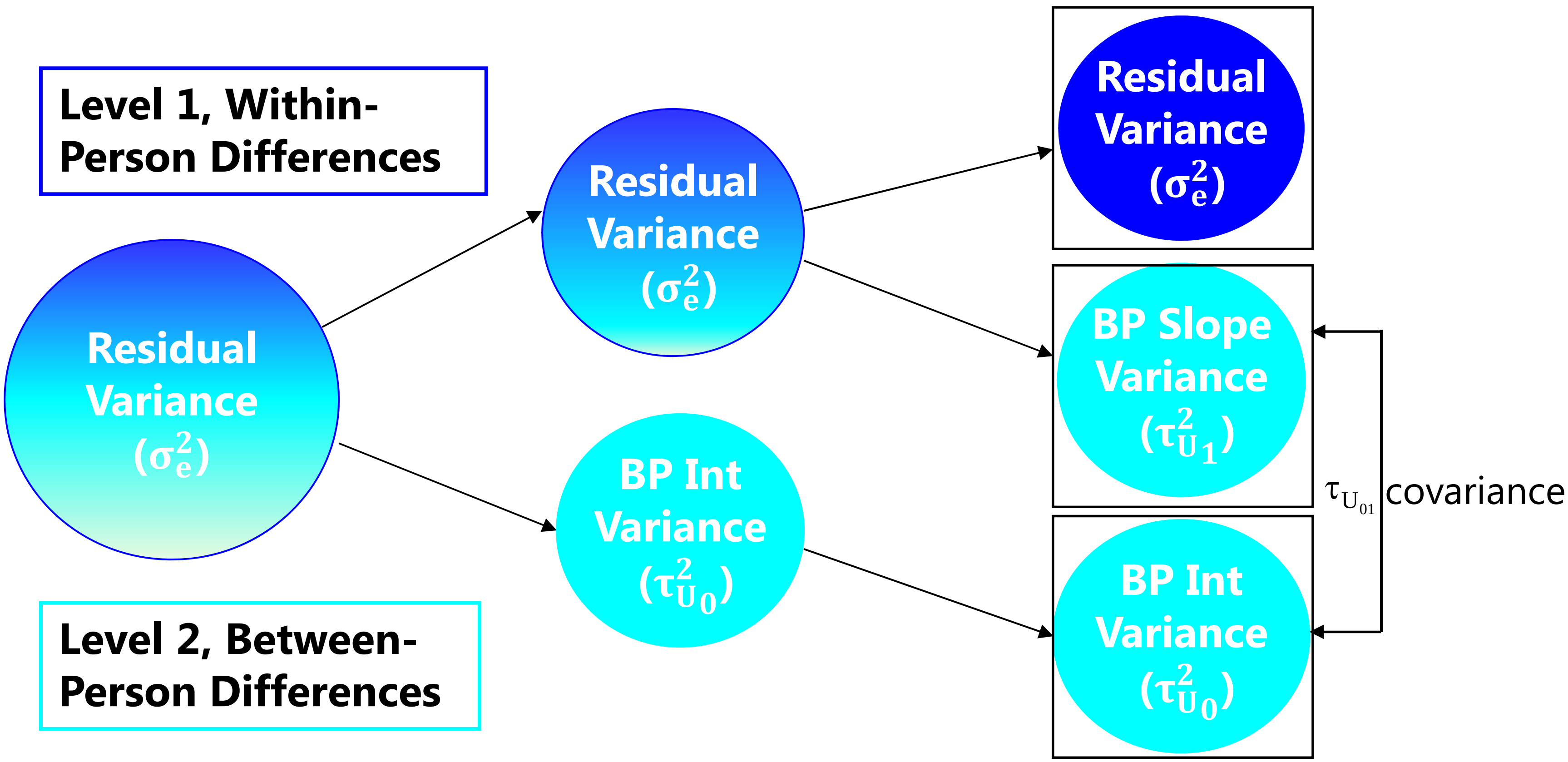
- **ICC for level 3 relative to level 2 (ignoring level 1):**

- $\mathrm{ICC_{L3}} = \frac{\mathrm{Between-Group}}{\mathrm{Between-Person}} = \frac{\mathrm{L3}}{\mathrm{L3+L2}} = \frac{\tau^2_{V_{00}}}{\tau^2_{V_{00}} + \tau^2_{U_0}}$

