

Applied Multilevel Models for Longitudinal and Clustered Data

QIPSR Workshop at the University of Kentucky
5/14/2013 – 5/16/2013

Presented by:

Dr. Lesa Hoffman
Associate Professor, Department of Psychology
University of Nebraska-Lincoln

Materials will be available for download at:
<http://psych.unl.edu/hoffman/Sheets/Longitudinal.htm>

For further resources and online course materials, please visit:
<http://psych.unl.edu/hoffman/HomePage.htm>

COURSE OVERVIEW

Multilevel models are known by many synonyms (i.e., hierarchical linear models, general linear mixed models). The defining feature of these models is their capacity to provide quantification and prediction of random variance due to multiple sampling dimensions (across occasions, persons, or groups). Multilevel models offer many advantages for analyzing longitudinal data, such as flexible ways for modeling individual differences in change, the examination of time-invariant or time-varying predictor effects, and the use of all available complete observations. Multilevel models are also useful in analyzing clustered data (e.g., persons nested in groups), in which one wishes to examine predictors pertaining to individuals or to groups. This workshop will serve as an applied introduction to multilevel models, beginning with longitudinal data, continuing onto clustered data, and concluding with clustered longitudinal data. Although generalized multilevel models are also available, this workshop will focus on general multilevel models (i.e., for conditionally normally distributed outcomes).

The first day will be spent reviewing general linear models (e.g., regression, ANOVA) and then introducing the multilevel model for change over time. The second day will be spent two-level conditional (predictor) models for longitudinal data, including both time-invariant and time-varying predictors. The third day will be spent examining two-level conditional models for clustered data, and then three-level models for clustered longitudinal data. The primary software package utilized for instruction will be STATA, but examples using SPSS and SAS will also be provided. Participants should be familiar with the general linear model, but no prior experience with multilevel models or knowledge of advanced mathematics (e.g., matrix algebra) is assumed.

TENTATIVE SCHEDULE OF TOPICS

| Day | Topic |
|---------|--|
| 5/14 AM | <u>Lecture 1: Introduction to Multilevel Models</u> <ul style="list-style-type: none"> • What is multilevel modeling? • Concepts in longitudinal data • From between-person to within-person models • Kinds of ANOVAs for longitudinal data |
| 5/14 PM | <u>Lecture 2: Describing Within-Person Change in Longitudinal Data</u> <ul style="list-style-type: none"> • Multilevel modeling notation and terminology • Fixed and random effects of linear time • Predicted variances and covariances from random slopes • Dependency and effect size in random effects models • Describing nonlinear change (polynomials, <i>piecewise</i>, <i>nonlinear</i>) • Fun with likelihood estimation and model comparisons • Data example 2 (data, syntax, and output provided) |
| 5/15 AM | <u>Lecture 3: Time-Invariant Predictors in Longitudinal Models</u> <ul style="list-style-type: none"> • Missing predictors in MLM • Effects of time-invariant predictors • Fixed, systematically varying, and random level-1 effects • Model building strategies and assessing significance • Data example 3 (data, syntax, and output provided) |
| 5/15 PM | <u>Lecture 4: Time-Varying Predictors in Longitudinal Models</u> <ul style="list-style-type: none"> • Time-varying predictors that fluctuate over time • Person-Mean-Centering (PMC) • Data example 4 (data, syntax, and output provided) • <i>Grand-Mean-Centering (GMC)</i> • <i>Model extensions under Person-MC vs. Grand-MC</i> • <i>Time-varying predictors that change over time</i> |
| 5/16 AM | <u>Lecture 5: Two-Level Models for Clustered Data</u> <ul style="list-style-type: none"> • Fixed vs. random effects for modeling clustered data • ICC and design effects in clustered data • Group-Mean-Centering vs. Grand-Mean Centering • Data example 5 (syntax and output provided only) • <i>Model extensions under Group-MC and Grand-MC</i> |
| 5/16 PM | <u>Lecture 6: Three-Level Models for Clustered Longitudinal Data</u> <ul style="list-style-type: none"> • Decomposing variation across three levels in clustered longitudinal data • Unconditional (time only) model specification • Data example 6 (syntax and output provided only) • <i>Conditional (other predictors) model specification</i> • <i>Other kinds of three-level designs</i> |

Topics in italics may be skipped due to time constraints.

Introduction to Multilevel Models

- Topics:
 - **What is multilevel modeling?**
 - Concepts in longitudinal data
 - From between-person to within-person models
 - Kinds of ANOVAs for longitudinal data

What is a Multilevel Model (MLM)?

- Same as other terms you have heard of:
 - **General Linear Mixed Model** (if you are from statistics)
 - *Mixed* = Fixed and Random effects
 - **Random Coefficients Model** (also if you are from statistics)
 - Random coefficients = Random effects
 - **Hierarchical Linear Model** (if you are from education)
 - Not the same as hierarchical regression
- Special cases of MLM:
 - Random Effects ANOVA or Repeated Measures ANOVA
 - (Latent) Growth Curve Model (where “Latent” → SEM)
 - Within-Person Fluctuation Model (e.g., for daily diary data)
 - Clustered/Nested Observations Model (e.g., for kids in schools)
 - Cross-Classified Models (e.g., “value-added” models)

The Two Sides of Any Model

- **Model for the Means:**

- Aka **Fixed Effects**, Structural Part of Model
- What you are used to **caring about for testing hypotheses**
- How the expected outcome for a given observation varies as a function of values on predictor variables

- **Model for the Variances:**

- Aka **Random Effects and Residuals**, Stochastic Part of Model
- What you are used to **making assumptions about** instead
- How residuals are distributed and related across observations (persons, groups, time, etc.) → these relationships are called "dependency" and ***this is the primary way that multilevel models differ from general linear models (e.g., regression)***

Dimensions for Organizing Models

- Outcome type: General (normal) vs. Generalized (not normal)
 - Dimensions of sampling: One (so one variance term per outcome) vs. **Multiple** (so multiple variance terms per outcome) → **OUR WORLD**
 - **General Linear Models**: conditionally normal outcome distribution, **fixed effects** (identity link; only one dimension of sampling)
 - **Generalized Linear Models**: **any conditional outcome distribution**, **fixed** effects through **link functions**, no random effects (one dimension)
 - **General Linear Mixed Models**: conditionally normal outcome distribution, **fixed and random effects** (identity link, but multiple sampling dimensions)
 - **Generalized Linear Mixed Models**: **any conditional outcome distribution**, **fixed and random effects** through **link functions** (multiple dimensions)
- "Linear" means the fixed effects predict the *link-transformed* DV in a linear combination of (effect*predictor) + (effect*predictor)...

Note: Least Squares is only for GLM

How We Will Learn MLM

- “Levels” are defined by the context of a study
 - **Level** \approx **a dimension of sampling** (can be nested or crossed)
- We will start with MLM for longitudinal data...
 - Level 1 = variation over time, Level 2 = variation over persons
 - More complex case because of the time dimension
- ...We will follow with MLM for clustered data...
 - Level 1 = variation over persons, Level 2 = variation over groups
- ... and conclude with MLM for clustered+longitudinal data
 - Time (Level 1) within persons (Level 2) within groups (Level 3)
 - Persons (Level 1) within occasions (Level 2) within groups (Level 3)

What can MLM do for you?

1. **Model dependency across observations**
 - Longitudinal, clustered, and/or cross-classified data? No problem!
 - Tailor your model of sources of correlation to your data
2. **Include categorical or continuous predictors at any level**
 - Time-varying, person-level, group-level predictors for each variance
 - Explore reasons for dependency, don't just control for dependency
3. **Does not require same data structure for each person**
 - Unbalanced or missing data? No problem!
4. **You already know how (or you will soon)!**
 - Use SPSS Mixed, SAS Mixed, **Stata**, Mplus, R, HLM, MLwiN...
 - What's an intercept? What's a slope? What's a pile of variance?

I. Model Dependency

- Sources of dependency depend on the sources of **variation** created by your sampling design: residuals for outcomes from the same unit are likely to be related, which violates the GLM “independence” assumption
- **“Levels” for dependency** = “levels of random effects”
 - Sampling dimensions can be **nested**
 - e.g., time within person, person within group, school within country
 - If you can't figure out the direction of your nesting structure, odds are good you have a **crossed sampling design** instead
 - e.g., persons crossed with items, raters crossed with targets
 - To have a “level”, there must be random outcome variation due to sampling that remains after including the model's fixed effects
 - e.g., treatment vs. control does not create another level of “group”

Dependency comes from...

- Mean differences across sampling units (persons, groups)
 - Creates constant dependency over time (or persons)
 - Will be represented by a random intercept in our models
- Individual/group differences in effects of predictors
 - Longitudinal: individual differences in growth, stress reactivity
 - Clustered: group differences in slopes of person predictors
 - Creates non-constant dependency, the size of which depends on the value of the predictor at each occasion or for each person
 - Will be represented by random slopes in our models
- Longitudinal data: non-constant within-person correlation for unknown reasons (time-specific autocorrelation)
 - Can add other patterns of correlation as needed for this (AR, TOEP)

Why care about dependency?

- In other words, what happens if we have the wrong model for the variances (assume independence instead)?
- **Validity of the tests of the predictors** depends on having the “most right” model for the variances
 - Estimates will usually be ok → come from model for the means
 - Standard errors (and thus p -values) can be inaccurate
- The sources of variation that exist in your outcome will dictate **what kinds of predictors** will be useful
 - Between-Person variation needs Between-Person predictors
 - Within-Person variation needs Within-Person predictors
 - Between-Group variation needs Between-Group predictors

2. Include categorical or continuous predictors at any level of analysis

- ANOVA: test differences among discrete IV factor levels
 - Between-Groups: Gender, Intervention Group, Age Groups
 - Within-Subjects (Repeated Measures): Condition, Time
 - Test main effects of continuous covariates (ANCOVA)
- Regression: test whether slopes relating predictors to outcomes are different from 0
 - Persons measured once, differ categorically or continuously on a set of time-invariant (person-level) covariates
- What if a predictor is assessed repeatedly (time-varying predictors) but can't be characterized by 'conditions'?
 - ANOVA or Regression won't work → need MLM

2. Include categorical or continuous predictors at any level of analysis

- Some things don't change over measurements...
 - Sex, Ethnicity
 - Time-Invariant Predictor = Person Level
- Some things do change over measurements...
 - Health Status, Stress Levels, Living Arrangements
 - Time-Varying Predictor = Time Level
- Some predictors might be measured at higher levels
 - Family SES, length of marriage, school size, country size
- Interactions between levels may be included, too
 - Does the effect of health status differ by gender and SES?

Level: Time Person Family

3. Does not require same data structure per person (by accident or by design)

RM ANOVA: uses **multivariate** (wide) data structure:

| ID | Sex | T1 | T2 | T3 | T4 |
|-----|-----|----|----|----|----|
| 100 | 0 | 5 | 6 | 8 | 12 |
| 101 | 1 | 4 | 7 | . | 11 |

People missing any data are excluded (data from ID 101 are not included at all)

MLM: uses **stacked** (long) data structure:

Only rows missing data are excluded

100 uses 4 cases
101 uses 3 cases

| ID | Sex | Time | Y |
|-----|-----|------|----|
| 100 | 0 | 1 | 5 |
| 100 | 0 | 2 | 6 |
| 100 | 0 | 3 | 8 |
| 100 | 0 | 4 | 12 |
| 101 | 1 | 1 | 4 |
| 101 | 1 | 2 | 7 |
| 101 | 1 | 3 | . |
| 101 | 1 | 4 | 11 |

Time can also be **unbalanced** across people such that each person can have his or her own measurement schedule: Time "0.9" "1.4" "3.5" "4.2"...

4. You already know how!

- If you can do GLM, you can do MLM
(and if you can do generalized linear models, you can do generalized multilevel models, too)
- How do you interpret an estimate for...
 - the intercept?
 - the effect of a continuous variable?
 - the effect of a categorical variable?
 - a variance component ("pile of variance")?

Introduction to Multilevel Models

- Topics:
 - What is multilevel modeling?
 - **Concepts in longitudinal data**
 - From between-person to within-person models
 - Kinds of ANOVAs for longitudinal data

Options for Longitudinal Models

- Although models and software are logically separate, longitudinal data can be analyzed via multiple analytic frameworks:
 - “Multilevel/Mixed Models”
 - Dependency over time, persons, groups, etc. is modeled via random effects (multivariate through “levels” using stacked/long data)
 - Builds on GLM, generalizes easier to additional levels of analysis
 - “Structural Equation Models”
 - Dependency over time *only* is modeled via latent variables (single-level analysis using multivariate/wide data)
 - Generalizes easier to broader analysis of latent constructs, mediation
 - Because random effects and latent variables are the same thing, many longitudinal models can be specified/estimated either way
 - And now “Multilevel Structural Equation Models” can do it all...

Data Requirements for Our Models

- A useful outcome variable:
 - Has an interval scale*
 - A one-unit difference means the same thing across all scale points
 - In subscales, each contributing item has an equivalent scale
 - **Other kinds of outcomes can be analyzed using generalized multilevel models instead, but estimation is more challenging*
 - Has scores with the same meaning over observations
 - Includes meaning of construct
 - Includes how items relate to the scale
 - Implies measurement invariance
- FANCY MODELS CANNOT SAVE BADLY MEASURED VARIABLES OR CONFOUNDED RESEARCH DESIGNS.

Requirements for Longitudinal Data

- Multiple OUTCOMES from the same sampling unit!
 - 2 is the minimum, but just 2 can lead to problems:
 - Only 1 kind of change is observable (1 difference)
 - Can't distinguish "real" individual differences in change from error
 - Repeated measures ANOVA is just fine for 2 observations
 - Necessary assumption of "sphericity" is satisfied with only 2 observations even if compound symmetry doesn't hold
 - More data is better (with diminishing returns)
 - More occasions → better description of the form of change
 - More persons → better estimates of amount of individual differences in change; better prediction of those individual differences
 - More items/stimuli → more power to show effects of differences between items/stimuli/conditions

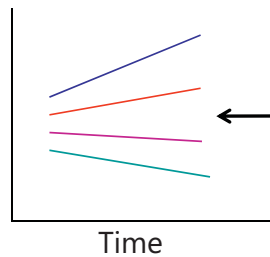
Levels of Analysis in Longitudinal Data

- Between-Person (BP) Variation:
 - **Level-2** – "**INTER**-individual Differences" – Time-Invariant
 - All longitudinal studies begin as cross-sectional studies
- Within-Person (WP) Variation:
 - **Level-1** – "**INTRA**-individual Differences" – Time-Varying
 - Only longitudinal studies can provide this extra information
- Longitudinal studies allow examination of both types of relationships simultaneously (and their interactions)
 - Any variable measured over time usually has both BP and WP variation
 - BP = more/less than other people; WP = more/less than one's average
- I use "person" here, but level-2 can be anything that is measured repeatedly (like animals, schools, countries...)

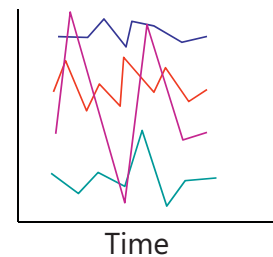
A Longitudinal Data Continuum

- **Within-Person Change:** Systematic change
 - Magnitude or direction of change can be different across individuals
 - “Growth curve models” → Time is meaningfully sampled
- **Within-Person Fluctuation:** No systematic change
 - Outcome just varies/fluctuates over time (e.g., emotion, stress)
 - Time is just a way to get lots of data per individual

Pure WP Change



Pure WP Fluctuation



Introduction to Multilevel Models

- Topics:
 - What is multilevel modeling?
 - Concepts in longitudinal data
 - **From between-person to within-person models**
 - Kinds of ANOVAs for longitudinal data

The Two Sides of a (BP) Model

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i + e_i$$

- **Model for the Means (Predicted Values):**

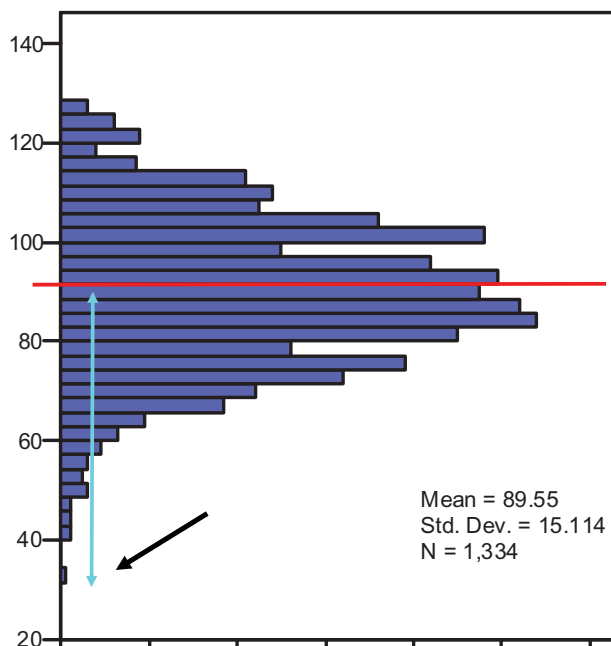
Our focus today

- Each person's expected (predicted) outcome is a weighted linear function of his/her values on X and Z (and here, their interaction), each measured once per person (i.e., this is a between-person model)
- Estimated parameters are called fixed effects (here, β_0 , β_1 , β_2 , and β_3)

- **Model for the Variance ("Piles" of Variance):**

- $e_i \sim N(0, \sigma_e^2) \rightarrow$ ONE residual (unexplained) deviation
- e_i has a mean of 0 with some estimated constant variance σ_e^2 , is normally distributed, is unrelated to X and Z, and is unrelated across people (across all observations, just people here)
- **Estimated parameter is residual variance only in above BP model**

An Empty Between-Person Model (i.e., Single-Level)



$$y_i = \beta_0 + e_i$$

Filling in values:

$$32 = 90 + -58$$

Y pred

Model
for the
Means

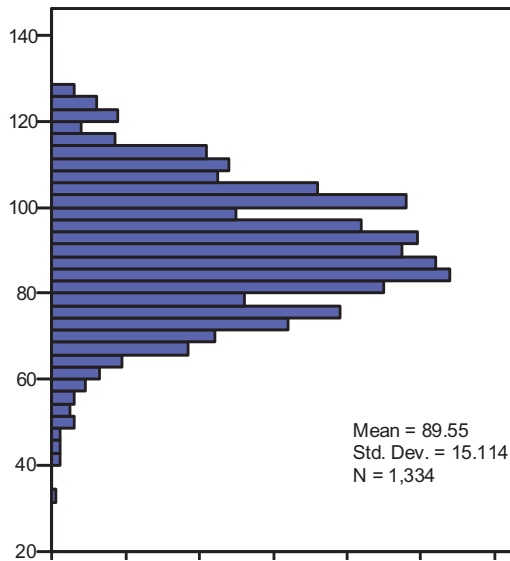
Y Error

Variance:

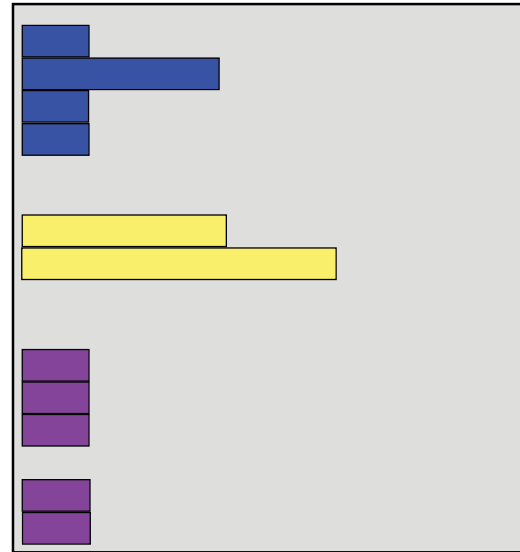
$$\frac{\sum (y - y_{\text{pred}})^2}{N - 1}$$

Adding Within-Person Information... (i.e., to become a Multilevel Model)

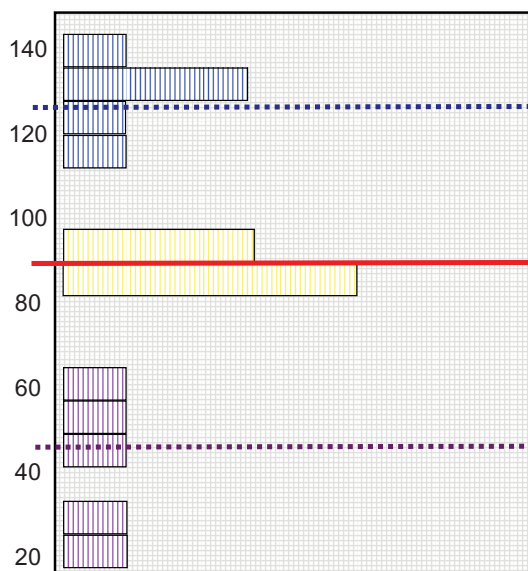
Full Sample Distribution



3 People, 5 Occasions each



Empty + Within-Person Model



**Start off with Mean of Y as
“best guess” for any value:**

= Grand Mean

= Fixed Intercept

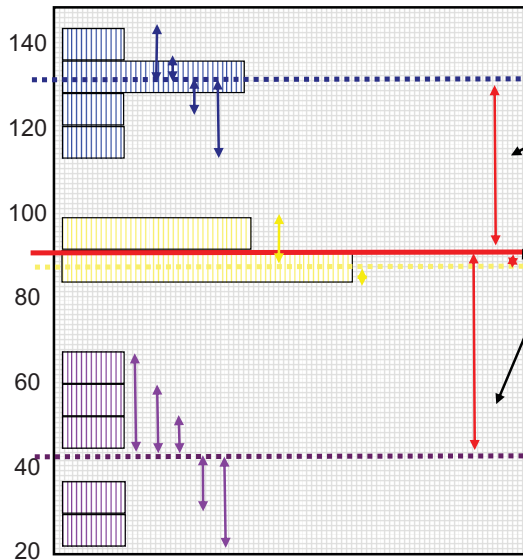
**Can make better guess by
taking advantage of
repeated observations:**

= Person Mean

→ Random Intercept

Empty + Within-Person Model

Variance of Y \rightarrow 2 sources:



Between-Person (BP) Variance:

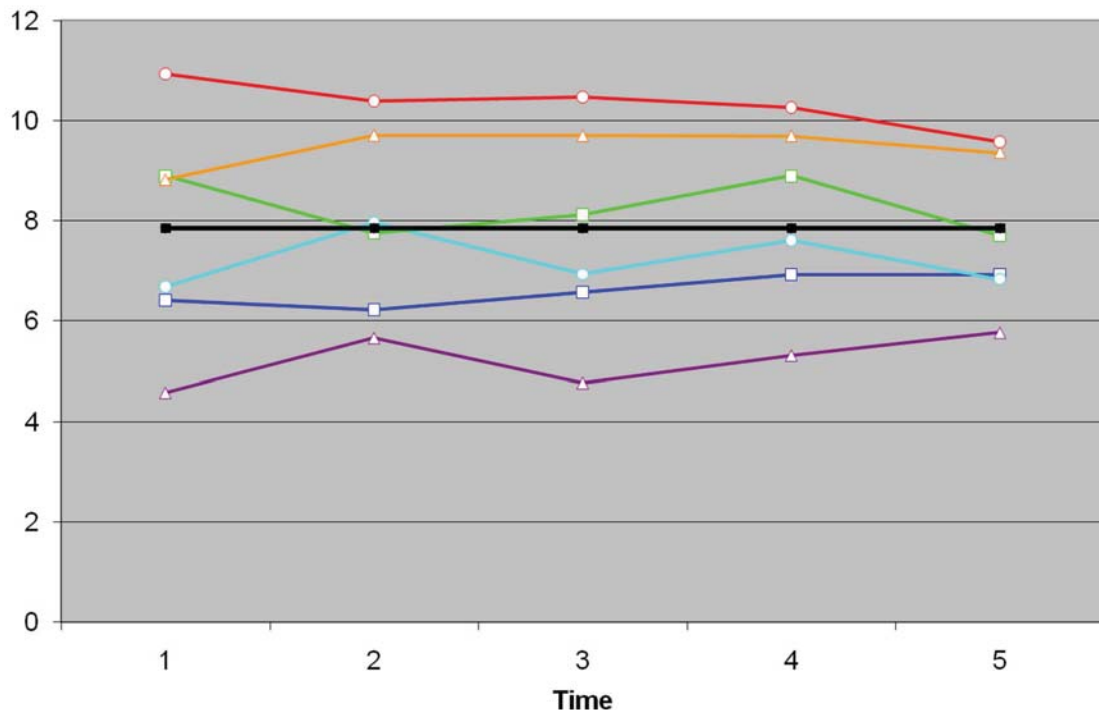
- \rightarrow Differences from **GRAND** mean
- \rightarrow **INTER**-Individual Differences

Within-Person (WP) Variance:

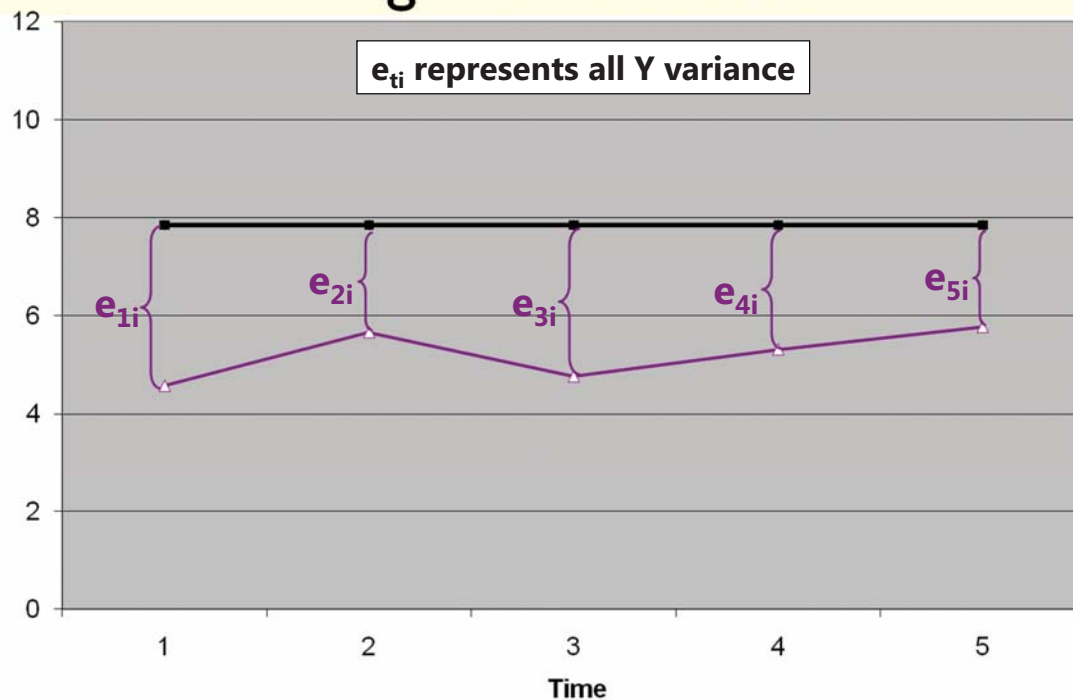
- \rightarrow Differences from **OWN** mean
- \rightarrow **INTRA**-Individual Differences
- \rightarrow This part is only observable through longitudinal data.

Now we have 2 piles of variance in Y to predict.

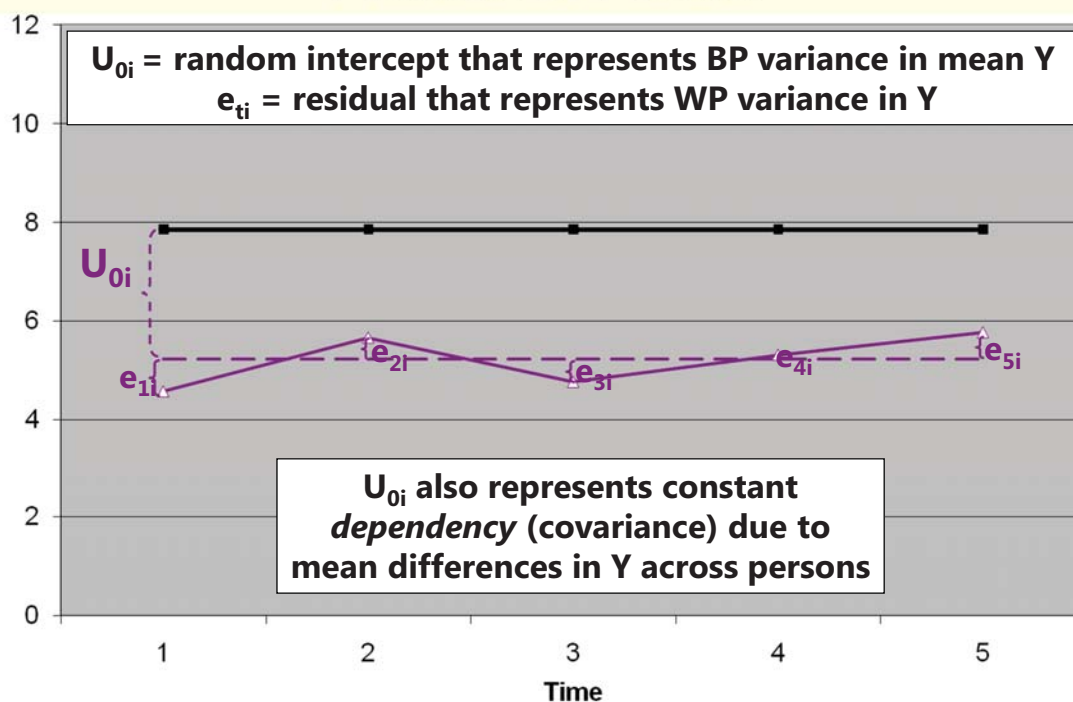
Hypothetical Longitudinal Data



“Error” in a BP Model for the Variance: Single-Level Model

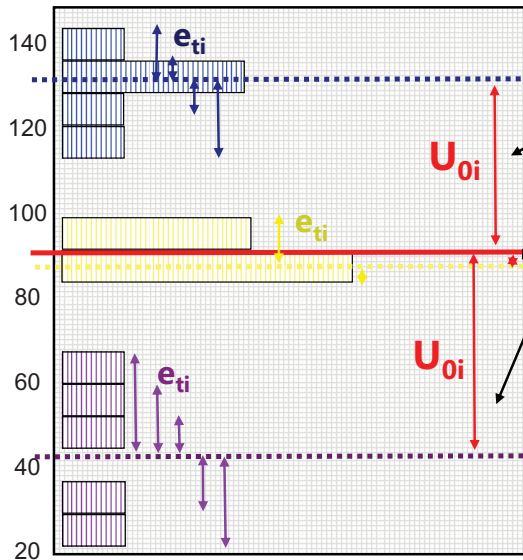


“Error” in a +WP Model for the Variance: Multilevel Model



Empty + Within-Person Model

Variance of Y → 2 sources:



Level 2 Random Intercept

Variance (of U_{0i} , as $\tau_{U_0}^2$):

- **Between**-Person Variance
- Differences from **GRAND** mean
- **INTER**-Individual Differences

Level 1 Residual Variance

(of e_{ti} , as σ_e^2):

- **Within**-Person Variance
- Differences from **OWN** mean
- **INTRA**-Individual Differences

BP vs. +WP Empty Models

- Empty **Between-Person** Model (used for 1 occasion):

$$y_i = \beta_0 + e_i$$

- β_0 = fixed intercept = grand mean
- e_i = residual deviation from GRAND mean

- Empty + **Within-Person** Model (>1 occasions):

$$y_{ti} = \beta_0 + U_{0i} + e_{ti}$$

- β_0 = fixed intercept = grand mean
- U_{0i} = random intercept = individual deviation from GRAND mean
- e_{ti} = time-specific residual deviation from OWN mean

Intraclass Correlation (ICC)

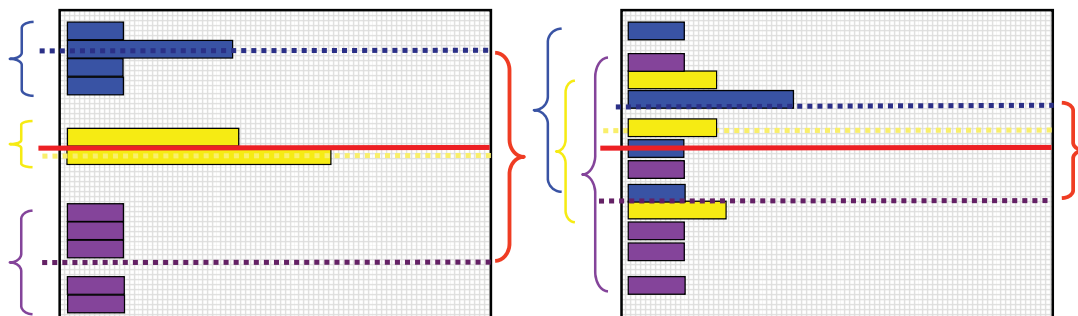
Intraclass Correlation (ICC):

$$\begin{aligned} \text{ICC} &= \frac{\text{BP}}{\text{BP} + \text{WP}} = \frac{\text{Intercept Variance}}{\text{Intercept Variance} + \text{Residual Variance}} \\ &= \frac{\text{BP Variance in Mean Outcome}}{\text{Total Outcome Variance}} = \frac{\tau_{U_0}^2}{\tau_{U_0}^2 + \sigma_e^2} \end{aligned}$$

- ICC = Proportion of total variance that is between persons
- ICC = Average correlation among occasions
- ICC is a standardized way of expressing how much we need to worry about *dependency due to person mean differences* (**i.e., ICC is an effect size for *constant* person dependency**)

$$\text{ICC} = \frac{\text{Between-Person}}{\text{Between-Person} + \text{Within-Person}}$$

Counter-Intuitive: Between-Person Variance is in the numerator, but the ICC is the correlation over time!



$$\text{ICC} = \text{BTW} / \text{BTW} + \text{within}$$

→ Large ICC

→ Large correlation over time

$$\text{ICC} = \text{btw} / \text{btw} + \text{WITHIN}$$

→ Small ICC

→ Small correlation over time

BP and +WP Conditional Models

- Multiple Regression, **Between-Person** ANOVA: **1 PILE**
 - $y_i = (\beta_0 + \beta_1 X_i + \beta_2 Z_i \dots) + e_i$
 - $e_i \rightarrow$ ONE residual, assumed uncorrelated with equal variance across observations (here, just persons) \rightarrow "**BP (all) variation**"
- Repeated Measures, **Within-Person** ANOVA: **2 PILES**
 - $y_{ti} = (\beta_0 + \beta_1 X_i + \beta_2 Z_i \dots) + U_{0i} + e_{ti}$
 - $U_{0i} \rightarrow$ A random intercept for differences in person means, assumed uncorrelated with equal variance across persons
 \rightarrow "**BP (mean) variation**" = $\tau_{U_0}^2$ is now "leftover" after predictors
 - $e_{ti} \rightarrow$ A residual that represents remaining time-to-time variation, usually assumed uncorrelated with equal variance across observations (now, persons and time) \rightarrow "**WP variation**"
= σ_e^2 is also now "leftover" after predictors

Introduction to Multilevel Models

- Topics:
 - What is multilevel modeling?
 - Concepts in longitudinal data
 - From between-person to within-person models
 - **Kinds of ANOVAs for longitudinal data**

ANOVA for longitudinal data?

- There are 3 possible “kinds” of ANOVAs we could use:
 - Between-Persons/Groups, Univariate RM, and Multivariate RM
- **NONE OF THEM ALLOW:**
 - **Missing occasions** (do listwise deletion due to least squares)
 - **Time-varying predictors** (covariates are BP predictors only)
- Each includes the same model for the means for time: all possible mean differences (so 4 parameters to get to 4 means)
 - **“Saturated means model”**: $\beta_0 + \beta_1(T_1) + \beta_2(T_2) + \beta_3(T_3)$
 - **The Time variable must be balanced and discrete in ANOVA!**
- These ANOVAs differ by what they predict for the correlation across outcomes from the same person in the model for the variances...
 - i.e., **how they “handle dependency”** due to persons, or what they says the variance and covariance of the y_{ti} residuals should look like...

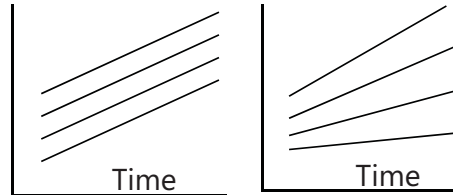
I. Between-Groups ANOVA

- **Uses e_{ti} only** (total variance = a single variance term of σ_e^2)
- **Assumes no covariance** at all among observations from the same person: *Dependency? What dependency?*
- Will usually be **very, very wrong** for longitudinal data
 - WP effects tested against wrong residual variance (significance tests will often be way too conservative)
 - Will also tend to be wrong for clustered data, but less so (*because the correlation among persons from the same group is not as strong as the correlation among occasions from the same person*)
- Predicts a variance-covariance matrix over time (here, 4 occasions) like this, called **“Variance Components”**:

$$\begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

2a. Univariate Repeated Measures

- Separates total variance into two sources:
 - **Between-Person** (mean differences due to U_{0i} or $\tau_{U_0}^2$)
 - **Within-Person** (remaining variance due to e_{ti} or σ_e^2)
- Predicts a variance-covariance matrix over time (here, 4 occasions) like this, called "**Compound Symmetry**":
 - **Mean differences from U_{0i} are the only reason why occasions are correlated**
- Will usually be at least somewhat wrong for longitudinal data
 - If people change at different rates, the variances and covariances over time have to change, too



The Problem with Univariate RM ANOVA

- Univ. RM ANOVA ($\tau_{U_0}^2 + \sigma_e^2$) predicts **compound symmetry**:
 - All variances and all covariances are equal across occasions
 - In other words, the amount of error observed should be the same at any occasion, so a single, pooled error variance term makes sense
 - If not, tests of fixed effects may be biased (i.e., sometimes tested against too much or too little error, if error is not really constant over time)
 - **COMPOUND SYMMETRY RARELY FITS FOR LONGITUDINAL DATA**
- But to get the correct tests of the fixed effects, the data must only meet a less restrictive assumption of **sphericity**:
 - In English → **pairwise differences** between adjacent occasions have equal variance and covariance (satisfied by default with only 2 occasions)
 - If compound symmetry is satisfied, so is sphericity (but see above)
 - Significance test provided in ANOVA for where data meet sphericity assumption
 - **Other RM ANOVA approaches are used when sphericity fails...**

The Other Repeated Measures ANOVAs...