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

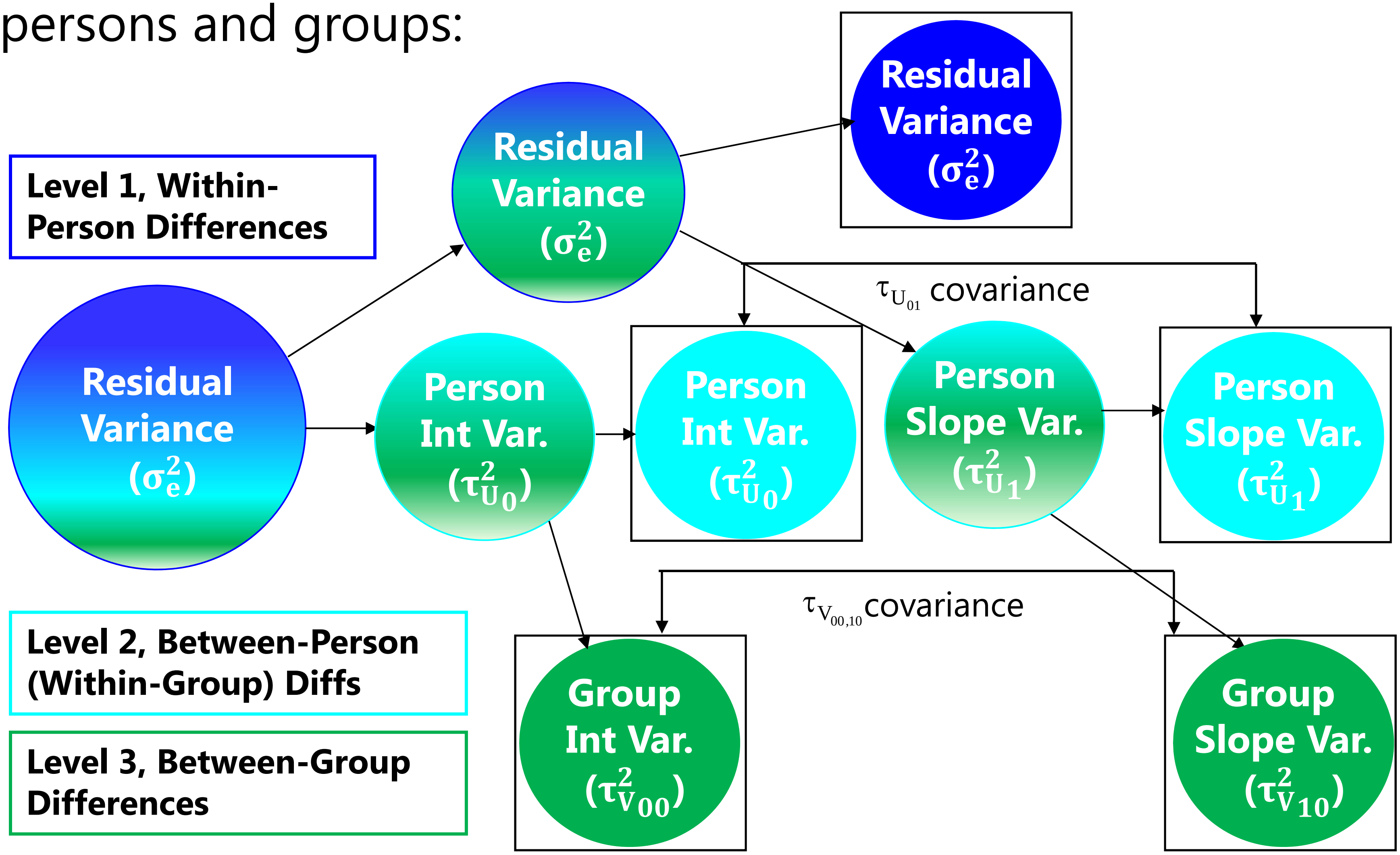
2-Level Random Slope Model

- What about time? After adding fixed effects of time, we can add random effects of time over persons in a 2-level model:

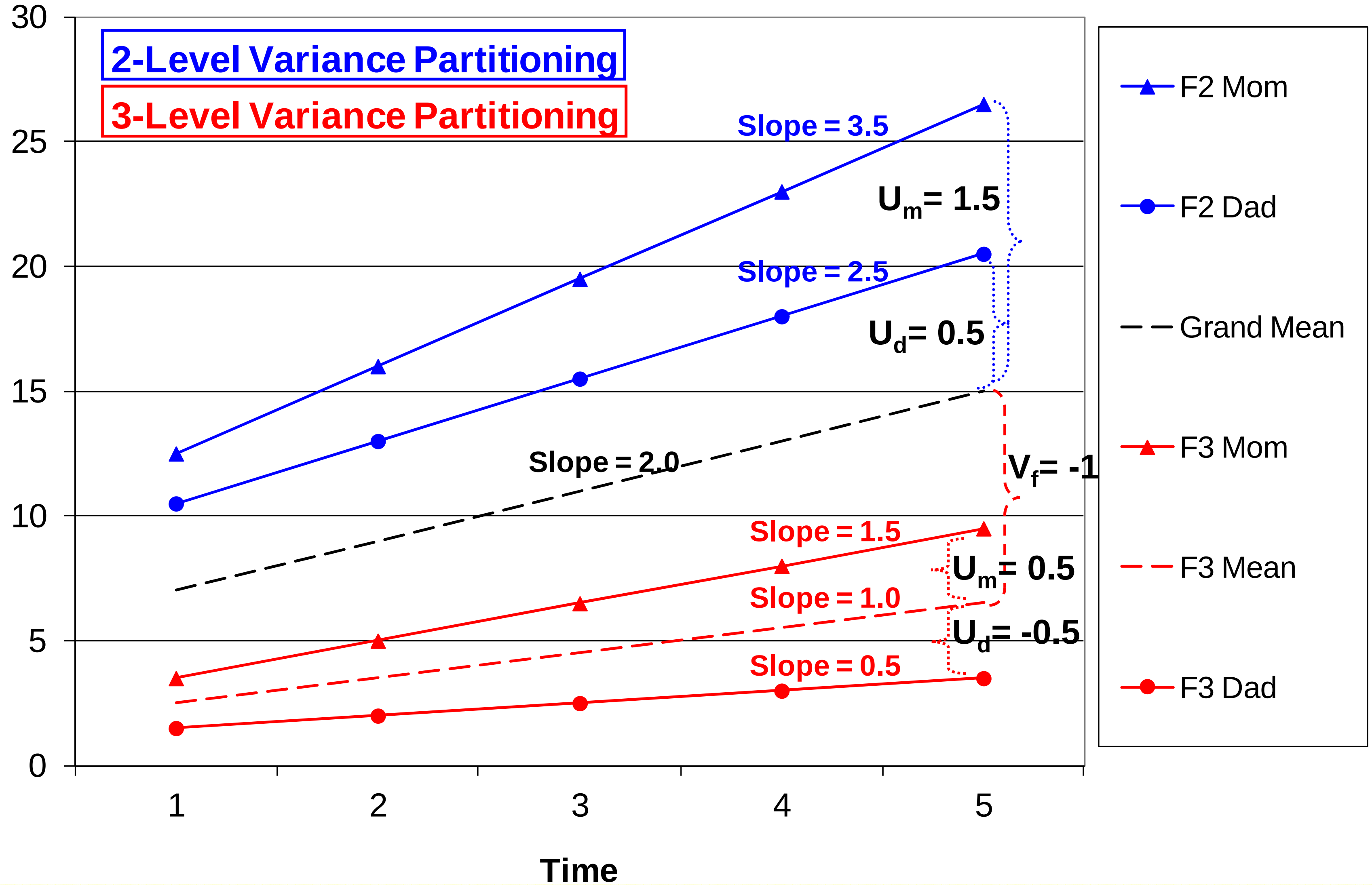


3-Level Random Slope Model

- In a 3-level model, we can have random effects of time over persons and groups:



Random Time Slopes at both Levels 2 AND 3? An example with family as group:



3-Level Random Time Slope Model

Notation: t = level-1 time, i = level-2 person, j = level-3 group

Level 1: $y_{tij} = \beta_{0ij} + \beta_{1ij}(Time_{tij}) + e_{tij}$

Residual = time-specific deviation from **person's** predicted growth line (σ_e^2)

Level 2: $\beta_{0ij} = \delta_{00j} + U_{0ij}$

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

Person Random Intercept and Slope = person-specific deviations from **group's** predicted intercept, slope ($\tau_{U_0}^2$, $\tau_{U_1}^2$, $\tau_{U_{01}}$)

Level 3: $\delta_{00j} = \gamma_{000} + V_{00j}$

$\delta_{10j} = \gamma_{100} + V_{10j}$

Group Random Intercept and Slope = group-specific deviations from fixed intercept, slope ($\tau_{V_{00}}^2$, $\tau_{V_{10}}^2$, $\tau_{V_{00,10}}$)

Fixed Intercept, Fixed Linear Time Slope

Composite equation (9 parameters):

$$y_{tij} = (\gamma_{000} + V_{00j} + U_{0ij}) + (\gamma_{100} + V_{10j} + U_{1ij})(Time_{tij}) + e_{tij}$$

ICCs for Random Intercepts and Slopes

- Once random slopes are included at both level-3 and level-2, ICCs can be computed for the random intercepts and slopes specifically (which would be the level-3 type of ICC)

$$\mathrm{ICC_{Int}} = \frac{\mathrm{Between-Group}}{\mathrm{Between-Person}} = \frac{\mathrm{L3\ Int}}{\mathrm{L3\ Int + L2\ Int}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{00}}}}{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}}}$$

$$\mathrm{ICC_{Slope}} = \frac{\mathrm{Between-Group}}{\mathrm{Between-Person}} = \frac{\mathrm{L3\ Slope}}{\mathrm{L3\ Slope + L2\ Slope}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{10}}}}{\boldsymbol{\tau}^2_{\mathbf{V_{10}}} + \boldsymbol{\tau}^2_{\mathbf{U_1}}}$$

- Can be computed for any level-1 slope that is random at both levels (e.g., linear and quadratic time, time-varying predictors)
- Be careful when the model is uneven across levels, though

$$\frac{\text{Random Level 2: int, linear, quad}}{\text{Random Level 3: int, linear}} \rightarrow \frac{\text{Linear is when time} = 0}{\text{Linear is at any occasion}}$$

More on Random Slopes in 3-Level Models

- Any level-1 predictor can have a random effect over level 2, level 3, or over both levels, but I recommend working your way **UP the higher levels** for assessing random effects…
  - e.g., Does the effect of time vary over persons?
  - If so, does the effect of time vary over groups, too? → Is there a commonality in how people from the same group change over time?
- … because random effects at level 3 only are possible but unlikely (e.g., means everyone in the group changes the same)