• 2b. **Univariate RM ANOVA with sphericity corrections**

- Based on $\epsilon \rightarrow$ how far off sphericity (from 0-1, 1=spherical)
- Applies an overall correction for model df based on estimated ϵ , but it doesn't really address the problem that data \neq model

• 3. **Multivariate Repeated Measures ANOVA**

- All variances and covariances are estimated separately over time (here, 4 occasions), called "**Unstructured**"—it's not a model, it IS the data reproduced directly:

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2 \end{bmatrix}$$

- Because it can never be wrong, UN can be useful for **complete and balanced longitudinal data** with few occasions (e.g., 2-4)
- Parameters = $\frac{\text{\#occasions} * (\text{\#occasions} + 1)}{2}$ so can be hard to estimate
- Unstructured can also be specified to include random intercept variance τ_{U0}^2
- Every other model for the variances is nested within Unstructured (we can do model comparisons to see if all other models are NOT WORSE)

Summary: ANOVA approaches for longitudinal data are "one size fits most"

• **Saturated Model for the Means** (balanced time required)

- All possible mean differences
- Unparsimonious, but best-fitting (is a description, not a model)

• **3 kinds of Models for the Variances** (complete data required)

- BP ANOVA (σ_e^2 only) \rightarrow assumes independence and constant variance over time
- Univ. RM ANOVA ($\tau_{U0}^2 + \sigma_e^2$) \rightarrow assumes constant variance and covariance
- Multiv. RM ANOVA (whatever) \rightarrow no assumptions; is a description, not a model

there is no structure that shows up in a scalar equation (i.e., the way $U_{0i} + e_{ti}$ does)

• **MLM will give us more flexibility in both parts of the model:**

- Fixed effects that *predict* the pattern of means (polynomials, pieces)
- Random intercepts and slopes and/or alternative covariance structures that *predict* intermediate patterns of variance and covariance over time

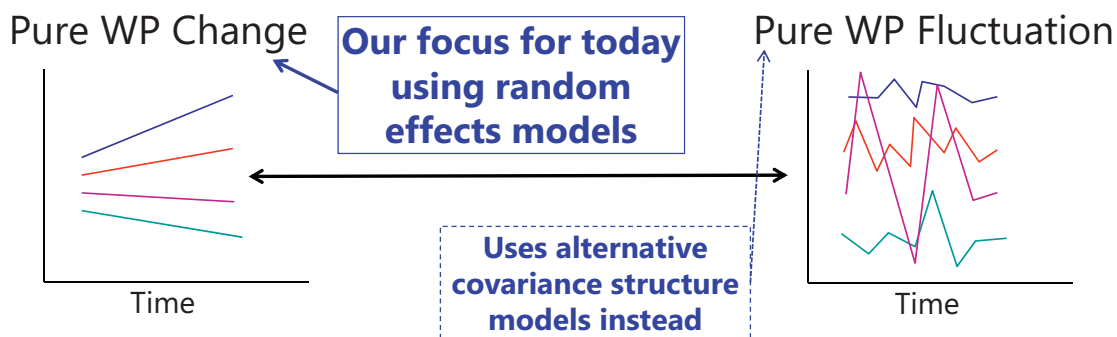
Describing Within-Person Change in Longitudinal Data

- Topics:
 - **Multilevel modeling notation and terminology**
 - Fixed and random effects of linear time
 - Predicted variances and covariances from random slopes
 - Dependency and effect size in random effects models
 - Describing nonlinear change: polynomial, piecewise models, and truly nonlinear models
 - Fun with likelihood estimation and model comparisons

Lecture 2

1

Modeling Change vs. Fluctuation



Model for the Means:

- **WP Change** → describe pattern of *average* change (over "time")
- WP Fluctuation → *may* not need anything (if no systematic change)

Model for the Variances:

- **WP Change** → describe *individual differences* in change (random effects)
 - this allows variances and covariances to differ over time
- WP Fluctuation → describe pattern of variances and covariances over time

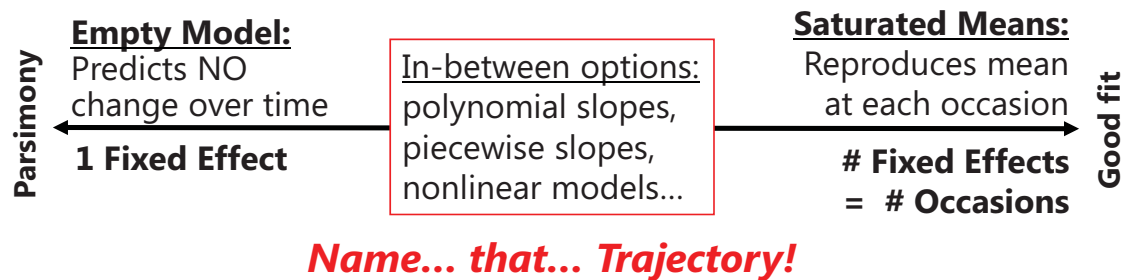
Lecture 2

2

The Big Picture of Longitudinal Data:

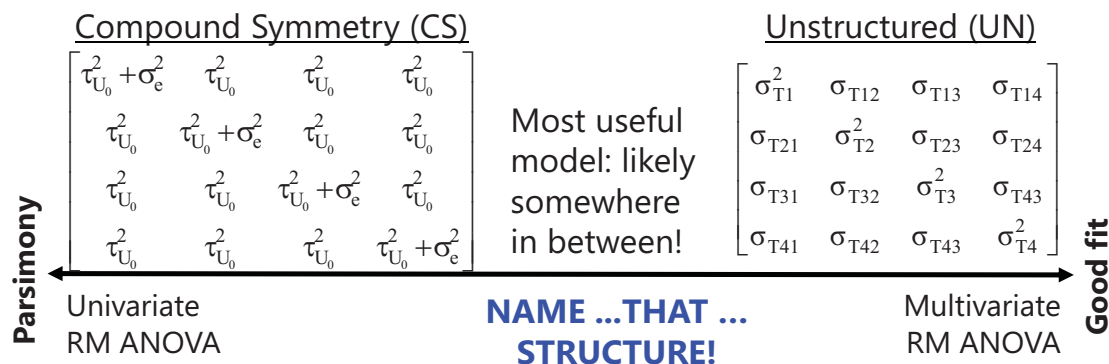
Models for the Means

- What kind of change occurs on average over "time"?
There are two baseline models to consider:
 - "Empty" → only a fixed intercept (predicts no change)
 - "Saturated" → all occasion mean differences from time 0
(ANOVA model that uses # fixed effects = n)
*** *may not be possible in unbalanced data*



The Big Picture of Longitudinal Data:

Models for the Variance

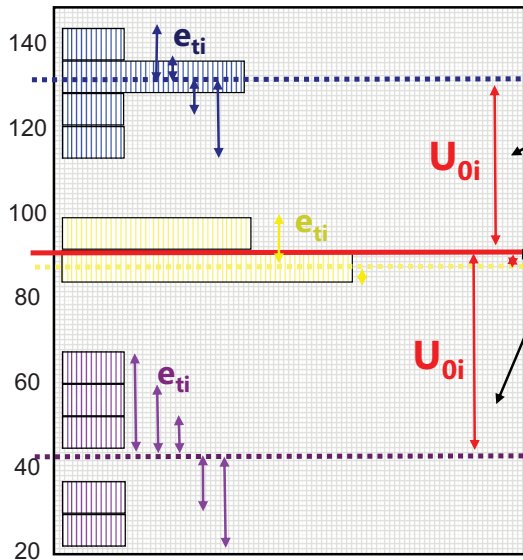


What is the pattern of variance and covariance over time?

CS and UN are just two of the many, many options available within MLM, including **random effects models** (for change) and **alternative covariance structure models** (for fluctuation).

Empty + Within-Person Model

Variance of Y → 2 sources:



Level 2 Random Intercept

Variance (of U_{0i} , as $\tau_{U_0}^2$):

- **Between**-Person Variance
- Differences from **GRAND** mean
- **INTER**-Individual Differences

Level 1 Residual Variance

(of e_{ti} , as σ_e^2):

- **Within**-Person Variance
- Differences from **OWN** mean
- **INTRA**-Individual Differences

Empty Means, Random Intercept Model

GLM Empty Model:

$$y_i = \beta_0 + e_i$$

MLM Empty Model:

- Level 1:

$$y_{ti} = \beta_{0i} + e_{ti}$$

- Level 2:

$$\beta_{0i} = \gamma_{00} + U_{0i}$$

Fixed Intercept
= grand mean
(because no
predictors yet)

Random Intercept
= individual-specific
deviation from
predicted intercept

Residual = time-specific deviation
from individual's predicted outcome

3 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

- Level-1 Variance of $e_{ti} \rightarrow \sigma_e^2$
- Level-2 Variance of $U_{0i} \rightarrow \tau_{U_0}^2$

Composite equation:

$$y_{ti} = (\gamma_{00} + U_{0i}) + e_{ti}$$

Saturated Means, Random Intercept Model

- Although rarely shown this way, a saturated means, random intercept model would be represented as a multilevel model like this (for $n = 4$ here, in which the time predictors are dummy codes to distinguish each occasion from time 0):

- Level 1:

$$y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time1}_{ti}) + \beta_{2i}(\text{Time2}_{ti}) + \beta_{3i}(\text{Time3}_{ti}) + e_{ti}$$

- Level 2:

$$\beta_{0i} = \gamma_{00} + u_{0i}$$

$$\beta_{1i} = \gamma_{10}$$

$$\beta_{2i} = \gamma_{20}$$

$$\beta_{3i} = \gamma_{30}$$

Composite equation (6 parameters):

$$y_{ti} = \gamma_{00} + \gamma_{10}(\text{Time1}_{ti}) + \gamma_{20}(\text{Time2}_{ti}) + \gamma_{30}(\text{Time3}_{ti}) + u_{0i} + e_{ti}$$

This model is also known as **univariate repeated measures ANOVA**. Although the means are perfectly predicted, the random intercept assumes parallel growth (and equal variance/covariance over time).

Describing Within-Person Change in Longitudinal Data

- Topics:
 - Multilevel modeling notation and terminology
 - **Fixed and random effects of linear time**
 - Predicted variances and covariances from random slopes
 - Dependency and effect size in random effects models
 - Describing nonlinear change: polynomial, piecewise models, and truly nonlinear models
 - Fun with likelihood estimation and model comparisons

Augmenting the empty means, random intercept model with *time*

- 2 questions about the possible effects of *time*:

1. Is there an effect of time on average?

- If the line describing the sample means not flat?
- Significant **FIXED** effect of time

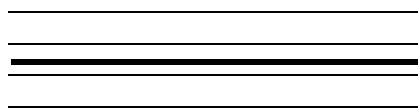
2. Does the average effect of time vary across individuals?

- Does each individual need his or her own line?
- Significant **RANDOM** effect of time

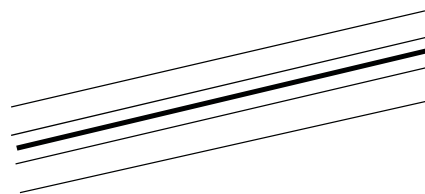
Fixed and Random Effects of Time

(Note: The intercept is random in every figure)

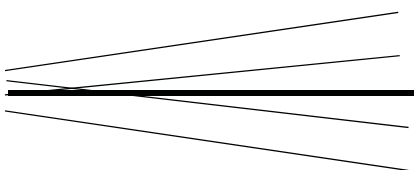
No Fixed, No Random



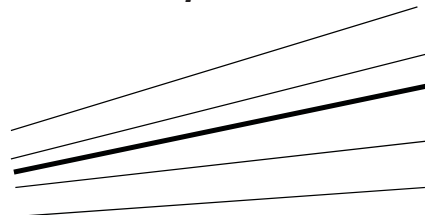
Yes Fixed, No Random



No Fixed, Yes Random



Yes Fixed, Yes Random



Fixed Linear Time, Random Intercept Model

(4 total parameters: effect of time is **FIXED** only)

Multilevel Model

Residual = time-specific deviation from individual's predicted outcome → estimated variance of σ_e^2

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti}$

Fixed Intercept
= predicted mean outcome at time 0

Fixed Linear Time Slope
= predicted mean rate of change per unit time

Level 2: $\beta_{0i} = \gamma_{00} + u_{0i}$ $\beta_{1i} = \gamma_{10}$

Random Intercept = individual-specific deviation from fixed intercept → estimated variance of $\tau_{u_0}^2$

Composite Model

$y_{ti} = (\underbrace{\gamma_{00} + u_{0i}}_{\beta_{0i}}) + (\underbrace{\gamma_{10}}_{\beta_{1i}})(\text{Time}_{ti}) + e_{ti}$

Because the effect of time is **fixed**, everyone is predicted to change at exactly the same rate.

Random Intercept Models Imply...

- **People differ from each other systematically in only ONE way**—in intercept (u_{0i}), which implies **ONE kind of BP variance**, which translates to **ONE source of person dependency** (covariance or correlation in the outcomes from the same person)
- If so, after controlling for BP intercept differences (by estimating the variance of u_{0i} as $\tau_{u_0}^2$ in the **G** matrix), the **e_{ti} residuals** (whose variance and covariance are estimated in the **R** matrix) should be **uncorrelated with homogeneous variance across time**, as shown:

Level-2
G matrix:
RANDOM
TYPE=UN

$$\begin{bmatrix} \tau_{u_0}^2 \end{bmatrix}$$

Level-1 **R** matrix:
REPEATED TYPE=VC

$$\begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

G and **R** matrices combine to create a total **V** matrix with CS pattern

$$\begin{bmatrix} \tau_{u_0}^2 + \sigma_e^2 & \tau_{u_0}^2 & \tau_{u_0}^2 & \tau_{u_0}^2 \\ \tau_{u_0}^2 & \tau_{u_0}^2 + \sigma_e^2 & \tau_{u_0}^2 & \tau_{u_0}^2 \\ \tau_{u_0}^2 & \tau_{u_0}^2 & \tau_{u_0}^2 + \sigma_e^2 & \tau_{u_0}^2 \\ \tau_{u_0}^2 & \tau_{u_0}^2 & \tau_{u_0}^2 & \tau_{u_0}^2 + \sigma_e^2 \end{bmatrix}$$

Matrices in a Random Intercept Model

Total predicted data matrix is called **V matrix**, created from the **G [TYPE=UN]** and **R [TYPE=VC]** matrices as follows:

$$V = Z * G * Z^T + R = V$$

$$V = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \tau_{U_0}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix} = \begin{bmatrix} \tau_{U_0}^2 + \sigma_e^2 & \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 \\ \tau_{U_0}^2 & \tau_{U_0}^2 + \sigma_e^2 & \tau_{U_0}^2 & \tau_{U_0}^2 \\ \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 + \sigma_e^2 & \tau_{U_0}^2 \\ \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 + \sigma_e^2 \end{bmatrix}$$

VCORR then provides the intraclass correlation, calculated as:

$$ICC = \tau_{U_0}^2 / (\tau_{U_0}^2 + \sigma_e^2)$$

$$\begin{bmatrix} 1 & ICC & ICC & ICC \\ ICC & 1 & ICC & ICC \\ ICC & ICC & 1 & ICC \\ ICC & ICC & ICC & 1 \end{bmatrix} \text{ assumes a constant correlation over time}$$

For any random effects model:

G matrix = BP variances/covariances

R matrix = WP variances/covariances

Z matrix = values of predictors with random effects (just intercept here), which can vary per person

V matrix = Total variance/covariance

Summary so far...

- Regardless of what kind of model for the means you have...
 - Empty means = 1 fixed intercept that predicts no change
 - Saturated means = 1 fixed intercept + $n-1$ fixed effects for mean differences that perfectly predict the means over time
 - Is a description, not a model, and may not be possible with unbalanced time
 - Fixed linear time = 1 fixed intercept, 1 fixed linear time slope that predicts linear average change across time
 - Is a model that works with balanced or unbalanced time
 - May cause an increase in the random intercept variance by explaining residual variance
- A random intercept model...
 - Predicts constant total variance and covariance over time in **V** using **G**
 - Should be possible in balanced or unbalanced data
 - Still has residual variance (always there via default **R** matrix TYPE=VC)
- Now we'll see what happens when adding other kinds of random effects, such as a random linear effect of time...

Random Linear Time Model (6 total parameters)

Multilevel Model

Residual = time-specific deviation from individual's predicted outcome → estimated variance of σ_e^2

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti}$

Fixed Intercept
= predicted mean outcome at time 0

Fixed Linear Time Slope
= predicted mean rate of change per unit time

Level 2: $\beta_{0i} = Y_{00} + U_{0i}$ $\beta_{1i} = Y_{10} + U_{1i}$

Random Intercept = individual-specific deviation from fixed intercept at time 0 → estimated variance of $\tau_{U_0}^2$

Random Linear Time Slope = individual-specific deviation from fixed linear time slope → estimated variance of $\tau_{U_1}^2$

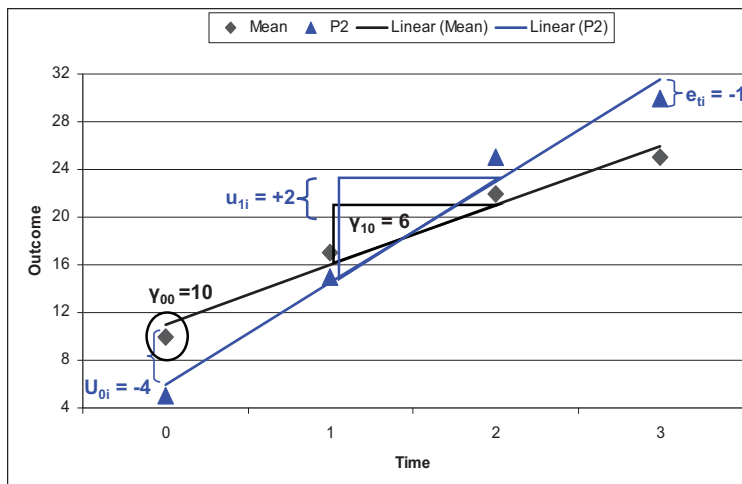
Also has an estimated covariance of random intercepts and slopes of $\tau_{U_{01}}$

Composite Model

$$y_{ti} = (\underbrace{Y_{00} + U_{0i}}_{\beta_{0i}}) + (\underbrace{Y_{10} + U_{1i}}_{\beta_{1i}})(\text{Time}_{ti}) + e_{ti}$$

Random Linear Time Model

$$y_{ti} = (\underbrace{Y_{00}}_{\text{Fixed Intercept}} + \underbrace{U_{0i}}_{\text{Random Intercept Deviation}}) + (\underbrace{Y_{10}}_{\text{Fixed Slope}} + \underbrace{U_{1i}}_{\text{Random Slope Deviation}})(\text{Time}_{ti}) + \underbrace{e_{ti}}_{\text{error for person } i \text{ at time } t}$$



6 Parameters:

2 Fixed Effects:

Y_{00} Intercept, Y_{10} Slope

2 Random Effects

Variances:

U_{0i} Intercept Variance

$$= \tau_{U_0}^2$$

U_{1i} Slope Variance

$$= \tau_{U_1}^2$$

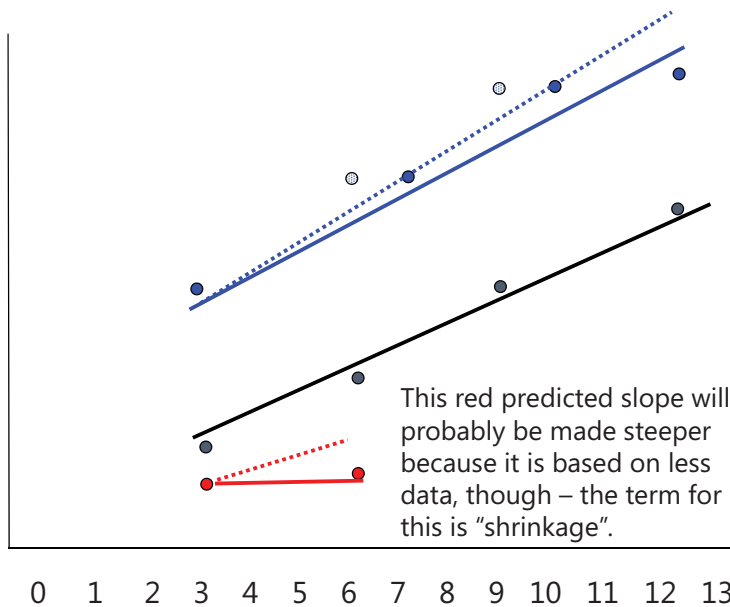
Int-Slope Covariance

$$= \tau_{U_{01}}$$

e_{ti} Residual Variance

$$= \sigma_e^2$$

Unbalanced Time → Different time occasions across persons? No problem!



Rounding time can lead to incorrect estimates!

Code time as **exactly** as possible.

MLM uses each complete time point for each person.

Summary: Sequential Models for Effects of Time

Level 1: $y_{ti} = \beta_{0i} + e_{ti}$

Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$

Composite: $y_{ti} = \gamma_{00} + U_{0i} + e_{ti}$

Empty Means,
Random Intercept Model:
3 parms = γ_{00} , σ_e^2 , $\tau_{U_0}^2$

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti}$

Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$
 $\beta_{1i} = \gamma_{10}$

Composite: $y_{ti} = (\gamma_{00} + U_{0i}) + \gamma_{10}(\text{Time}_{ti}) + e_{ti}$

Fixed Linear Time,
Random Intercept Model:
4 parms = γ_{00} , γ_{10} , σ_e^2 , $\tau_{U_0}^2$

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti}$

Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$
 $\beta_{1i} = \gamma_{10} + U_{1i}$

Composite: $y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(\text{Time}_{ti}) + e_{ti}$

Random Linear Time Model:
6 parms = γ_{00} , γ_{10} , σ_e^2 , $\tau_{U_0}^2$,
 $\tau_{U_1}^2$, $\tau_{U_{01}}$ (→ cov of U_{0i} and U_{1i})

Describing Within-Person Change in Longitudinal Data

- Topics:
 - Multilevel modeling notation and terminology
 - Fixed and random effects of linear time
 - **Predicted variances and covariances from random slopes**
 - Dependency and effect size in random effects models
 - Describing nonlinear change: polynomial, piecewise models, and truly nonlinear models
 - Fun with likelihood estimation and model comparisons

Lecture 2

19

Random Linear Time Models Imply:

- **People differ from each other systematically in TWO ways**—in intercept (U_{0i}) and slope (U_{1i}), which implies **TWO kinds of BP variance**, which translates to **TWO sources of person dependency** (covariance or correlation in the outcomes from the same person)
- If so, after controlling for both BP intercept and slope differences (by estimating the $\tau_{U_0}^2$ and $\tau_{U_1}^2$ variances in the **G** matrix), the **e_{ti} residuals** (whose variance and covariance are estimated in the **R** matrix) should be **uncorrelated with homogeneous variance across time**, as shown:

Level-2
G matrix:
RANDOM
TYPE=UN

$$\begin{bmatrix} \tau_{U_0}^2 & \tau_{U_{01}} \\ \tau_{U_{01}} & \tau_{U_1}^2 \end{bmatrix}$$

Level-1 **R** matrix:
REPEATED TYPE=VC

$$\begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

G and **R** combine to create a total **V** matrix whose per-person structure depends on the specific time occasions each person has (very flexible for unbalanced time)

Lecture 2

20

Random Linear Time Model

(6 total parameters: effect of time is now **RANDOM**)

- How the model predicts each element of the **V** matrix:

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti}$

Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$
 $\beta_{1i} = \gamma_{10} + U_{1i}$

Composite Model: $y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(\text{Time}_{ti}) + e_{ti}$

Predicted *Time-Specific* Variance:

$$\begin{aligned} \text{Var}[y_{ti}] &= \text{Var}[(\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(\text{Time}_i) + e_{ti}] \\ &= \text{Var}[U_{0i} + (U_{1i} * \text{Time}_i) + e_{ti}] \\ &= \{\text{Var}(U_{0i})\} + \{\text{Var}(U_{1i} * \text{Time}_i)\} + \{2 * \text{Cov}(U_{0i}, U_{1i} * \text{Time}_i)\} + \{\text{Var}(e_{ti})\} \\ &= \{\text{Var}(U_{0i})\} + \{\text{Time}_i^2 * \text{Var}(U_{1i})\} + \{2 * \text{Time}_i * \text{Cov}(U_{0i}, U_{1i})\} + \{\text{Var}(e_{ti})\} \\ &= \{\tau_{U_0}^2\} + \{\text{Time}_i^2 * \tau_{U_1}^2\} + \{2 * \text{Time}_i * \tau_{U_{01}}\} + \{\sigma_e^2\} \end{aligned}$$

Random Linear Time Model

(6 total parameters: effect of time is now **RANDOM**)

- How the model predicts each element of the **V** matrix:

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti}$

Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$
 $\beta_{1i} = \gamma_{10} + U_{1i}$

Composite Model: $y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(\text{Time}_{ti}) + e_{ti}$

Predicted *Time-Specific* Covariances (Time A with Time B):

$$\begin{aligned} \text{Cov}[y_{Ai}, y_{Bi}] &= \text{Cov}[(\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(A_i) + e_{Ai}, (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(B_i) + e_{Bi}] \\ &= \text{Cov}[U_{0i} + (U_{1i}A_i), U_{0i} + (U_{1i}B_i)] \\ &= \text{Cov}[U_{0i}, U_{0i}] + \text{Cov}[U_{0i}, U_{1i}B_i] + \text{Cov}[U_{1i}A_i, U_{0i}] + \text{Cov}[U_{1i}A_i, U_{1i}B_i] \\ &= \{\text{Var}(U_{0i})\} + \{(A_i + B_i) * \text{Cov}(U_{0i}, U_{1i})\} + \{(A_i B_i) \text{Var}(U_{1i})\} \\ &= \{\tau_{U_0}^2\} + \{(A_i + B_i) \tau_{U_{01}}\} + \{(A_i B_i) \tau_{U_1}^2\} \end{aligned}$$

Random Linear Time Model

(6 total parameters: effect of time is now **RANDOM**)

- Scalar "mixed" model equation per person:

$$Y_i = X_i * \gamma + Z_i * U_i + E_i$$

$$\begin{bmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} U_{0i} \\ U_{1i} \end{bmatrix} + \begin{bmatrix} e_{0i} \\ e_{1i} \\ e_{2i} \\ e_{3i} \end{bmatrix}$$

$$\begin{bmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{bmatrix} = \begin{bmatrix} \gamma_{00} + \gamma_{10}(0) \\ \gamma_{00} + \gamma_{10}(1) \\ \gamma_{00} + \gamma_{10}(2) \\ \gamma_{00} + \gamma_{10}(3) \end{bmatrix} + \begin{bmatrix} U_{0i} + U_{1i}(0) \\ U_{0i} + U_{1i}(1) \\ U_{0i} + U_{1i}(2) \\ U_{0i} + U_{1i}(3) \end{bmatrix} + \begin{bmatrix} e_{0i} \\ e_{1i} \\ e_{2i} \\ e_{3i} \end{bmatrix}$$

$$\begin{bmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{bmatrix} = \begin{bmatrix} \gamma_{00} + \gamma_{10}(0) + U_{0i} + U_{1i}(0) + e_{0i} \\ \gamma_{00} + \gamma_{10}(1) + U_{0i} + U_{1i}(1) + e_{1i} \\ \gamma_{00} + \gamma_{10}(2) + U_{0i} + U_{1i}(2) + e_{2i} \\ \gamma_{00} + \gamma_{10}(3) + U_{0i} + U_{1i}(3) + e_{3i} \end{bmatrix}$$

$X_i = n \times k$ values of **predictors with fixed effects**, so can differ per person ($k = 2$: intercept, linear time)

$\gamma = k \times 1$ estimated **fixed effects**, so will be the same for all persons (γ_{00} = intercept, γ_{10} = linear time)

$Z_i = n \times u$ values of **predictors with random effects**, so can differ per person ($u = 2$: intercept, linear time)

$U_i = u \times 2$ estimated individual **random effects**, so can differ per person

$E_i = n \times n$ time-specific residuals, so can differ per person

Random Linear Time Model

(6 total parameters: effect of time is now **RANDOM**)

- Predicted total variances and covariances per person:

$$V_i = Z_i * G_i * Z_i^T + R_i$$

$$V_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \tau_{U_0}^2 & \tau_{U_{01}} \\ \tau_{U_{01}} & \tau_{U_1}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

V_i matrix: Variance $[y_{\text{time}}]$

V_i matrix = complicated ☺

$$= \tau_{U_0}^2 + \left[(\text{time})^2 \tau_{U_1}^2 \right] + \left[2(\text{time}) \tau_{U_{01}} \right] + \sigma_e^2$$

V_i matrix: Covariance $[y_A, y_B]$

$$= \tau_{U_0}^2 + \left[(A + B) \tau_{U_{01}} \right] + \left[(AB) \tau_{U_1}^2 \right]$$

$Z_i = n \times u$ values of **predictors with random effects**, so can differ per person ($u = 2$: int., time slope)

$Z_i^T = u \times n$ values of predictors with random effects (just Z_i transposed)

$G_i = u \times u$ estimated **random effects variances and covariances**, so will be the same for all persons ($\tau_{U_0}^2$ = int. var., $\tau_{U_1}^2$ = slope var.)

$R_i = n \times n$ **time-specific residual variances and covariances**, so will be same for all persons (here, just diagonal σ_e^2)

Building **V** across persons: Random Linear Time Model

- **V** for two persons with **unbalanced time** observations:

$$V = Z * G * Z^T + R$$

$$V = \begin{bmatrix} 1 & 0.0 & 0 & 0 \\ 1 & 1.0 & 0 & 0 \\ 1 & 2.0 & 0 & 0 \\ 1 & 3.0 & 0 & 0 \\ 0 & 0 & 1 & 0.2 \\ 0 & 0 & 1 & 1.4 \\ 0 & 0 & 1 & 2.3 \\ 0 & 0 & 1 & 3.5 \end{bmatrix} \begin{bmatrix} \tau_{U_0}^2 & \tau_{U_{01}} & 0 & 0 \\ \tau_{U_{01}} & \tau_{U_1}^2 & 0 & 0 \\ 0 & 0 & \tau_{U_0}^2 & \tau_{U_{01}} \\ 0 & 0 & \tau_{U_{01}} & \tau_{U_1}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0.0 & 1.0 & 2.0 & 3.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0.2 & 1.4 & 2.3 & 3.5 \end{bmatrix} + \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_e^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_e^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

- The giant combined **V** matrix across persons is how the multilevel or mixed model is actually estimated
- Known as “**block diagonal**” structure → predictions are given for each person, but 0’s are given for the elements that describe relationships between persons (because persons are supposed to be independent here!)

Building **V** across persons: Random Linear Time Model

- **V** for two persons also with **different n** per person:

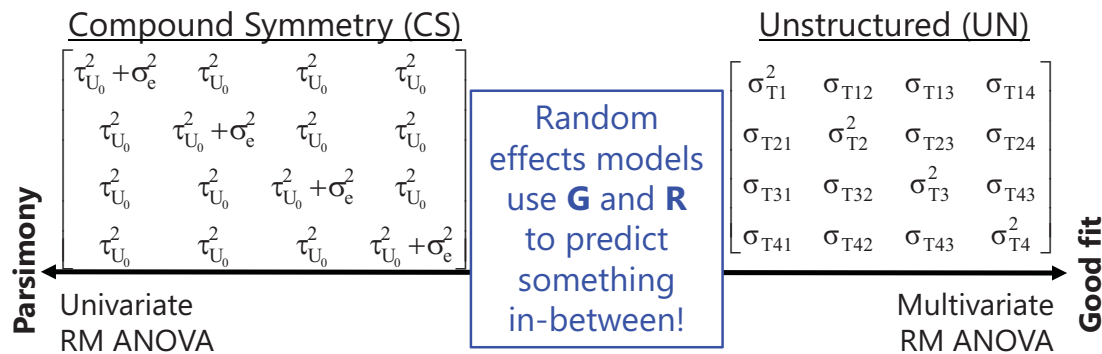
$$V = Z * G * Z^T + R$$

$$V = \begin{bmatrix} 1 & 0.0 & 0 & 0 \\ 1 & 1.0 & 0 & 0 \\ 1 & 2.0 & 0 & 0 \\ 1 & 3.0 & 0 & 0 \\ 0 & 0 & 1 & 0.2 \\ 0 & 0 & 1 & 1.4 \\ 0 & 0 & 1 & 2.3 \\ 0 & 0 & 1 & 3.5 \end{bmatrix} \begin{bmatrix} \tau_{U_0}^2 & \tau_{U_{01}} & 0 & 0 \\ \tau_{U_{01}} & \tau_{U_1}^2 & 0 & 0 \\ 0 & 0 & \tau_{U_0}^2 & \tau_{U_{01}} \\ 0 & 0 & \tau_{U_{01}} & \tau_{U_1}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0.0 & 1.0 & 2.0 & 3.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0.2 & 1.4 & 2.3 & 3.5 \end{bmatrix} + \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_e^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_e^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

- The “block diagonal” does not need to be the same size or contain the same time observations per person...
- **R** matrix can also include non-0 covariance or differential residual variance across time (as in ACS models), although the models based on the idea of a “lag” won’t work for unbalanced or unequal-interval time

G, R, and V: The Take-Home Point

- The partitioning of variance into piles...
 - **Level 2 = BP** → **G** matrix of random effects variances/covariances
 - **Level 1 = WP** → **R** matrix of residual variances/covariances
 - **G** and **R** combine via **Z** to create **V** matrix of total variances/covariances
 - Many flexible options that allows the variances and covariances to vary in a time-dependent way that better matches the actual data
 - Can allow variance and covariance due to other predictors, too



Describing Within-Person Change in Longitudinal Data

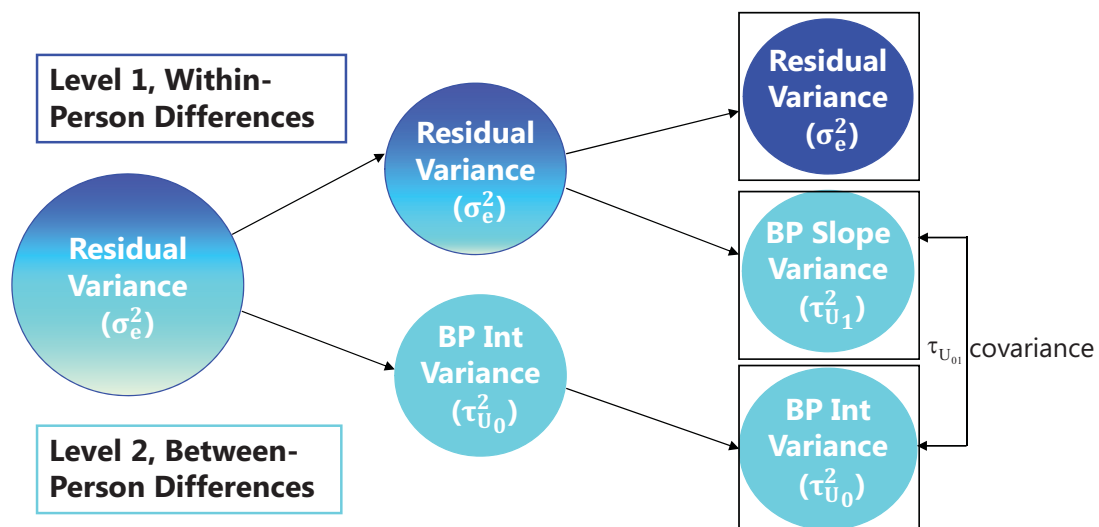
- Topics:
 - Multilevel modeling notation and terminology
 - Fixed and random effects of linear time
 - Predicted variances and covariances from random slopes
 - **Dependency and effect size in random effects models**
 - Describing nonlinear change: polynomial, piecewise models, and truly nonlinear models
 - Fun with likelihood estimation and model comparisons

How MLM “Handles” Dependency

- Common description of the purpose of MLM is that it “addresses” or “handles” correlated (dependent) data...
- But where does this correlation come from?
3 places (here, an example with health as an outcome):
 1. *Mean differences across persons*
 - Some people are just healthier than others (at every time point)
 - This is what a random intercept is for
 2. *Differences in effects of predictors across persons*
 - Does *time* (or *stress*) affect health more in some persons than others?
 - This is what random slopes are for
 3. Non-constant within-person correlation for unknown reasons
 - Occasions closer together may just be more related
 - This is what ACS models are for

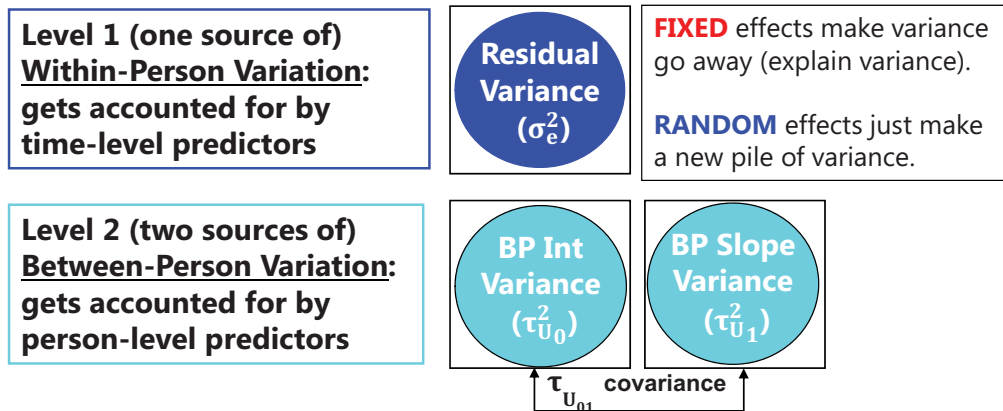
MLM “Handles” Dependency

- Where does each kind of person dependency go? Into a new random effects variance component (or “pile” of variance):



Piles of Variance

- By adding a random slope, we **carve up** our total variance into 3 piles:
 - BP (error) variance around intercept
 - BP (error) variance around slope
 - WP (error) residual variance
- These 2 piles are 1 pile of "error variance" in Univ. RM ANOVA
- But making piles does NOT make error variance go away...**



Fixed vs. Random Effects of Persons

- Person dependency: via **fixed effects in the model for the means** or via **random effects in the model for the variance**?
 - Individual intercept differences can be included as:
 - N-1 person dummy code fixed main effects** OR **1 random U_{0i}**
 - Individual time slope differences can be included as:
 - N-1*time person dummy code interactions** OR **1 random U_{1i} *time_{ti}**
 - Either approach would appropriately control for dependency (fixed effects are used in some programs that 'control' SEs for sampling)
- Two important advantages of **random effects**:
 - Quantification: Direct measure of how much of the outcome variance is due to person differences (in intercept or in effects of predictors)
 - Prediction: Person differences (main effects and effects of time) then become predictable quantities – this can't happen using fixed effects
 - Summary: Random effects give you predictable control of dependency**

Explained Variance from Fixed Linear Time

- Most common measure of effect size in MLM is Pseudo- R^2
 - Is supposed to be variance accounted for by predictors
 - Multiple piles of variance mean multiple possible values of pseudo R^2 (can be calculated per variance component or per model level)
 - A fixed linear effect of time will reduce level-1 residual variance σ_e^2 in **R**
 - By how much is the residual variance σ_e^2 reduced?

$$\text{Pseudo } R_e^2 = \frac{\text{residual variance}_{\text{fewer}} - \text{residual variance}_{\text{more}}}{\text{residual variance}_{\text{fewer}}}$$

- If time varies between persons, then level-2 random intercept variance $\tau_{U_0}^2$ in **G** may also be reduced:

$$\text{Pseudo } R_{U_0}^2 = \frac{\text{random intercept variance}_{\text{fewer}} - \text{random intercept variance}_{\text{more}}}{\text{random intercept variance}_{\text{fewer}}}$$

- But you are likely to see a (net) INCREASE in $\tau_{U_0}^2$ instead.... Here's why:

Increases in Random Intercept Variance

- Level-2 random intercept variance $\tau_{U_0}^2$ will often increase as a consequence of reducing level-1 residual variance σ_e^2
- Observed level-2 $\tau_{U_0}^2$ is NOT just between-person variance
 - Also has a small part of within-person variance (level-1 σ_e^2), or:
Observed $\tau_{U_0}^2 = \text{True } \tau_{U_0}^2 + (\sigma_e^2/n)$
 - As n occasions increases, bias of level-1 σ_e^2 is minimized
 - Likelihood-based estimates of "true" $\tau_{U_0}^2$ use (σ_e^2/n) as correction factor:
True $\tau_{U_0}^2 = \text{Observed } \tau_{U_0}^2 - (\sigma_e^2/n)$
- For example: observed level-2 $\tau_{U_0}^2 = 4.65$, level-1 $\sigma_e^2 = 7.06$, $n = 4$
 - True $\tau_{U_0}^2 = 4.65 - (7.06/4) = 2.88$ in empty means model
 - Add fixed linear time slope → reduce σ_e^2 from 7.06 to 2.17 ($R^2 = .69$)
 - But now True $\tau_{U_0}^2 = 4.65 - (2.17/4) = 4.10$ in fixed linear time model

Quantification of Random Effects Variances

- We can test if a random effect variance is significant, but the variance estimates are not likely to have inherent meaning
 - e.g., "I have a significant fixed linear time effect of $\gamma_{10} = 1.72$, so people increase by 1.72/time on average. I also have a significant random linear time slope variance of $\tau_{U_1}^2 = 0.91$, so people need their own slopes (people change differently). But how much is a variance of 0.91 , really?"
- **95% Random Effects Confidence Intervals** can tell you
 - Can be calculated for each effect that is random in your model
 - Provide range around the fixed effect within which 95% of your sample is predicted to fall, based on your random effect variance:
$$\text{Random Effect 95\% CI} = \text{fixed effect} \pm (1.96 * \sqrt{\text{Random Variance}})$$
$$\text{Linear Time Slope 95\% CI} = \gamma_{10} \pm (1.96 * \sqrt{\tau_{U_1}^2}) \rightarrow 1.72 \pm (1.96 * \sqrt{0.91}) = -0.15 \text{ to } 3.59$$
 - So although people improve on average, individual slopes are predicted to range from -0.15 to 3.59 (so some people may actually decline)

Describing Within-Person Change in Longitudinal Data

- Topics:
 - Multilevel modeling notation and terminology
 - Fixed and random effects of linear time
 - Predicted variances and covariances from random slopes
 - Dependency and effect size in random effects models
 - **Describing nonlinear change: polynomial, piecewise models, and truly nonlinear models**
 - Fun with likelihood estimation and model comparisons

Summary: Modeling Means and Variances

- We have two tasks in describing within-person change:
- **Choose a Model for the Means**
 - What kind of change in the outcome do we have **on average**?
 - What kind and how many **fixed effects** do we need to predict that mean change as parsimoniously but accurately as possible?
- **Choose a Model for the Variances**
 - What pattern do the variances and covariances of the outcome show over time because of **individual differences** in change?
 - What kind and how many **random effects** do we need to predict that pattern as parsimoniously but accurately as possible?