- Level-2 predictors can also have random effects over level 3
  - e.g., Does the effect of a person characteristic vary over groups?
- Level-1, level-2, and level-1 by level-2 cross-level interactions can all have random effects over level 3, too
  - But tread carefully! The more random effects you have, the more likely you are to have convergence problems ("not positive definite")

Conditional Model Specification

- Remember separating between- and within-person effects? <u>Now there are 3 potential effects for any level-1 predictor!</u>
  - Example: Effect of stress on wellbeing, both measured over time within person within families:
  - **Level 1** (Time): During **Times** of more stress, people have lower (time-specific) wellbeing than in times of less stress
  - **Level 2** (Person): **People** in the family who have more stress have lower (person average) wellbeing than people in the family who have less stress
  - **Level 3** (Family): **Families** who have more stress have lower (family average) wellbeing than families who have less stress

- <u>2 potential effects for any level-2 predictor, also</u>
  - Example: Effect of baseline level of person coping skills in same design:
  - **Level 2** (Person): **People** in the family who cope better have better (person average) wellbeing than people in the family who cope worse
  - **Level 3** (Family): **Families** who cope better have better (family average) wellbeing than families who cope worse

Separate Total Effects Per Level Using Person/Group-Mean-Centering

- **Level 1 (Time):** *Time-varying stress relative to person mean*
  - → $WPstress_{tij}$ = $Stress_{tij}$ – $PersonMeanStress_{ij}$
  - → Directly tests if within-person effect ≠ 0?
  - → **Total** within-person effect of having more stress ***than usual*** ≠ 0?

- **Level 2 (Person):** *Person mean stress relative to family*
  - → $WFstress_{ij}$ = $PersonMeanStress_{ij}$ – $FamilyMeanStress_{j}$
  - → Directly tests if 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 (from constant)*
  - → $BFstress_{j}$ = $FamilyMeanStress_{j}$ – C
  - → Directly tests if between-family effect ≠ 0?
  - → **Total** effect of having more stress ***than other families*** ≠ 0?

Separate Total Effects Per Level Using Person/Group-Mean-Centering

Notation: t = level-1 time, i = level-2 person, j = level-3 group
PM = person mean, FM = family mean, C = centering constant

Level 1: $y_{tij} = \beta_{0ij} + \beta_{1ij}(Time_{tij}) + \beta_{2ij}(Stress_{tij} - PMstress_{ij}) + e_{tij}$

Level 2: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(PMstress_{ij} - FMstress_{j}) + U_{0ij}$

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

$\beta_{2ij} = \delta_{20j} + (U_{2ij})$

Level 3: $\delta_{00j} = \gamma_{000} + \gamma_{001}(FMstress_{j} - C) + V_{00j}$

$\delta_{01j} = \gamma_{010} + (V_{01j})$

$\delta_{10j} = \gamma_{100} + V_{10j}$

$\delta_{20j} = \gamma_{200} + (V_{20j})$

Fixed intercept, Between-family stress main effect

Within-family stress main effect

Time main effect

Within-person stress main effect

Contextual Effects Per Level Using Grand-Mean-Centering

- **Level 1 (Time):** *Time-varying stress (relative to sample constant)*
  - → TVstress_{tij} = Stress_{tij} – C
  - → Directly tests if 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)*
  - → BPstress_{ij} = $\text{PersonMeanStress}_{ij}$ – C
  - → Directly tests if 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 (from constant)*
  - → BFstress_{j} = $\text{FamilyMeanStress}_{j}$ – C
  - → Directly tests if within-family and between-family effects ≠ ?
  - → **Contextual** effect of having more stress ***than other families*** ≠ 0?

Contextual Effects Per Level Using Grand-Mean-Centering

Notation: t = level-1 time, i = level-2 person, j = level-3 group
PM = person mean, FM = family mean, C = centering constant

Level 1: $y_{tij} = \beta_{0ij} + \beta_{1ij}(Time_{tij}) + \beta_{2ij}(Stress_{tij} - C) + e_{tij}$

Level 2: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(PMstress_{ij} - C) + U_{0ij}$

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

$\beta_{2ij} = \delta_{20j} + (U_{2ij})$

Level 3: $\delta_{00j} = \gamma_{000} + \gamma_{001}(FMstress_{j} - C) + V_{00j}$

$\delta_{01j} = \gamma_{010} + (V_{01j})$

$\delta_{10j} = \gamma_{100} + V_{10j}$

$\delta_{20j} = \gamma_{200} + (V_{20j})$

Fixed intercept, <u>Contextual</u> family stress main effect

<u>Contextual</u> within-family stress main effect

Time main effect

Within-person stress main effect

What does it mean to omit higher-level effects under each centering method?