Name that trajectory... Polynomial?

- Predict **mean change** with **polynomial fixed effects of time**:
 - Linear → *constant* amount of change (up or down)
 - Quadratic → *change* in linear rate of change (acceleration/deceleration)
 - Cubic → *change* in acceleration/deceleration of linear rate of change (known in physics as jerk, surge, or jolt)
 - Terms work together to describe curved trajectories
- **Can have polynomial fixed time slopes UP TO: $n - 1$ ***
 - 3 occasions = 2nd order (time²) = Fixed Quadratic Time or less
 - 4 occasions = 3rd order (time³) = Fixed Cubic Time or less
- Interpretable polynomials past cubic are rarely seen in practice
- * $n - 1$ rule can be broken in unbalanced data (but cautiously)

Interpreting Quadratic Fixed Effects

A Quadratic time effect is a two-way interaction: time*time

- Fixed quadratic time = "**half the rate of acceleration/deceleration**"
- So to interpret it as how the linear time effect changes per unit time, **you must multiply the quadratic coefficient by 2**
- If fixed linear time slope = 4 at time 0, with quadratic slope = 0.3?
 - Instantaneous linear rate of Δ at time 0 = 4.0, at time 1 = 4.6...
- The "twice" part comes from taking the derivatives of the function:

| | |
|---|--|
| Intercept (Position) at Time T: | $\hat{y}_T = 50.0 + 4.0T + 0.3T^2$ |
| First Derivative (Velocity) at Time T: | $\frac{d\hat{y}_T}{d(T)} = 4.0 + 0.6T$ |
| Second Derivative (Acceleration) at Time T: | $\frac{d^2\hat{y}_T}{d(T)} = 0.6$ |

Interpreting Quadratic Fixed Effects

A Quadratic time effect is a two-way interaction: time*time

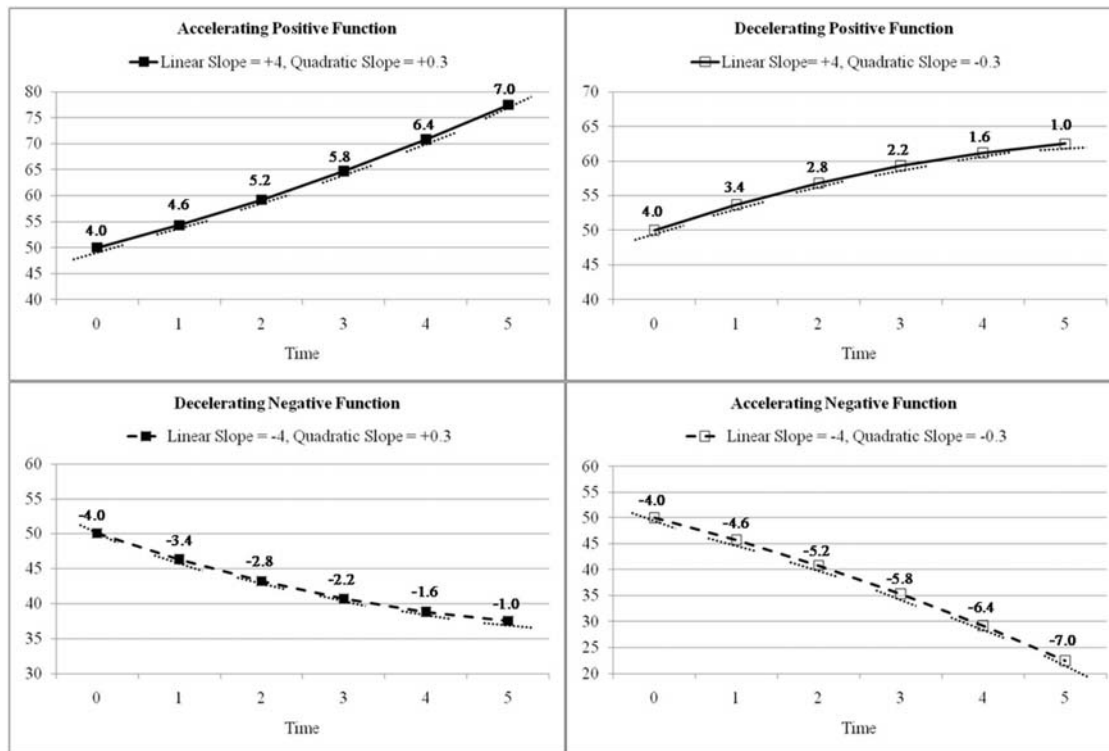
- Fixed quadratic time = "**half the rate of acceleration/deceleration**"
- So to interpret it as how the linear time effect changes per unit time, **you must multiply the quadratic coefficient by 2**
- If fixed linear time slope = 4 at time 0, with quadratic slope = 0.3?
 - Instantaneous linear rate of Δ at time 0 = 4.0, at time 1 = 4.6...

- The "twice" part also comes from what you remember about the role of interactions with respect to their constituent main effects:

| |
|--|
| $\hat{y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ$ |
| Effect of X = $\beta_1 + \beta_3 Z$ |
| Effect of Z = $\beta_2 + \beta_3 X$ |
| $\hat{y}_T = \beta_0 + \beta_1 \text{Time}_T + \text{_____} + \beta_3 \text{Time}_T^2$ |
| Effect of Time _T = $\beta_1 + 2\beta_3 \text{Time}_T$ |

- Because time is interacting with itself, there is no second main effect in the model for the interaction to modify as usual. So the quadratic time effect gets applied twice to the one (main) linear effect of time.

Examples of Fixed Quadratic Time Effects



Lecture 2

41

Conditionality of Polynomial Fixed Time Effects

- We've seen how main effects become conditional simple effects once they are part of an interaction
- The same is true for polynomial **fixed effects of time**:
 - **Fixed Intercept Only?**
 - Fixed Intercept = predicted mean of Y for any occasion (= grand mean)
 - **Add Fixed Linear Time?**
 - Fixed Intercept = **now** predicted mean of Y from linear time at time=0 (would be different if time was centered elsewhere)
 - Fixed Linear Time = mean linear rate of change across all occasions (would be the same if time was centered elsewhere)
 - **Add Fixed Quadratic Time?**
 - Fixed Intercept = still predicted mean of Y at time=0 (but from quadratic model) (would be different if time was centered elsewhere)
 - Fixed Linear Time = **now** mean linear rate of change at time=0 (would be different if time was centered elsewhere)
 - Fixed Quadratic Time = half the mean rate of acceleration or deceleration of change across all occasions (i.e., the linear slope changes the same over time)

Lecture 2

42

Polynomial **Fixed** vs. **Random** Time Effects

- **Polynomial fixed effects** combine to describe mean trajectory over time (can have fixed slopes up to $n - 1$):
 - Fixed Intercept = Predicted mean level (at time 0)
 - Fixed Linear Time = Mean linear rate of change (at time 0)
 - Fixed Quadratic Time = Half of mean acceleration/deceleration in linear rate of change (2*quad is how the linear time slope changes per unit time if quadratic is highest order fixed effect of time)
- **Polynomial random effects** (individual deviations from the fixed effect) describe individual differences in those change parameters (can have random slopes up to $n - 2$):
 - Random Intercept = BP variance in level (at time 0)
 - Random Linear Time = BP variance in linear time slope (at time 0)
 - Random Quadratic Time = BP variance in half the rate of acceleration/deceleration of linear time slope (across all time if quadratic is highest-order random effect of time)

Lecture 2

43

Random Quadratic Time Model

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}\text{Time}_{ti} + \beta_{2i}\text{Time}_{ti}^2 + e_{ti}$

Level 2 Equations (one per β):

$$\beta_{0i} = \overset{\text{Intercept for person } i}{\uparrow} \overset{\text{Fixed (mean) Intercept}}{\uparrow} Y_{00} + \overset{\text{Random (Deviation) Intercept}}{\uparrow} U_{0i}$$

Fixed Effect Subscripts:

1st = which Level 1 term

2nd = which Level 2 term

$$\beta_{1i} = \overset{\text{Linear Slope for person } i}{\uparrow} \overset{\text{Fixed (mean) Linear Slope}}{\uparrow} Y_{10} + \overset{\text{Random (Deviation) Linear Slope}}{\uparrow} U_{1i}$$

Number of Possible Slopes by Number of Occasions (n):

Fixed slopes = $n - 1$

Random slopes = $n - 2$

$$\beta_{2i} = \overset{\text{Quad Slope for person } i}{\uparrow} \overset{\text{Fixed (mean) Quad Slope}}{\uparrow} Y_{20} + \overset{\text{Random (Deviation) Quad Slope}}{\uparrow} U_{2i}$$

Need $n = 4$ occasions to fit random quadratic time model

Lecture 2

44

Conditionality of Polynomial Random Effects

- We saw previously that lower-order fixed effects of time are conditional on higher-order polynomial fixed effects of time
- The same is true for polynomial **random effects of time**:
 - **Random Intercept Only?**
 - Random Intercept = BP variance *for any occasion* in predicted mean Y
(= variance in grand mean because individual lines are parallel)
 - **Add Random Linear Time?**
 - Random Intercept = **now** BP variance *at time=0* in predicted mean Y
(*would be different if time was centered elsewhere*)
 - Random Linear Time = BP variance *across all occasions* in linear rate of change
(*would be the same if time was centered elsewhere*)
 - **Add Random Quadratic Time?**
 - Random Intercept = still BP variance *at time=0* in predicted mean Y
 - Random Linear Time = **now** BP variance *at time=0* in linear rate of change
(*would be different if time was centered elsewhere*)
 - Random Quadratic Time = BP variance *across all occasions* in half of accel/decel of change
(*would be the same if time was centered elsewhere*)

Random Effects Allowed by #Occasions

| | <u>Data</u> | <u>G Matrix</u> | <u>R Matrix</u> | <u>Variance Model # Parameters</u> |
|--|---|--|--|------------------------------------|
| <u>$n=2$ occasions</u> 3 unique pieces of information | $\begin{bmatrix} \sigma_1^2 & & \\ \sigma_{21} & \sigma_2^2 & \\ & & \end{bmatrix}$ | $\begin{bmatrix} \tau_{U_0}^2 \\ \\ \text{Random Intercept only} \end{bmatrix}$ | $\begin{bmatrix} \sigma_e^2 & 0 \\ 0 & \sigma_e^2 \end{bmatrix}$ | 2 |
| <u>$n=3$ occasions</u> 6 unique pieces of information | $\begin{bmatrix} \sigma_1^2 & & & \\ \sigma_{21} & \sigma_2^2 & & \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \\ & & & \end{bmatrix}$ | $\begin{bmatrix} \tau_{U_0}^2 & & \\ \tau_{U_{01}} & \tau_{U_1}^2 & \\ \text{Up to 1 Random slope} & & \end{bmatrix}$ | $\begin{bmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_e^2 & 0 \\ 0 & 0 & \sigma_e^2 \end{bmatrix}$ | 4 |
| <u>$n=4$ occasions</u> 10 unique pieces of information | $\begin{bmatrix} \sigma_1^2 & & & & \\ \sigma_{21} & \sigma_2^2 & & & \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & & \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 & \\ & & & & \end{bmatrix}$ | $\begin{bmatrix} \tau_{U_0}^2 & & & \\ \tau_{U_{01}} & \tau_{U_1}^2 & & \\ \tau_{U_{02}} & \tau_{U_{12}} & \tau_{U_2}^2 & \\ \text{Up to 2 Random slopes} & & & \end{bmatrix}$ | $\begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$ | 7 |

Predicted **V** Matrix from Polynomial Random Effects Models

- **Random linear model?** Variance has a **quadratic** dependence on time
 - Variance will be at a minimum when $\text{time} = -\text{Cov}(U_0, U_1) / \text{Var}(U_1)$, and will increase parabolically and symmetrically over time
 - **Predicted variance** at each occasion and covariance between A and B:

$$\text{Var}(y_{\text{time}}) = \text{Var}(e_t) + \text{Var}(U_0) + 2\text{Cov}(U_0, U_1)(\text{time}_t) + \text{Var}(U_1)(\text{time}_t^2)$$

$$\text{Cov}(y_A, y_B) = \text{Var}(U_0) + \text{Cov}(U_0, U_1)(A + B) + \text{Var}(U_1)(AB)$$
- **Random quadratic model?** Variance has a **quartic** dependence on time

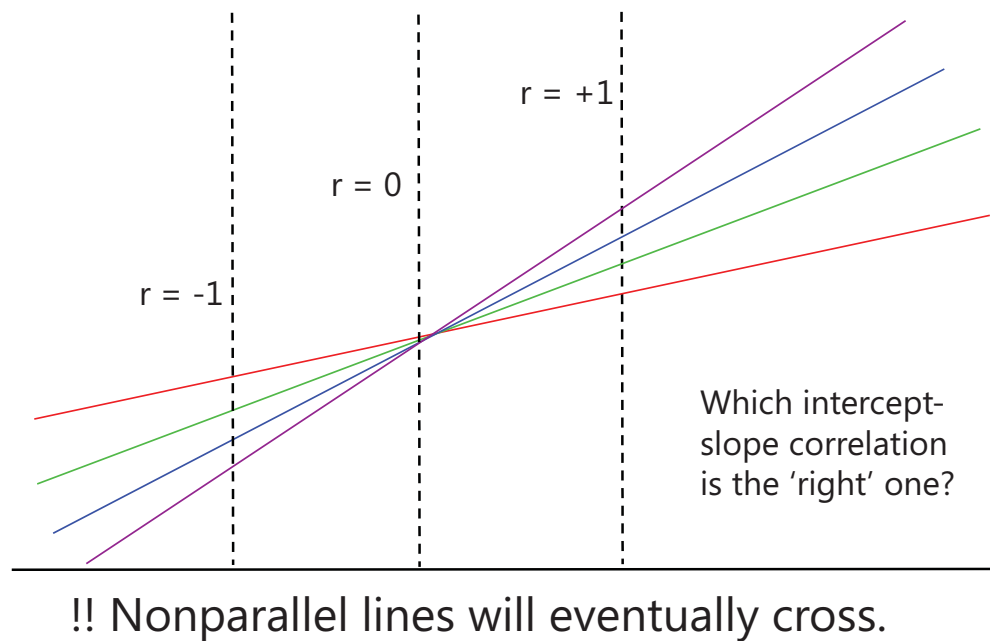
$$\text{Var}(y_{\text{time}}) = \text{Var}(e_t) + \text{Var}(U_0) + 2\text{Cov}(U_0, U_1)(\text{time}_t) + \text{Var}(U_1)(\text{time}_t^2) + 2\text{Cov}(U_0, U_2)(\text{time}_t^2) + 2\text{Cov}(U_1, U_2)(\text{time}_t^3) + \text{Var}(U_2)(\text{time}_t^4)$$

$$\text{Cov}(y_A, y_B) = \text{Var}(U_0) + \text{Cov}(U_0, U_1)(A + B) + \text{Var}(U_1)(AB) + \text{Cov}(U_0, U_2)(A^2 + B^2) + \text{Cov}(U_1, U_2)[(AB^2) + (A^2B)] + \text{Var}(U_2)(A^2B^2)$$
- *The point of the story: random effects of time are a way of allowing the variances and covariances to differ over time in specific, time-dependent patterns (that result from differential individual change over time).*

Rules for Polynomial Models (and in general for fixed and random effects)

- On the same side of the model (means or variances side), lower-order effects stay in EVEN IF NONSIGNIFICANT (for correct interpretation)
 - e.g., Significant *fixed* quadratic? Keep the *fixed* linear
 - e.g., Significant *random* quadratic? Keep the *random* linear
- Also remember—you can have a significant random effect EVEN IF the corresponding fixed effect is not significant (keep it anyway):
 - e.g., Fixed linear not significant, but random linear is significant?
 - No linear change *on average*, but significant individual differences in change
- Language: A random effect supersedes a fixed effect:
 - If Fixed = intercept, linear, quad; Random = intercept, linear, quad?
 - Call it a "Random quadratic model" (implies everything beneath those terms)
 - If Fixed = intercept, linear, quad; Random = intercept, linear?
 - Call it a "Fixed quadratic, random linear model" (distinguishes no random quad)
- Intercept-slope correlation depends largely on centering of time...

Correlation between Random Intercept and Random Linear Slope depends on time 0

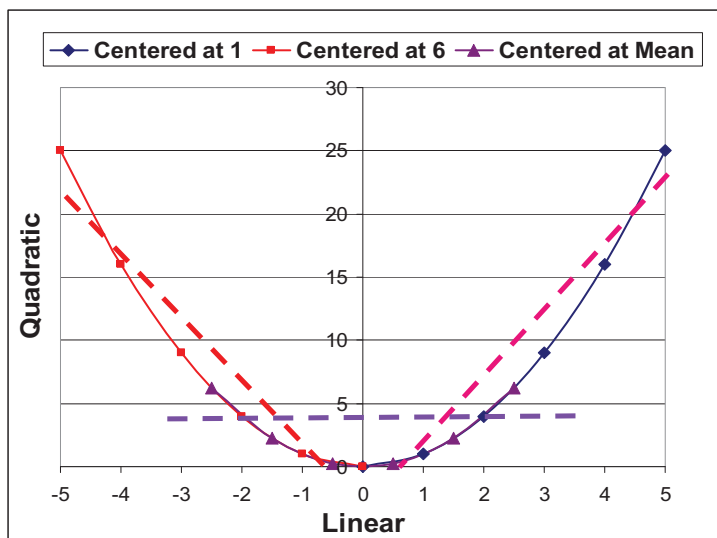


Correlations among polynomial slopes

| Session | Linear | Quadratic |
|---------|--------|-----------|
| 1 | 0 | 0 |
| 2 | 1 | 1 |
| 3 | 2 | 4 |
| 4 | 3 | 9 |
| 5 | 4 | 16 |
| 6 | 5 | 25 |

| Session | Linear | Quadratic |
|---------|--------|-----------|
| 1 | -5 | 25 |
| 2 | -4 | 16 |
| 3 | -3 | 9 |
| 4 | -2 | 4 |
| 5 | -1 | 1 |
| 6 | 0 | 0 |

| Session | Linear | Quadratic |
|---------|--------|-----------|
| 1 | -2.5 | 6.25 |
| 2 | -1.5 | 2.25 |
| 3 | -0.5 | 0.25 |
| 4 | 0.5 | 0.25 |
| 5 | 1.5 | 2.25 |
| 6 | 2.5 | 6.25 |

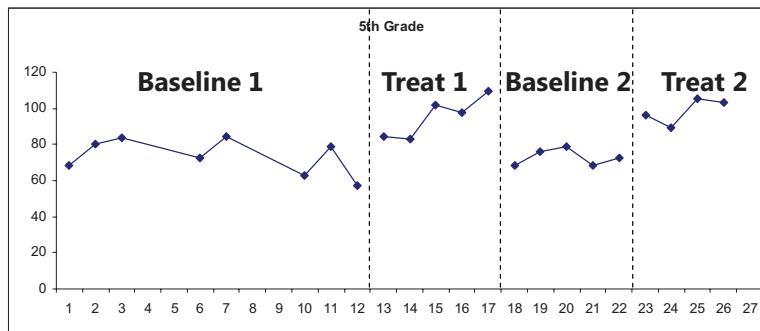


Correlations among polynomial effects of time can be induced by centering time near the start or near the end.

Therefore, these correlations will be **most** interpretable when centering time at its mean instead.