- **Person-MC**: Removing terms means the effect at that level <u>does not exist</u> (= 0)
  - Remove L3 effect? Assume L3 Between-Family effect = 0
    - *L1 effect = Within-Person effect, L2 effect = Within-Family effect*
  - Then remove L2 effect? Assume L2 Within-Family effect = 0
    - *L1 effect = Within-Person effect*

- **Grand-MC**: Removing terms means the effect at that level is <u>equivalent to</u> the effect at the level beneath it
  - Remove L3 effect? Assume L3 Between-Family = L2 Within-Family effect
    - *L1 effect = Within-Person effect, L2 effect = 'smushed' WF and BF effects*
  - Then remove L2 effect? Assume L2 Between-Person effect = L1 effect
    - *L1 'smushed' = Within-Person, Within-Family, and Between-Family effects*

Interactions belong at each level, too…

- Example: Is the effect of stress on wellbeing moderated by time-invariant person coping? Using person/group-MC…
- **Stress Effects**
  - **Level 1 (Time):** $WPstress_{tij} = Stress_{tij} - PersonMeanStress_{ij}$
  - **Level 2 (Person):** $WFstress_{ij} = PersonMeanStress_{ij} - FamilyMeanStress_{j}$
  - **Level 3 (Family):** $BFstress_{j} = FamilyMeanStress_{j} - C$
- **Coping Effects**
  - **Level 2 (Person):** $WFcope_{ij} = Cope_{ij} - FamilyMeanCope_{j}$
  - **Level 3 (Family):** $BFcope_{j} = FamilyMeanCope_{j} - C$
- **Interaction Effects**
- With level 1 stress: $WPstress_{tij} * WFcope_{ij}$, $WPstress_{tij} * BFcope_{j}$
- With level 2 stress: $WFstress_{ij} * WFcope_{ij}$, ($WFstress_{ij} * BFcope_{j}$)
- With level 3 stress: $BFstress_{j} * BFcope_{j}$, ($BFstress_{j} * WFcope_{ij}$)

Interactions belong at each level, too…

Notation: t = level-1 time, i = level-2 person, j = level-3 group
PM = person mean, FM = family mean, C = centering constant

Level 1: $y_{tij} = \beta_{0ij} + \beta_{1ij}(Time_{tij}) + \beta_{2ij}(Stress_{tij} - PMstress_{ij}) + e_{tij}$

Level 2:

$$\beta_{0ij} = \delta_{00j} + \delta_{01j}(PMstress_{ij} - FMstress_{j}) + \delta_{02j}(Cope_{ij} - FMcope_{j}) + \delta_{03j}(PMstress_{ij} - FMstress_{j})(Cope_{ij} - FMcope_{j}) + U_{0ij}$$

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

$$\beta_{2ij} = \delta_{20j} + \delta_{21j}(Cope_{ij} - FMcope_{j}) + (U_{2ij})$$

Level 3:

$$\delta_{00j} = \gamma_{000} + \gamma_{001}(FMstress_{j} - C) + \gamma_{002}(FMcope_{j} - C) + \gamma_{003}(FMstress_{j} - C)(FMcope_{j} - C) + V_{00j}$$

$$\delta_{01j} = \gamma_{010} + (V_{01j}) \quad \delta_{02j} = \gamma_{020} + (V_{02j}) \quad \delta_{03j} = \gamma_{030} + (V_{03j})$$

$$\delta_{10j} = \gamma_{100} + V_{10j}$$

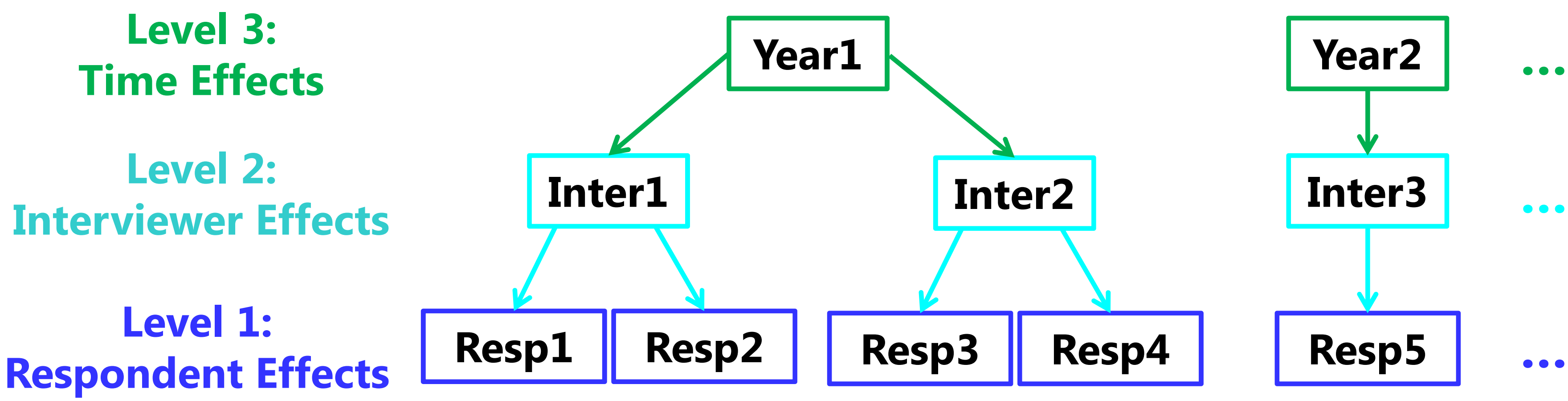
$$\delta_{20j} = \gamma_{200} + \gamma_{202}(FMcope_{j} - C) + (V_{20j}) \quad \delta_{21j} = \gamma_{210} + (V_{21j})$$

Summary: Clustered Longitudinal 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….