Other Random Effects Models of Change

- **Piecewise models:** Discrete slopes for discrete phases of time
 - Separate terms describe sections of overall trajectories
 - Useful for examining change in intercepts and slopes before/after discrete events (changes in policy, interventions)
 - **Must know where the break point is ahead of time!**

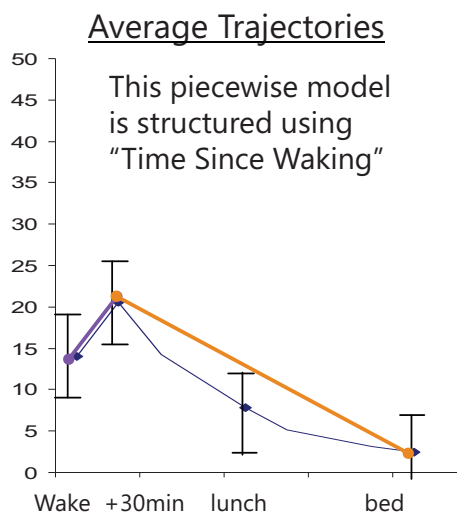


Piecewise Model:

4 slopes
(one per phase)

3 "jumps"
(shift in intercept
between phases)

Example of Daily Cortisol Fluctuation: Morning Rise and Afternoon Decline



SAS Code to create two piecewise slopes from continuous time of day in stacked data:

```
IF occasion=1 THEN DO;
  P1=0; P2=0; END;
IF occasion=2 THEN DO;
  P1= time2-time1; P2=0; END;
IF occasion=3 THEN DO;
  P1= time2-time1; P2=time3-time2; END;
IF occasion=4 THEN DO;
  P1= time2-time1; P2=time4-time2; END;
```

Note that a quadratic slope may be necessary for the afternoon decline slope!

Random Two-Slope Piecewise Model

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}\text{Slope1}_{ti} + \beta_{2i}\text{Slope2}_{ti} + e_{ti}$

Level 2 Equations (one per β):

$$\beta_{0i} = \underset{\substack{\uparrow \\ \text{Intercept} \\ \text{for person } i}}{\beta_{0i}} = \underset{\substack{\uparrow \\ \text{Fixed (mean)} \\ \text{Intercept}}}{Y_{00}} + \underset{\substack{\uparrow \\ \text{Random} \\ \text{(Deviation)} \\ \text{Intercept}}}{U_{0i}}$$

Fixed Effect Subscripts:

1st = which Level 1 term

2nd = which Level 2 term

$$\beta_{1i} = \underset{\substack{\uparrow \\ \text{Slope1} \\ \text{for person } i}}{\beta_{1i}} = \underset{\substack{\uparrow \\ \text{Fixed (mean)} \\ \text{Slope1}}}{Y_{10}} + \underset{\substack{\uparrow \\ \text{Random} \\ \text{(Deviation)} \\ \text{Slope1}}}{U_{1i}}$$

Number of Possible Slopes by Number of Occasions (n):

Fixed slopes = $n - 1$

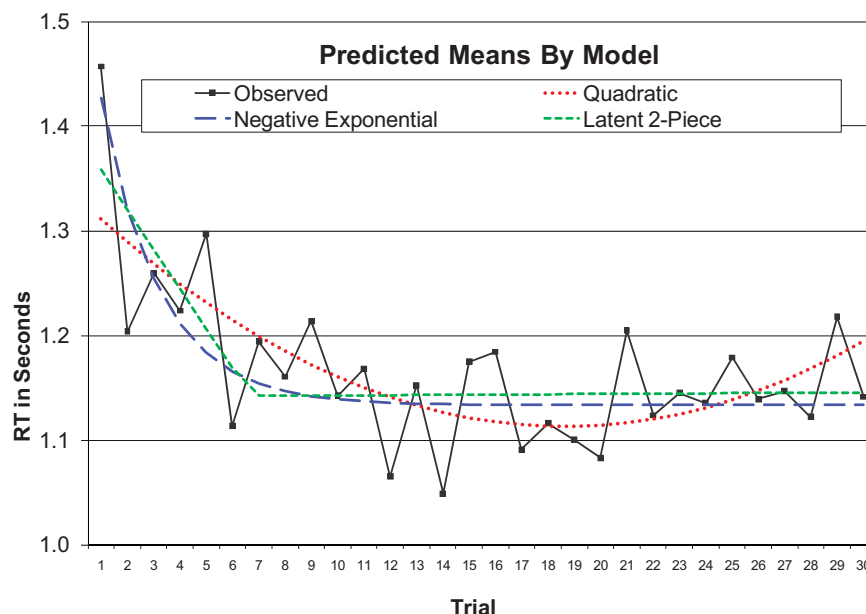
Random slopes = $n - 2$

$$\beta_{2i} = \underset{\substack{\uparrow \\ \text{Slope2} \\ \text{for person } i}}{\beta_{2i}} = \underset{\substack{\uparrow \\ \text{Fixed (mean)} \\ \text{Slope2}}}{Y_{20}} + \underset{\substack{\uparrow \\ \text{Random} \\ \text{(Deviation)} \\ \text{Slope2}}}{U_{2i}}$$

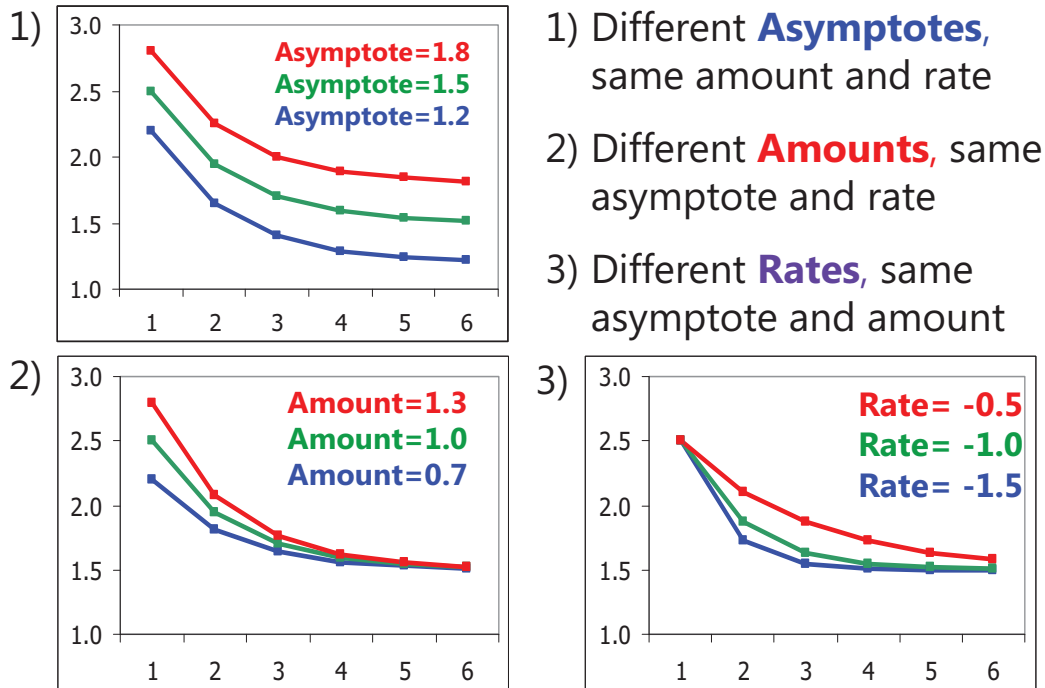
Need $n = 4$ occasions to fit random two-slope model

Other Random Effects for Change

- **Truly nonlinear models:** Non-additive terms to describe change
 - Models can include **asymptotes** (so change can "shut off" as needed)
 - Include **power** and **exponential** functions (see chapter 6 for references)



(Negative) Exponential Model Parameters



Lecture 2

55

Exponential Model (3 Random Effects)

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i} \cdot \exp(\beta_{2i} \cdot \text{Time}_{ti}) + e_{ti}$

Level 2 Equations (one per β):

$$\beta_{0i} = Y_{00} + U_{0i}$$

Asymptote for person i Fixed (mean) Asymptote Random (Deviation) Asymptote

$$\beta_{1i} = Y_{10} + U_{1i}$$

Amount for person i Fixed (mean) Amount Random (Deviation) Amount

$$\beta_{2i} = Y_{20} + U_{2i}$$

Rate for person i Fixed (mean) Rate Random (Deviation) Rate

Fixed Effect Subscripts:

1st = which Level 1 term

2nd = which Level 2 term

Number of Possible Slopes by Number of Occasions (n):

Fixed slopes = $n - 1$

Random slopes = $n - 2$

Also need 4 occasions to fit random exponential model

(Likely need way more occasions to find U_{2i} , though)

Lecture 2

56

Describing Within-Person Change in Longitudinal Data

- Topics:
 - Multilevel modeling notation and terminology
 - Fixed and random effects of linear time
 - Predicted variances and covariances from random slopes
 - Dependency and effect size in random effects models
 - Describing nonlinear change: polynomial, piecewise models, and truly nonlinear models
 - **Fun with likelihood estimation and model comparisons**

3 Decision Points for Model Comparisons

1. Are the models **nested** or **non-nested**?
 - Nested: have to add OR subtract effects to go from one to other
 - Can conduct significance tests for improvement in fit
 - Non-nested: have to add AND subtract effects
 - No significance tests available for these comparisons
2. Differ in model for the **means, variances, or both**?
 - Means? Can only use ML $-2\Delta LL$ tests (or p -value of each fixed effect)
 - Variances? Can use ML (or preferably REML) $-2\Delta LL$ tests, no p -values
 - Both sides? Can only use ML $-2\Delta LL$ tests
3. Models estimated using **ML** or **REML**?
 - ML: All model comparisons are ok
 - REML: Model comparisons are ok for the variance parameters only

Likelihood-Based Model Comparisons

- Relative model fit is indexed by a “**deviance**” statistic → **-2LL**
 - Log of likelihood (**LL = total data height**) of observing the data given model parameters, $-2*LL$ so that the differences between model LL values follow $\sim\chi^2$
 - **-2LL is a measure of BADNESS of fit, so smaller values = better models**
 - Models are compared using their deviance values (significance tests)
 - Two estimation flavors (labeled as $-2 \log$ likelihood in SAS, SPSS, but given as LL instead in STATA): Maximum Likelihood (**ML**) or Restricted (Residual) ML (**REML**)
- Fit is also indexed by **Information Criteria** that reflect **-2LL** deviance AND # parameters used and/or sample size
 - **AIC** = Akaike IC = $-2LL + 2 * (\#parameters)$
 - **BIC** = Bayesian IC = $-2LL + \log(N) * (\#parameters)$ → penalty for complexity
 - In ML → #parameters = all parameters (means and variances models)
 - In REML → #parameters = variance model parameters only (except in STATA!)
 - No significance tests or critical values, just “smaller is better”

-2ΔLL (i.e., LRT, Deviance) Tests: (models must use the same estimator & N)

1. Calculate $-2\Delta LL$: $(-2LL_{\text{fewer}}) - (-2LL_{\text{more}})$
 2. Calculate Δdf : $(\# \text{Parms}_{\text{more}}) - (\# \text{Parms}_{\text{fewer}})$
 3. Compare $-2\Delta LL$ to χ^2 distribution with $df = \Delta df$
CHIDIST function in excel will give exact p-values for the difference test; so will STATA
1. & 2. must be positive values!
- Fixed effects $p < .05$: $-2\Delta LL(1) > 3.84$, $-2\Delta LL(2) > 5.99$, $-2\Delta LL(3) > 7.82$
 - Some controversy about $-2\Delta LL$ tests when testing random effects variances that cannot be negative (i.e., the “boundary problem”)
 - χ^2 is not distributed as usual (mean=df) → is actually a mixture χ^2 with df and df-1, so using the critical χ^2 for actual df results in conservative model comparison test
 - e.g., $-2\Delta LL(df=2) > 5.99$, whereas $-2\Delta LL(df=\text{mixture of } 1,2) > 5.14$
 - Two proposed solutions when testing random effects variances:
 - For random intercepts, can use a 1-tailed test (χ^2 for $p < .10$): $-2\Delta LL(1) > 2.71$
 - Use mixture p-value = $0.5 * \text{prob}(\chi^2_{df-1} > -2\Delta LL) + 0.5 * \text{prob}(\chi^2_{df} > -2\Delta LL)$
 - In practice these assume no relationship among how well variance parameters are estimated, which is suspect → I tend to just use the conservative test and call it good

Critical Values for 50:50 χ^2 Mixtures

| df (q) | Significance Level | | | | |
|-----------|--------------------|-------|-------|-------|-------|
| | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
| 0 vs. 1 | 1.64 | 2.71 | 3.84 | 5.41 | 6.63 |
| 1 vs. 2 | 3.81 | 5.14 | 6.48 | 8.27 | 9.63 |
| 2 vs. 3 | 5.53 | 7.05 | 8.54 | 10.50 | 11.97 |
| 3 vs. 4 | 7.09 | 8.76 | 10.38 | 12.48 | 14.04 |
| 4 vs. 5 | 8.57 | 10.37 | 12.10 | 14.32 | 15.97 |
| 5 vs. 6 | 10.00 | 11.91 | 13.74 | 16.07 | 17.79 |
| 6 vs. 7 | 11.38 | 13.40 | 15.32 | 17.76 | 19.54 |
| 7 vs. 8 | 12.74 | 14.85 | 16.86 | 19.38 | 21.23 |
| 8 vs. 9 | 14.07 | 16.27 | 18.35 | 20.97 | 22.88 |
| 9 vs. 10 | 15.38 | 17.67 | 19.82 | 22.52 | 24.49 |
| 10 vs. 11 | 16.67 | 19.04 | 21.27 | 24.05 | 26.07 |

This may work ok if only one new parameter is bounded ... for example:

+ Random Intercept
df=1: 2.71 vs. 3.84

+ Random Linear
df=2: 5.14 vs. 5.99

+ Random Quad
df=3: 7.05 vs. 7.82

Critical values such that the right-hand tail probability =
 $0.5 \times \Pr(\chi^2_q > c) + 0.5 \times \Pr(\chi^2_{q+1} > c)$

Source: Appendix C (p. 484) from Fitzmaurice, Laird, & Ware (2004).
Applied Longitudinal Analysis. Hoboken, NJ: Wiley

ML vs. REML

Remember "population"
vs. "sample" formulas for
calculating variance?

Population: $\sigma_e^2 = \frac{\sum_{i=1}^N (y_i - \mu)^2}{N}$

Sample: $\sigma_e^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N-1}$

| All comparisons must have same N!!! | ML | REML |
|--|--|---|
| To select, type... | METHOD=ML (-2 log likelihood) | METHOD=REML <i>default</i> (-2 res log likelihood) |
| In estimating variances, it treats fixed effects as... | Known (df for having to also estimate fixed effects is not factored in) | Unknown (df for having to estimate fixed effects is factored in) |
| So, in small samples, L2 variances will be... | Too small (less difference after N=30-50 or so) | Unbiased (correct) |
| But because it indexes the fit of the... | Entire model (means + variances) | Variances model only |
| You can compare models differing in... | Fixed and/or random effects (either/both) | Random effects only (same fixed effects) |

Rules for Comparing Multilevel Models

All observations must be the same across models!

Compare Models Differing In:

| Type of Comparison: | Means Model (Fixed) Only | Variance Model (Random) Only | Both Means and Variances Model (Fixed and Random) |
|--|--|--|---|
| <u>Nested?</u> YES, can do significance tests via... | Fixed effect p -values from ML or REML -- OR -- ML $-2\Delta LL$ only (NO REML $-2\Delta LL$) | NO p -values REML $-2\Delta LL$ (ML $-2\Delta LL$ is ok if big N) | ML $-2\Delta LL$ only (NO REML $-2\Delta LL$) |
| <u>Non-Nested?</u> NO signif. tests, instead see... | ML AIC, BIC (NO REML AIC, BIC) | REML AIC, BIC (ML ok if big N) | ML AIC, BIC only (NO REML AIC, BIC) |

Nested = one model is a direct subset of the other

Non-Nested = one model is not a direct subset of the other

Summary: Model Comparisons

- Significance of **fixed effects** can be tested with EITHER their **p -values** OR **ML $-2\Delta LL$** (LRT, deviance difference) tests
 - p -value → Is EACH of these effects significant? (fine under ML or REML)
 - ML $-2\Delta LL$ test → Does this SET of predictors make my model better?
 - REML $-2\Delta LL$ tests are *WRONG* for comparing models differing in fixed effects
- Significance of **random effects** can only be tested with **$-2\Delta LL$ tests** (preferably using REML; here ML is not wrong, but results in too small variance components and fixed effect SEs in smaller samples)
 - Can get p -values as part of output but **shouldn't** use them
 - #parms added (df) should always include the random effect covariances
- My recommended approach to building models:
 - Stay in REML (for best estimates), test new fixed effects with their p -values
 - THEN add new random effects, testing $-2\Delta LL$ against previous model

Example Sequence for Testing Fixed and Random Polynomial Effects of Time

Build up fixed and random effects simultaneously:

1. Empty Means, Random Intercept → to calculate ICC
2. Fixed Linear, Random Intercept → check fixed linear p -value
3. Random Linear → check $-2\Delta LL(df \approx 2)$ for random linear variance
4. Fixed Quadratic, Random Linear → check fixed quadratic p -value
5. Random Quadratic → check $-2\Delta LL(df \approx 3)$ for random quadratic variance
6.

*** In general: Can use **REML** for all models, so long as you:

- Test significance of new **fixed** effects by their **p -values**
- Test significance of new **random** effects in separate step by **$-2\Delta LL$**
- Also see if AIC and BIC are smaller when adding random effects

Time-Invariant Predictors in Longitudinal Models

- Topics:
 - **Missing predictors in MLM**
 - Effects of time-invariant predictors
 - Fixed, systematically varying, and random level-1 effects
 - Model building strategies and assessing significance

Summary of Steps in Unconditional Longitudinal Modeling

For all outcomes:

1. Empty Model; Calculate ICC
2. Decide on a metric of time
3. Decide on a centering point
4. Estimate means model and plot individual trajectories

If your outcome shows systematic change:

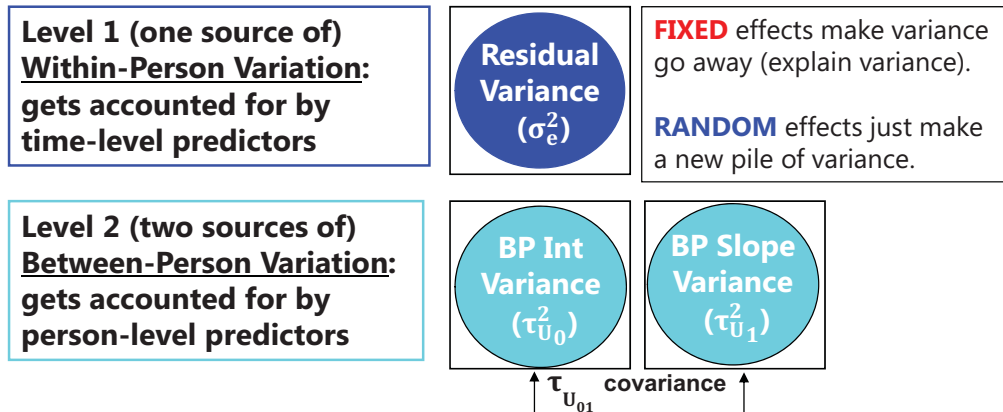
5. Evaluate fixed and random effects of time
6. Still consider possible alternative models for the residuals (**R** matrix)

If your outcome does NOT show ANY systematic change:

5. Evaluate alternative models for the variances (**G+R**, or **R**)

Random Effects Models for the Variance

- Each source of correlation or dependency goes into a new variance component (or pile of variance) until each source meets the usual assumptions of GLM: normality, independence, constant variance
- **Example 2-level longitudinal model:**



Now we get to add predictors to account for each pile!

Missing Data in MLM Software

- Common misconceptions about how MLM “handles” missing data
- Most MLM programs analyze only COMPLETE CASES
 - Does NOT require listwise deletion of *whole persons*
 - DOES delete any incomplete cases (occasions within a person)
- Observations missing predictors OR outcomes are not included!
 - **Time** is (probably) measured for **everyone**
 - **Predictors may NOT be measured for everyone**
 - *N* may change due to missing data for different predictors across models
- You may need to think about what predictors you want to examine PRIOR to model building, and pre-select your sample accordingly
 - Models and model fit statistics –2LL, AIC, and BIC are only **directly comparable** if they include the **exact same observations (LL is sum of each height)**
 - Will have less statistical power as a result of removing incomplete cases

Lecture 3

4

Be Careful of Missing Predictors!