Other 3-Level Designs

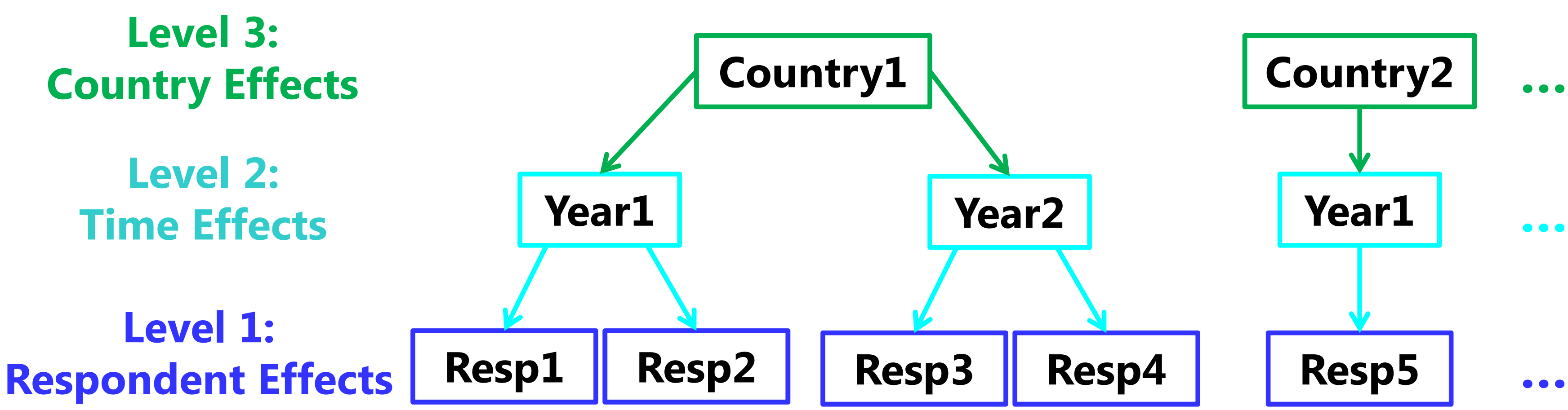
- The sampling design for the outcome (not the predictors) dictates what your levels will be, **so time may not always be level 1**
- Example: Predicting answer compliance in respondents nested in interviewers, collected over several years (<u>all different people</u>)



- Based on the sampling of time, time may be modeled…
  - As fixed effects in the model for the means → 2-level model instead
    - Best to use dummy codes for time if few occasions OR no time-level predictors of interest
  - As a random effect in the model for the variance → 3-level model
    - Then differences in compliance rates over time can be predicted by time-level predictors

Other 3-Level Designs

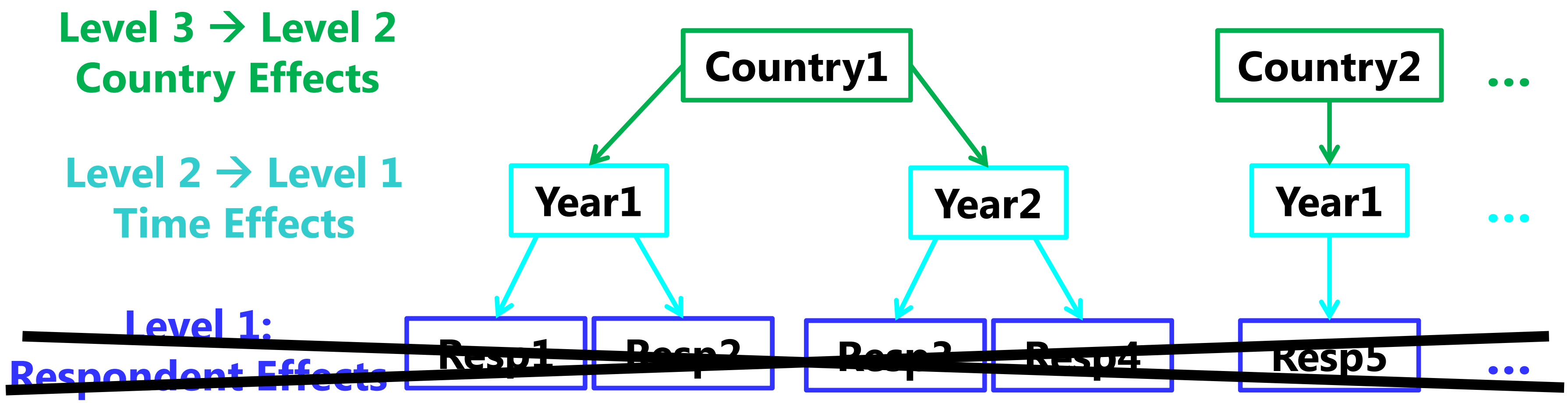
- Another example: Predicting **time-specific respondent outcomes** for people nested in countries, collected over several years (<u>all different people</u>, but the <u>same countries</u> measured over time)



- Before including any fixed effects of time, country and time are actually crossed, not nested as shown here
  - Are nested after controlling for which occasion is which via fixed effects (using dummy codes per mean or a time trend that describes the means)
  - Time is still a level because not all countries change the same way

3-Level Designs: Predictors vs. Outcomes

- Same example: What if, instead of respondent outcomes, we wanted to predict **time-varying country outcomes**?