**Multivariate
(wide) data
→ stacked
(long) data**

| ID | T1 | T2 | T3 | T4 | Person Pred | T1 Pred | T2 Pred | T3 Pred | T4 Pred |
|-----|----|----|----|----|----------------|------------|------------|------------|------------|
| 100 | 5 | 6 | 8 | 12 | 50 | 4 | 6 | 7 | . |
| 101 | 4 | 7 | . | 11 | . | 7 | . | 4 | 9 |

| Row | ID | Time | DV | Person Pred | Time Pred |
|-----|-----|------|----|----------------|--------------|
| 1 | 100 | 1 | 5 | 50 | 4 |
| 2 | 100 | 2 | 6 | 50 | 6 |
| 3 | 100 | 3 | 8 | 50 | 7 |
| 4 | 100 | 4 | 12 | 50 | . |
| 5 | 101 | 1 | 4 | . | 7 |
| 6 | 101 | 2 | 7 | . | . |
| 7 | 101 | 3 | . | . | 4 |
| 8 | 101 | 4 | 11 | . | 9 |

**Only rows with complete data
get used – for each model, which
rows get used in MIXED?**

Model with Time → DV: 1-6, 8

Model with Time,
Time Pred → DV: 1-3, 5, 8

Model with Time,
Person Pred → DV: 1-4

Model with Time,
Time Pred, &
Person Pred → DV: 1-3

Lecture 3

5

So what does this mean for missing data in MLM?

- **Missing outcomes are assumed MAR**
 - Because the likelihood function is for predicted Y, just estimated on whatever Y responses a person does have (can be incomplete)
- **Missing time-varying predictors are MAR-to-MCAR ish**
 - Would be MCAR because X is not in the likelihood function (is Y given X instead), but other occasions may have predictors (so MAR-ish)
- **Missing time-invariant predictors are assumed MCAR**
 - Because the predictor would be missing for all occasions, whole people will be deleted (may lead to bias)
- Missingness on predictors can be accommodated:
 - In Multilevel SEM with certain assumptions (\approx outcomes then)
 - Via multilevel multiple imputation in Mplus v 6.0+ (but careful!)
 - Must preserve all effects of potential interest in imputation model, including random effects; $-2\Delta LL$ tests are not done in same way

Time-Invariant Predictors in Longitudinal Models

- Topics:
 - Missing predictors in MLM
 - **Effects of time-invariant predictors**
 - Fixed, systematically varying, and random level-1 effects
 - Model building strategies and assessing significance

Modeling Time-Invariant Predictors

What independent variables can be time-invariant predictors?

- Also known as “person-level” or “level-2” predictors
- Include substantive predictors, controls, and predictors of missingness
- Can be anything that **does not change across time** (e.g., Biological Sex)
- Can be anything that **is not likely to change across the study**, but you may have to argue for this (e.g., Parenting Strategies, SES)
- Can be anything that **does change across the study**...
 - But you have **only measured once**
 - Limit conclusions to variable’s status at time of measurement
 - e.g., “Parenting Strategies at age 10”
 - Or **is perfectly correlated with time** (age, time to event)
 - Would use Age at Baseline, or Time to Event *from Baseline* instead

Centering Time-Invariant Predictors

- Very useful to center all predictors such that 0 is a meaningful value:
 - Same significance level of main effect, different interpretation of intercept
 - Different (more interpretable) main effects within higher-order interactions
 - With interactions, main effects = simple effects when other predictor = 0
- Choices for centering **continuous** predictors:
 - At Mean: Reference point is *average level of predictor within the sample*
 - Useful if predictor is on arbitrary metric (e.g., unfamiliar test)
 - Better → At Meaningful Point: Reference point is *chosen level of predictor*
 - Useful if predictor is already on a meaningful metric (e.g., age, education)
- Choices for centering **categorical** predictors:
 - Re-code group so that your chosen reference group = **reference (0) category!** (highest is the default in SAS and SPSS; lowest is default in STATA)
 - I do not recommend mean-centering categorical predictors (because who is at the mean of a categorical variable !!?)

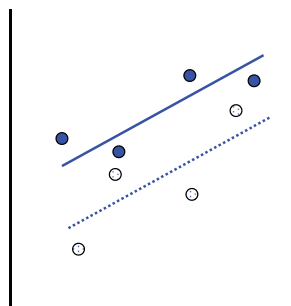
Main Effects of Predictors within Interactions

- Main effects of predictors within interactions should remain in the model regardless of whether or not they are significant
 - An interaction is an over-additive (enhancing) or under-additive (dampening) effect, so *what it is additive to* must be included
- The role of a two-way interaction is to adjust its main effects...
- However, the idea of a "main effect" no longer applies... each main effect is **conditional** on the interacting predictor = 0
- e.g., Model of $Y = W, X, Z, X*Z$:
 - The effect of W is still a "main effect" because it is not part of an interaction
 - The effect of X is now the conditional main effect of X *specifically when $Z=0$*
 - The effect of Z is now the conditional main effect of Z *specifically when $X=0$*
- The trick is keeping track of what 0 means for every interacting predictor, which depends on the way each predictor is being represented, as determined by you, or by the software without you!

The Role of Time-Invariant Predictors in the **Model for the Means**

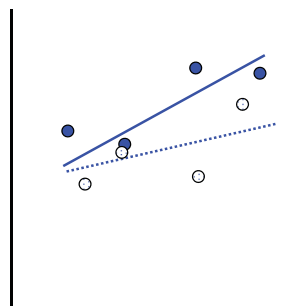
- **In Within-Person Change Models** → Adjust growth curve

Main effect of X , No interaction with time



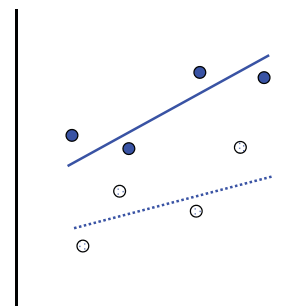
← Time →

Interaction with time, Main effect of X ?



← Time →

Main effect of X , and Interaction with time

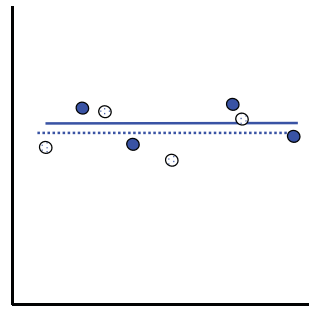


← Time →

The Role of Time-Invariant Predictors in the **Model for the Means**

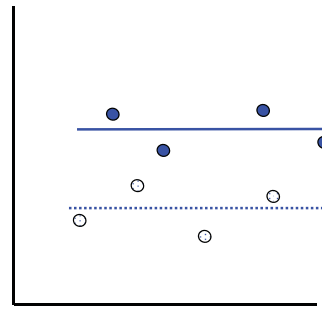
- In **Within-Person Fluctuation Models** → Adjust mean level

No main effect of X



← Time →

Main effect of X



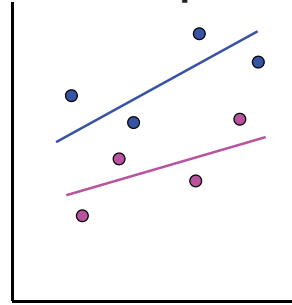
← Time →

The Role of Time-Invariant Predictors in the **Model for the Variance**

- In addition to fixed effects in the model for the means, time-invariant predictors can allow be used to allow **heterogeneity of variance** at their level or below
- e.g., Sex as a predictor of heterogeneity of variance:
 - **At level 2:** amount of individual differences in intercepts/slopes differs between boys and girls (i.e., one group is more variable)
 - **At level 1:** amount of within-person residual variation differs between boys and girls
 - In within-person **fluctuation** model: differential fluctuation over time
 - In within-person **change** model: differential fluctuation/variation remaining after controlling for fixed and random effects of time
- These models are harder to estimate and may require custom software (e.g., NLMIXED in SAS)

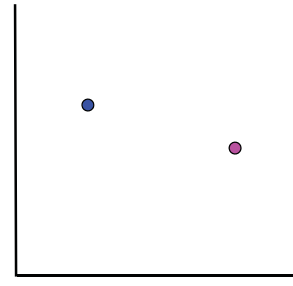
Why Level-2 Predictors Cannot Have Random Effects in 2-Level Models

Random Slopes for Time



Time
(or Any Level-1 Predictor)

Random Slopes for Sex?



Sex
(or any Level-2 Predictor)

You cannot make a line out of a dot, so level-2 effects cannot vary randomly over persons.

Education as a Time-Invariant Predictor: Example using a Random Quadratic Time Model

- Main Effect of Education = Education*Intercept Interaction
 - Moderates the intercept → Increase or decrease in expected outcome at time 0 for every year of education
- Effect of Education on Linear Time = Education*Time Interaction
 - Moderates the linear time slope → Increase or decrease in expected rate of change at time 0 for every year of education
- Effect of Education on Quadratic Time = Education*Time² Interaction
 - Moderates the quadratic time slope → Increase or decrease in half of expected acceleration/deceleration of linear rate of change for every year of education

Education (12 years = 0) as a Time-Invariant Predictor: Example using a Random Quadratic Time Model

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}\text{Time}_{ti} + \beta_{2i}\text{Time}_{ti}^2 + e_{ti}$

Level 2 Equations (one per β):

$$\beta_{0i} = \underset{\substack{\uparrow \\ \text{Intercept} \\ \text{for person } i}}{\beta_{0i}} = \underset{\substack{\uparrow \\ \text{Fixed Intercept} \\ \text{when Time}=0 \\ \text{and Ed}=12}}{Y_{00}} + \underset{\substack{\uparrow \\ \Delta \text{ in Intercept} \\ \text{per unit } \Delta \text{ in Ed}}}{Y_{01}\text{Ed}_i} + \underset{\substack{\uparrow \\ \text{Random (Deviation)} \\ \text{Intercept after} \\ \text{controlling for Ed}}}{U_{0i}}$$

$$\beta_{1i} = \underset{\substack{\uparrow \\ \text{Linear Slope} \\ \text{for person } i}}{\beta_{1i}} = \underset{\substack{\uparrow \\ \text{Fixed Linear} \\ \text{Time Slope} \\ \text{when Time}=0 \\ \text{and Ed}=12}}{Y_{10}} + \underset{\substack{\uparrow \\ \Delta \text{ in Linear Time} \\ \text{Slope per unit } \Delta \\ \text{in Ed (}=\text{Ed*time)}}}{Y_{11}\text{Ed}_i} + \underset{\substack{\uparrow \\ \text{Random (Deviation)} \\ \text{Linear Time Slope after} \\ \text{controlling for Ed}}}{U_{1i}}$$

$$\beta_{2i} = \underset{\substack{\uparrow \\ \text{Quad Slope} \\ \text{for person } i}}{\beta_{2i}} = \underset{\substack{\uparrow \\ \text{Fixed Quad} \\ \text{Time Slope} \\ \text{when Ed} = 12}}{Y_{20}} + \underset{\substack{\uparrow \\ \Delta \text{ in Quad Time} \\ \text{Slope per unit } \Delta \\ \text{in Ed (}=\text{Ed*time}^2)}}{Y_{21}\text{Ed}_i} + \underset{\substack{\uparrow \\ \text{Random (Deviation)} \\ \text{Quad Time Slope after} \\ \text{controlling for Ed}}}{U_{2i}}$$

Lecture 3

16

Education (12 years = 0) as a Time-Invariant Predictor: Example using a Random Quadratic Time Model

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}\text{Time}_{ti} + \beta_{2i}\text{Time}_{ti}^2 + e_{ti}$

Level 2 Equations (one per β):

$$\beta_{0i} = Y_{00} + Y_{01}\text{Ed}_i + U_{0i}$$

$$\beta_{1i} = Y_{10} + Y_{11}\text{Ed}_i + U_{1i}$$

$$\beta_{2i} = Y_{20} + Y_{21}\text{Ed}_i + U_{2i}$$

• Composite equation:

$$y_{ti} = (Y_{00} + Y_{01}\text{Ed}_i + U_{0i}) + (Y_{10} + Y_{11}\text{Ed}_i + U_{1i})\text{Time}_{ti} + (Y_{20} + Y_{21}\text{Ed}_i + U_{2i})\text{Time}_{ti}^2 + e_{ti}$$

Y_{11} and Y_{21} are known as
"cross-level" interactions
(level-1 predictor by
level-2 predictor)

Lecture 3

17

Time-Invariant Predictors in Longitudinal Models

- Topics:
 - Missing predictors in MLM
 - Effects of time-invariant predictors
 - **Fixed, systematically varying, and random level-1 effects**
 - Model building strategies and assessing significance

Fixed Effects of Time-Invariant Predictors

- Question of interest: Why do people change differently?
 - We're trying to predict individual differences in intercepts and slopes (i.e., reduce level-2 random effects variances)
 - So level-2 random effects variances become 'conditional' on predictors
→ actually random effects variances *left over*

$$\begin{array}{lcl}
 \beta_{0i} = \mathbf{Y}_{00} + \mathbf{U}_{0i} & \longrightarrow & \beta_{0i} = \mathbf{Y}_{00} + \mathbf{Y}_{01}\mathbf{Ed}_i + \mathbf{U}_{0i} \\
 \beta_{1i} = \mathbf{Y}_{10} + \mathbf{U}_{1i} & & \beta_{1i} = \mathbf{Y}_{10} + \mathbf{Y}_{11}\mathbf{Ed}_i + \mathbf{U}_{1i} \\
 \beta_{2i} = \mathbf{Y}_{20} + \mathbf{U}_{2i} & & \beta_{2i} = \mathbf{Y}_{20} + \mathbf{Y}_{21}\mathbf{Ed}_i + \mathbf{U}_{2i}
 \end{array}$$

- Can calculate pseudo- R^2 for each level-2 random effect variance between models with *fewer* versus *more* parameters as:

$$\text{Pseudo } R^2 = \frac{\text{random variance}_{\text{fewer}} - \text{random variance}_{\text{more}}}{\text{random variance}_{\text{fewer}}}$$

Fixed Effects of Time-Invariant Predictors

- What about predicting level-1 effects with no random variance?
 - If the random linear time slope is n.s., can I test interactions with time?

This should be ok to do...

$$\beta_{0i} = \gamma_{00} + \gamma_{01}Ed_i + U_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}Ed_i + U_{1i}$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21}Ed_i + U_{2i}$$

Is this still ok to do?

$$\beta_{0i} = \gamma_{00} + \gamma_{01}Ed_i + U_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}Ed_i$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21}Ed_i$$

- YES, surprisingly enough....
- **In theory**, if a level-1 effect does not vary randomly over individuals, then it has "no" variance to predict (so cross-level interactions with that level-1 effect are not necessary)
- However, because power to detect random effects is often lower than power to detect fixed effects, fixed effects of predictors can still be significant even if there is "no" (≈ 0) variance for them to predict
- Just make sure you test for random effects BEFORE testing any cross-level interactions with that level-1 predictor!

3 Types of Effects: **Fixed**, **Random**, and **Systematically (Non-Randomly) Varying**

Let's say we have a significant fixed linear effect of time.
What happens after we test a sex*time interaction?

| | Non-Significant Sex*Time effect? | Significant Sex*Time effect? |
|---|--|--|
| Random time slope initially not significant | Linear effect of time is FIXED | Linear effect of time is systematically varying |
| Random time initially sig, not sig. after sex*time | --- | Linear effect of time is systematically varying |
| Random time initially sig, still sig. after sex*time | Linear effect of time is RANDOM | Linear effect of time is RANDOM |

The effects of level-1 predictors (time-level) can be fixed, random, or systematically varying. The effects of level-2 predictors (person-level) can only be fixed or systematically varying (nothing to be random over...yet).

Variance Accounted For By Level-2 Time-Invariant Predictors

- **Fixed effects of level 2 predictors *by themselves*:**
 - L2 (BP) main effects (e.g., sex) reduce L2 (BP) random intercept variance
 - L2 (BP) interactions (e.g., sex by ed) also reduce L2 (BP) random intercept variance
- **Fixed effects of *cross-level interactions* (level 1* level 2):**
 - If the interacting level-1 predictor is random, any cross-level interaction with it will reduce its corresponding level-2 BP random slope variance
 - e.g., if *time* is random, then *sex*time*, *ed*time*, and *sex*ed*time* can each reduce the random linear time slope variance
 - If the interacting level-1 predictor not random, any cross-level interaction with it will reduce the level-1 WP residual variance instead
 - e.g., if *time*² is fixed, then *sex*time*², *ed*time*², and *sex*ed*time*² will reduce the L1 (WP) residual variance → Different quadratic slopes from sex and ed will allow better trajectories, reduce the variance around trajectories

Variance Accounted for... For Real

- **Pseudo-R²** is named that way for a reason... piles of variance can shift around, such that it can actually be negative
 - Sometimes a sign of model mis-specification
 - Hard to explain to readers when it happens!
- **One last simple alternative: Total R²**
 - Generate model-predicted y's from fixed effects only (NOT including random effects) and correlate with observed y's
 - Then square correlation → total R²
 - Total R² = total reduction in overall variance of y across levels
 - Can be "unfair" in models with large unexplained sources of variance
- **MORAL OF THE STORY:** Specify EXACTLY which kind of pseudo-R² you used—give the formula and the reference!!

Time-Invariant Predictors in Longitudinal Models

- Topics:
 - Missing predictors in MLM
 - Effects of time-invariant predictors
 - Fixed, systematically varying, and random level-1 effects
 - **Model building strategies and assessing significance**

Model-Building Strategies

- It may be helpful to examine predictor effects in separate models at first, including interactions with all growth terms to see the total pattern of effects for a single predictor
 - Question: Does age matter at all in predicting change over time?
 - e.g., random quadratic model + age, age*time, age*time²
- Then predictor effects can be combined in layers in order to examine unique contributions (and interactions) of each
 - Question: Does age *still* matter after considering reasoning?
 - random quadratic + age, age*time, age*time²,
+ reason, reason*time, reason*time²
 - Potentially also + age*reason, age*reason*time, age*reason*time²
- Sequence of predictors should be guided by theory and research questions—there may not be a single “best model”
 - One person’s “control” is another person’s “question”, so may not end up in the same place given different orders of predictor inclusion

Evaluating Statistical Significance of Multiple New Fixed Effects at Once

- Compare nested models with ML $-2\Delta LL$ test
- Useful for 'borderline' cases - example:
 - Ed*time² interaction at $p = .04$, with nonsignificant ed*time and ed*Intercept (main effect of ed) terms?
 - Is it worth keeping a marginal higher-order interaction that requires two (possibly non-significant) lower-order terms?
 - ML $-2\Delta LL$ test on $df=3$: $-2\Delta LL$ must be > 7.82
 - **REML is WRONG for $-2\Delta LL$ tests for models with different fixed effects, regardless of nested or non-nested**
 - Because of this, it may be more convenient to switch to ML when focusing on modeling fixed effects of predictors
- Compare non-nested models with ML AIC & BIC instead

Evaluating Statistical Significance of New Individual Fixed Effects

Fixed effects can be tested via **Wald** tests: the ratio of its estimate/SE forms a statistic we compare to a distribution

| | Denominator DF is assumed infinite | Denominator DF is estimated instead |
|-------------------------------|--|--|
| Numerator DF = 1 | use z distribution (Mplus, STATA) | use t distribution (SAS, SPSS) |
| Numerator DF > 1 | use χ^2 distribution (Mplus, STATA) | use F distribution (SAS, SPSS) |
| Denominator DF (DDFM) options | not applicable, so DDF is not given | SAS: BW and KR SAS and SPSS: Satterthwaite |

Denominator DF (DDF) Methods

- **Between-Within** (DDFM=BW in SAS, not in SPSS):
 - Total DDF (T) comes from total number of observations, separated into level-2 for N persons and level-1 for n occasions
 - **Level-2 DDF** = $N - \text{\#level-2 fixed effects}$
 - **Level-1 DDF** = Total DDF – Level-2 DDF – #level-1 fixed effects
 - Level-1 effects with random slopes still get level-1 DDF
- **Satterthwaite** (DDFM=Satterthwaite in SAS, default in SPSS):
 - More complicated, but analogous to two-group t -test given unequal residual variances and unequal group sizes
 - Incorporates contribution of variance components at each level
 - Level-2 DDF will resemble Level-2 DDF from BW
 - Level-1 DDF will resemble Level-1 DDF from BW if the level-1 effect is not random, but will resemble level-2 DDF if it is random

Denominator DF (DDF) Methods

- **Kenward-Roger** (DDFM=KR in SAS, not in SPSS):
 - Adjusts the sampling covariance matrix of the fixed effects and variance components to reflect the uncertainty introduced by using large-sample techniques of ML/REML in small N samples
 - This creates different (larger) SEs for the fixed effects
 - Then uses Satterthwaite DDF, new SEs, and t to get p -values
- In an unstructured variance model, all effects use level-2 DDF
- Differences in inference not likely to matter often in practice
 - e.g., critical t -value at DDF=20 is 2.086, at infinite DDF is 1.960
- When in doubt, use KR (is overkill at worst, becomes Satterthwaite)
 - I used Satterthwaite in the book to maintain comparability across programs

Wrapping Up...