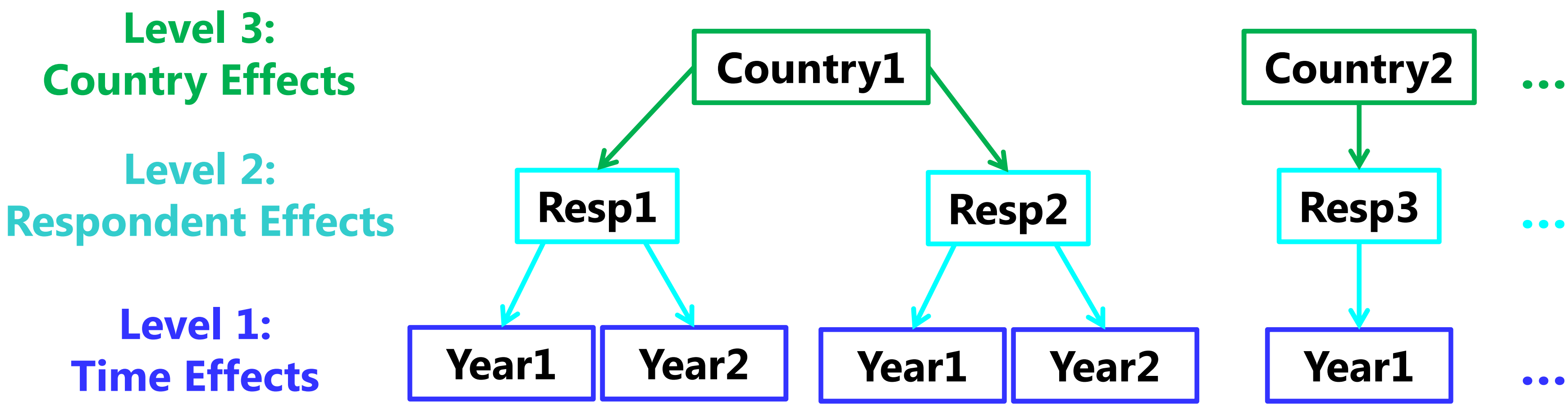


Because the outcome was measured at level 2 (country per time):

- Respondents are no longer a level at all (no outcomes for them)
- So there is nothing for respondent predictors to do, except at higher levels
  - **Time-specific averages** of respondent predictors → time-level outcome variation
  - **Across time, country averages** of respondent predictors → country-level outcome variation

Other 3-Level Designs: Predictors by Level

- Last example: Predicting **time-specific respondent outcomes** for people nested in countries, collected over several years (<u>all same people </u>and <u>same countries</u> are measured over time)



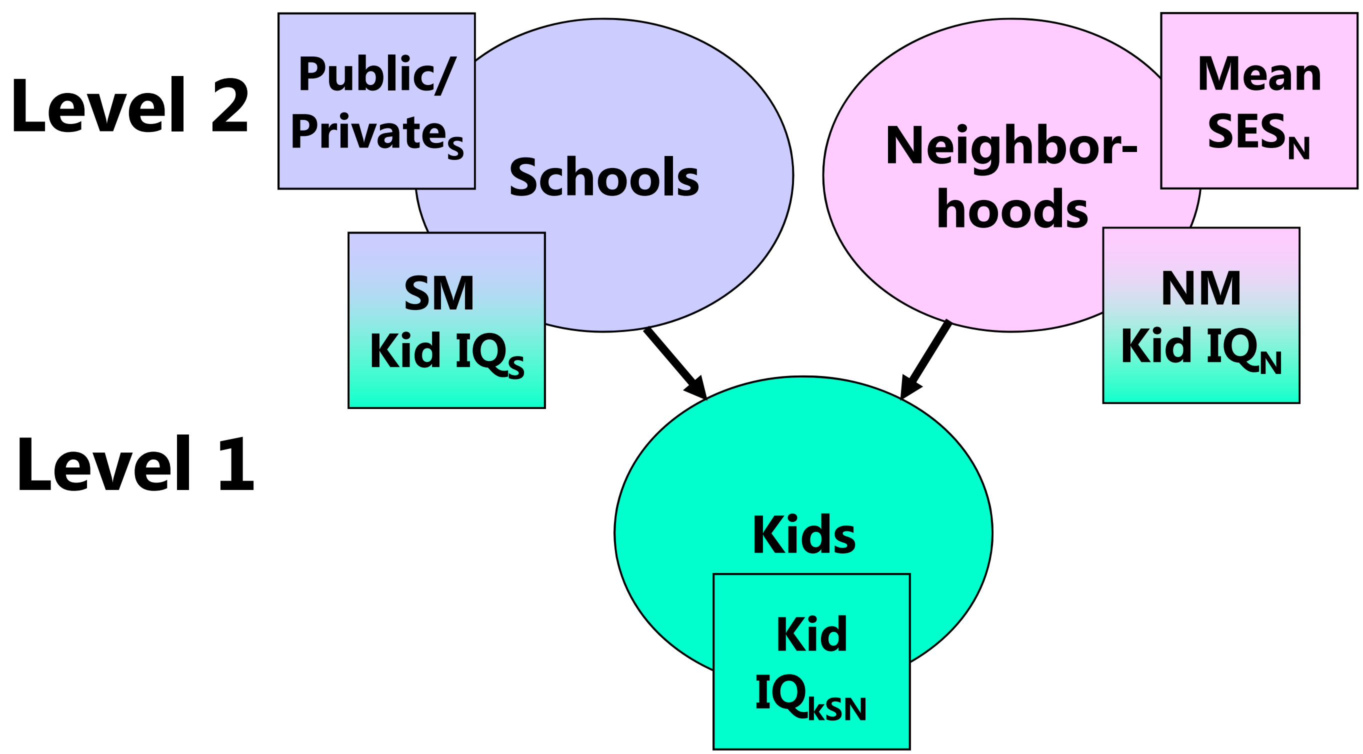
- Country predictors can be included at level 3 only (no random effects)
- Person predictors should be included at levels 2 and 3 (+random over 3)
- What about effects of <u>time-varying predictors</u>?
  - For <u>People</u>: effects should be included at all 3 levels (+random over 2 and 3)
  - For <u>Countries</u>: effects are only possible at levels 1 and 3 (+random over 3)

More Complex Multilevel Designs

- Multilevel models are specified based on the relevant dimensions by which observations differ each other, and how the units are organized
- Two-level models have at least two piles of variance, in which level-1 units are nested within level-2 units:
  - Longitudinal Data: Time nested within Persons
  - Students nested within Classes
- Three-level models have at least three piles of variance, in which level-2 units are nested within level-3 units:
  - Time nested within Persons within Families
  - Student nested within Classes within Schools
- In other designs, multiple sources of systematic variation may be present, but the sampling may be crossed instead…
  - Same idea as crossed random effects (i.e., as for persons and items), but these are known as "cross-classified" models in the clustered data world
  - Here are a few examples on when this might happen…

Kids, Schools, and Neighborhoods

- Kids are nested within schools AND within neighborhoods
- Not all kids from same neighborhood live in same school, so schools and neighborhoods are crossed at level 2
- Can include predictors for each source of variation



Kids, Schools, and Neighborhoods

$y_{kSN} = \gamma_{000}$ → *fixed intercept (all x's = 0)*

$+ \gamma_{010}(Private_S) + \gamma_{020}(SMIQ_S)$ → *school effects*

$+ \gamma_{001}(SES_N) + \gamma_{002}(NMIQ_N)$ → *neighborhood effects*

$+ \gamma_{100}(KidIQ_{kSN})$ → *kid effects*

$+ U_{0S0}$ → *random effect of school*

$+ U_{00N}$ → *random effect of neighborhood*

$+ e_{kSN}$ → *residual kid-to-kid variation*

Time, Kids, and Classrooms

- Kids are nested within classroom at each occasion…
- But kids move into different classrooms across time…
  - So Time is nested within Kid, Kid is crossed with Classroom
- How to model a time-varying random classroom effect?
  - This is the basis of so-called "value-added models"

- (At least) Two options:

  - Temporary classroom effect: Random effect for classroom that operates only at the point when the kid is in that classroom
    - e.g., Classroom effect ← teacher bias
    - Once out of classroom, effect is no longer present

  - Cumulative classroom effect: Random effect for classroom that operates at the point when the kid is in that classroom forwards
    - e.g., Classroom effect ← differential learning
    - Effect stays with the kid in the future

More on Cross-Classified Models

- In crossed models, lower-level predictors can have random slopes of over higher levels AND random slopes of the other crossed factor at the same level
  - Example: Kids, Schools, and Neighborhoods (data permitting)
    - Kid effects could vary over schools AND/OR neighborhoods
    - School effects could vary over neighborhoods (both level 2)
    - Neighborhood effects could vary over schools (both level 2)

- Concerns about smushing still apply over both level-2's
  - Separate contextual effects of kid predictors for schools and neighborhoods (e.g., after controlling for how smart you are, it matters incrementally whether you go to a smart school AND if you live in a neighborhood with smart kids)

Summary: Nested or Crossed Models

- Dimensions of sampling can result in systematic differences (i.e., dependency) that needs to be accounted for in the model for the variances
  - Sometimes this dependency is from nested sampling
  - Sometimes this dependency is from crossed sampling
- Multilevel models that include crossed random effects (or cross-classified models):
  - Can address this dependency (statistical motivation)
  - Can quantify and predict the amount of variation due to each source (substantive motivation)
  - Can include simultaneous hypothesis tests pertaining to each source of variation (substantive motivation)