- MLM uses ONLY rows of data that are COMPLETE: both predictors AND outcomes must be there!
 - Using whatever data you do have for each person will likely lead to better inferences and more statistical power than using only complete persons (listwise deletion)
- Time-invariant predictors modify the level-1 created growth curve → predict individual intercepts and slopes
 - They account for random effect variances (the predictors are the reasons WHY people need their own intercepts and slopes)
 - If a level-1 effect is not random, it can still be moderated by a cross-level interaction with a time-invariant predictor...
 - ... but then it will predict L1 residual variance instead

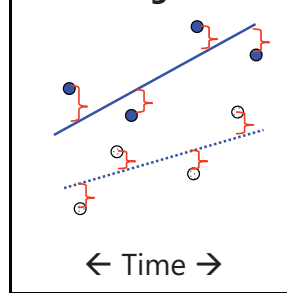
Time-Varying Predictors in Longitudinal Models

- Topics:
 - **Time-varying predictors that fluctuate over time**
 - Person-Mean-Centering (PMC)
 - Grand-Mean-Centering (GMC)
 - Model extensions under Person-MC vs. Grand-MC
 - Time-varying predictors that change over time

The Joy of Time-Varying Predictors

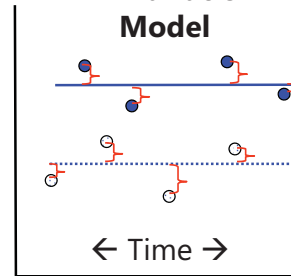
- TV predictors predict leftover **WP (residual) variation**:

WP Change Model



If model for time works, then residuals should look like this →

WP Variation Model



- Modeling time-varying predictors is complicated because they represent an **aggregated effect**:
 - Effect of the *between-person* variation in the predictor x_{ti} on Y
 - Effect of the *within-person* variation in the predictor x_{ti} on Y
 - Here we are assuming the predictor x_{ti} only **fluctuates** over time...
 - We will need a different model if x_{ti} changes systematically over time...

The Joy of Time-Varying Predictors

- Time-varying (TV) predictors usually carry 2 kinds of effects because they are really 2 predictor variables, not 1
- Example: Stress measured daily
 - Some days are worse than others:
 - **WP variation in stress** (represented as deviation from own mean)
 - Some people just have more stress than others all the time:
 - **BP variation in stress** (represented as person mean predictor over time)
- Can quantify each source of variation with an ICC
 - $ICC = (BP \text{ variance}) / (BP \text{ variance} + WP \text{ variance})$
 - $ICC > 0$? TV predictor has BP variation (so it *could* have a BP effect)
 - $ICC < 1$? TV predictor has WP variation (so it *could* have a WP effect)

Between-Person vs. Within-Person Effects

- Between-person and within-person effects in SAME direction
 - Stress → Health?
 - **BP:** People with more chronic stress than other people may have worse general health than people with less chronic stress
 - **WP:** People may feel worse than usual when they are currently under more stress than usual (regardless of what “usual” is)
- Between-person and within-person effects in OPPOSITE directions
 - Exercise → Blood pressure?
 - **BP:** People who exercise more often generally have lower blood pressure than people who are more sedentary
 - **WP:** During exercise, blood pressure is higher than during rest
- Variables have different **meanings** at different levels!
- Variables have different **scales** at different levels

3 Kinds of Effects for TV Predictors

- **Is the Between-Person (BP) effect significant?**
 - Are people with higher predictor values than other people (on average over time) also higher on Y than other people (on average over time), such that the person mean of the TV predictor accounts for level-2 random intercept variance (τ_{U0}^2)?
- **Is the Within-Person (WP) effect significant?**
 - If you have higher predictor values than usual (at this occasion), do you also have higher outcomes values than usual (at this occasion), such that the within-person deviation of the TV predictor accounts for level-1 residual variance (σ_e^2)?
- **Are the BP and WP effects different sizes: Is there a contextual effect?**
 - After controlling for the absolute value of TV predictor at each occasion, is there still an incremental contribution from having a higher person mean of the TV predictor (i.e., does one's general tendency predict τ_{U0}^2 above and beyond)?
 - If there is no contextual effect, then the BP and WP effects of the TV predictor show convergence, such that their effects are of equivalent magnitude

Modeling TV Predictors (labeled as x_{ti})

- **Level-2 effect of x_{ti} :**

- The level-2 effect of x_{ti} is usually represented by the person's mean of time-varying x_{ti} across time (labeled as **PMx_i** or \bar{X}_i)
- PMx_i should be centered at a CONSTANT (grand mean or other) so that 0 is meaningful, just like any other time-invariant predictor

- **Level-1 effect of x_{ti} can be included two different ways:**

- **"Group-mean-centering"** → **"person-mean-centering"** in longitudinal, in which level-1 predictors are centered using a level-2 VARIABLE
- **"Grand-mean-centering"** → level-1 predictors are centered using a CONSTANT (not necessarily the grand mean; it's just called that)
- Note that these 2 choices do NOT apply to the level-2 effect of x_{ti} !
 - But the interpretation of the level-2 effect of x_{ti} WILL DIFFER based on which centering method you choose for the level-1 effect of x_{ti} !

Time-Varying Predictors in Longitudinal Models

- Topics:

- Time-varying predictors that fluctuate over time
- **Person-Mean-Centering (PMC)**
- Grand-Mean-Centering (GMC)
- Model extensions under Person-MC vs. Grand-MC
- Time-varying predictors that change over time

Person-Mean-Centering (P-MC)

- In P-MC, we decompose the TV predictor x_{ti} into 2 variables that directly represent its BP (level-2) and WP (level-1) sources of variation, and include those variables as the predictors instead:
- **Level-2, PM predictor = person mean of x_{ti}**
 - $PMx_i = \bar{X}_i - C$
 - PMx_i is centered at a constant C , chosen so 0 is meaningful
 - PMx_i is positive? Above sample mean → "more than other people"
 - PMx_i is negative? Below sample mean → "less than other people"
- **Level-1, WP predictor = deviation from person mean of x_{ti}**
 - $WPx_{ti} = x_{ti} - \bar{X}_i$ (note: uncentered person mean \bar{X}_i is used to center x_{ti})
 - WPx_{ti} is NOT centered at a constant; is centered at a VARIABLE
 - WPx_{ti} is positive? Above your own mean → "more than usual"
 - WPx_{ti} is negative? Below your own mean → "less than usual"

Within-Person Fluctuation Model with Person-Mean-Centered Level-1 x_{ti}

→ WP and BP Effects directly through separate parameters

x_{ti} is person-mean-centered into WPx_{ti} , with PMx_i at L2:

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(WPx_{ti}) + e_{ti}$

$WPx_{ti} = x_{ti} - \bar{X}_i \rightarrow$ it has only Level-1 WP variation

Level 2: $\beta_{0i} = Y_{00} + Y_{01}(PMx_i) + U_{0i}$

$\beta_{1i} = Y_{10}$

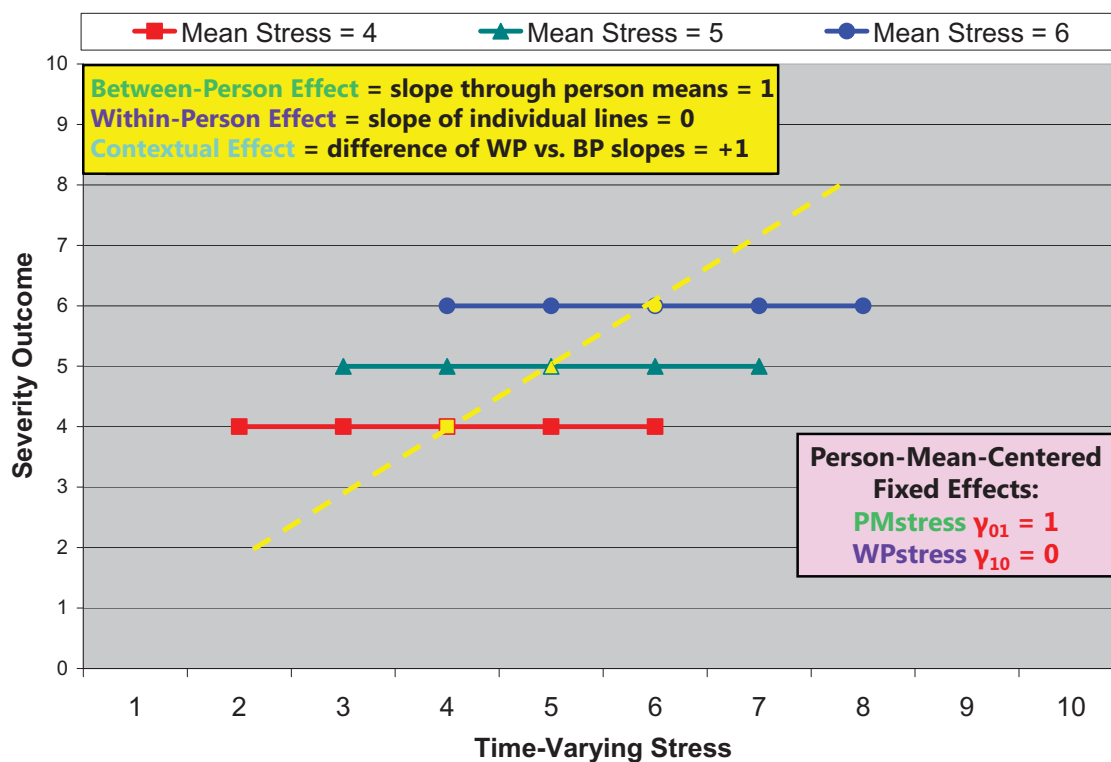
$PMx_i = \bar{X}_i - C \rightarrow$ it has only Level-2 BP variation

Y_{10} = WP main effect of having more x_{ti} than usual

Y_{01} = BP main effect of having more \bar{X}_i than other people

Because WPx_{ti} and PMx_i are uncorrelated, each gets the total effect for its level (WP=L1, BP=L2)

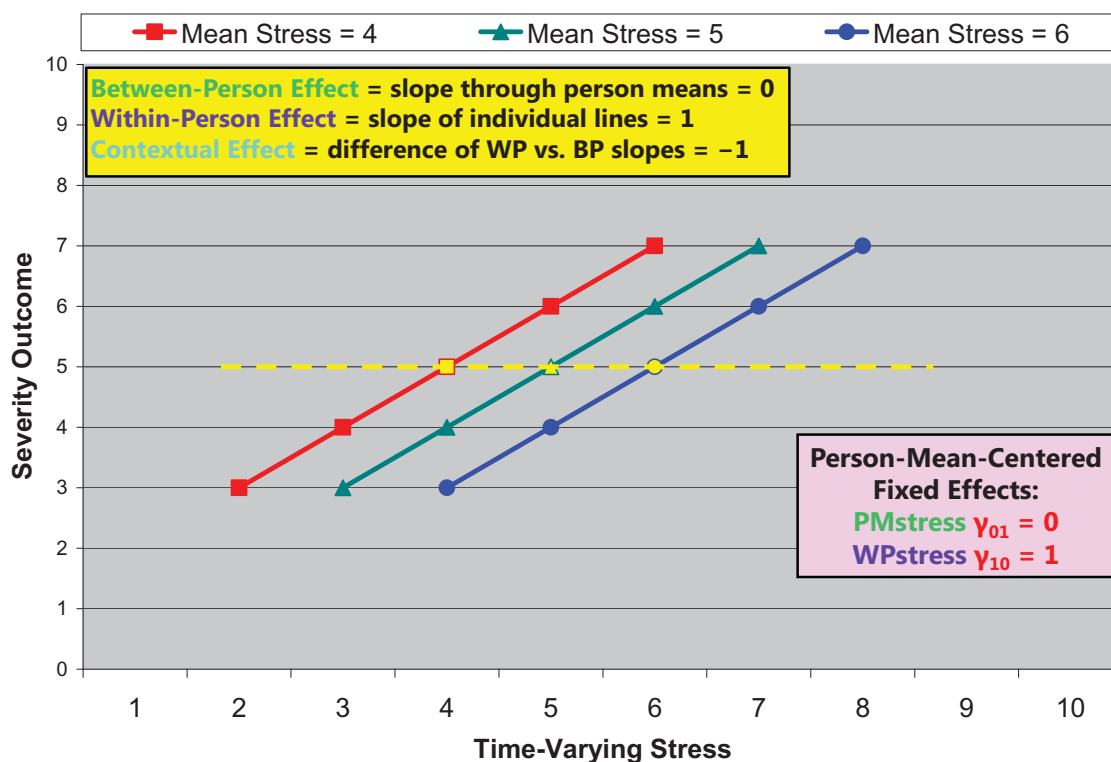
ALL Between-Person Effect, NO Within-Person Effect



Lecture 4

10

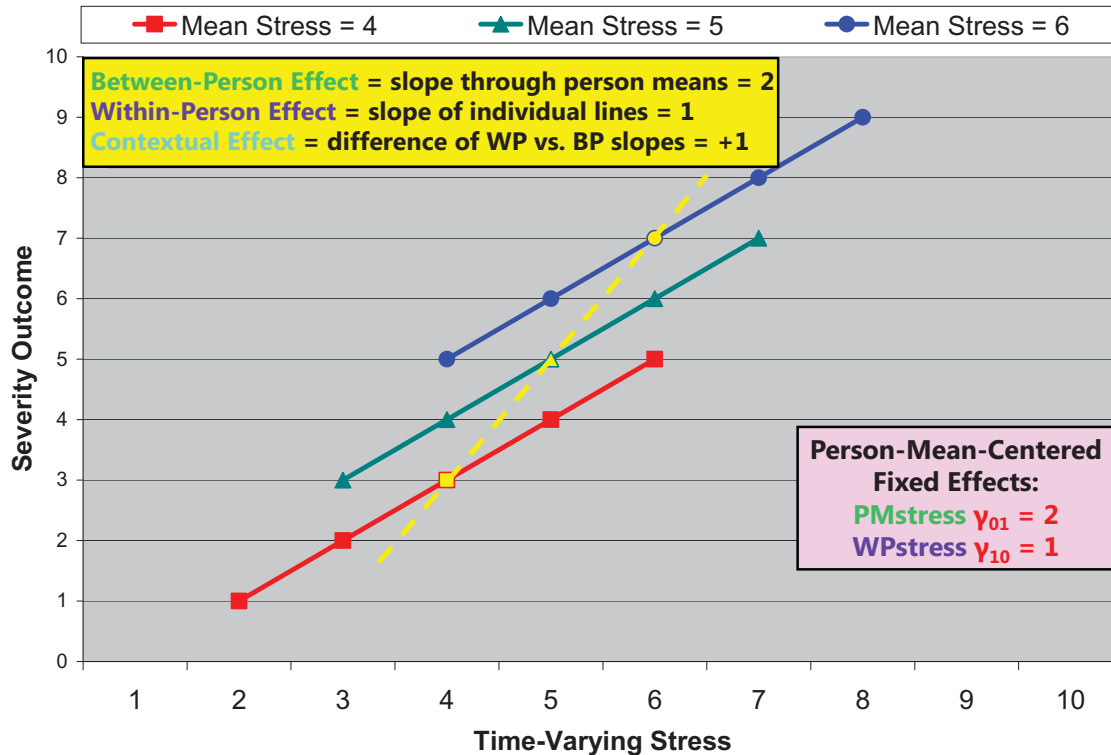
NO Between-Person Effect, ALL Within-Person Effect



Lecture 4

11

Between-Person Effect > Within-Person Effect



Lecture 4

12

Within-Person Fluctuation Model with Person-Mean-Centered Level-1 x_{ti}

→ WP and BP Effects directly through separate parameters

x_{ti} is person-mean-centered into WPx_{ti} , with PMx_i at L2:

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(WPx_{ti}) + e_{ti}$

$WPx_{ti} = x_{ti} - \bar{X}_i \rightarrow$ it has only Level-1 WP variation

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

$PMx_i = \bar{X}_i - C \rightarrow$ it has only Level-2 BP variation

$\beta_{1i} = \gamma_{10} + \gamma_{11}(PMx_i) + U_{1i}$

U_{1i} is a random slope for the WP effect of x_{ti}

γ_{10} = WP simple main effect of having more x_{ti} than usual for $PMx_i = 0$

γ_{01} = BP simple main effect of having more \bar{X}_i than other people for people at their own mean ($WPx_{ti} = x_{ti} - \bar{X}_i \rightarrow 0$)

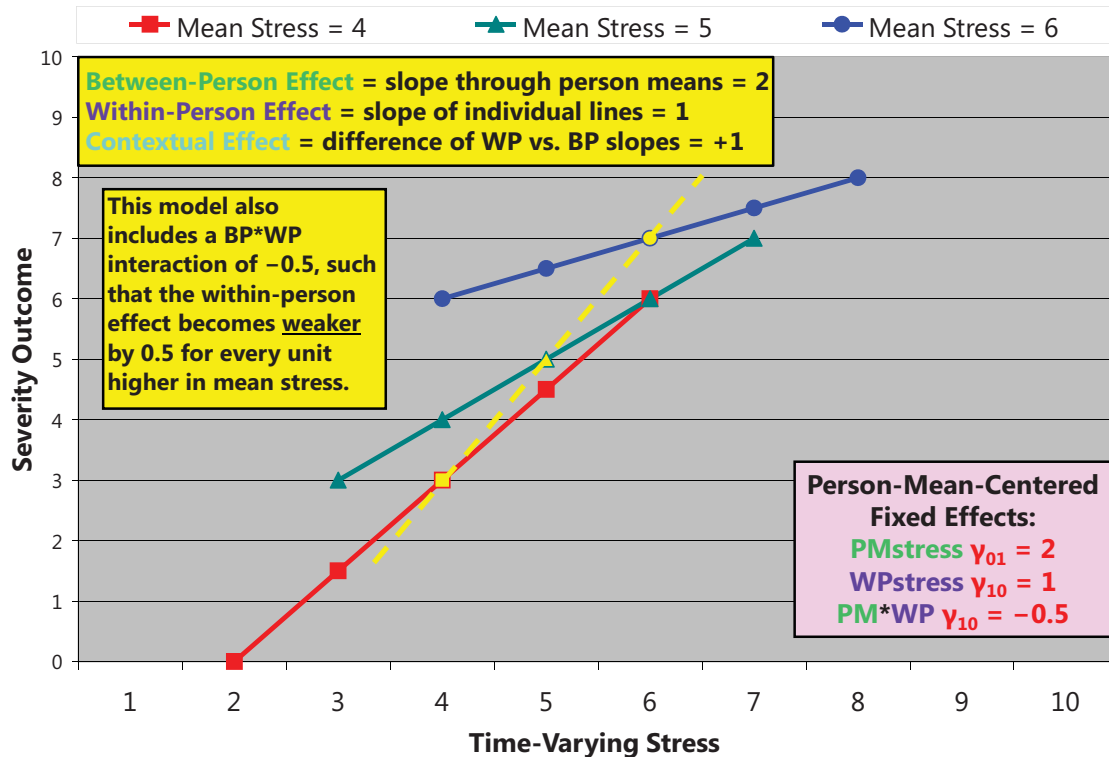
γ_{11} = BP*WP interaction: how the effect of having more x_{ti} than usual differs by how much \bar{X}_i you have

Note: this model should also test γ_{02} for $PMx_i * PMx_i$ (stay tuned)

Lecture 4

13

Between-Person x Within-Person Interaction



Lecture 4

14

Time-Varying Predictors in Longitudinal Models

- Topics:
 - Time-varying predictors that fluctuate over time
 - Person-Mean-Centering (PMC)
 - **Grand-Mean-Centering (GMC)**
 - Model extensions under Person-MC vs. Grand-MC
 - Time-varying predictors that change over time

Lecture 4

15

3 Kinds of Effects for TV Predictors

- **What Person-Mean-Centering tells us directly:**
- **Is the Between-Person (BP) effect significant?**
 - Are people with higher predictor values than other people (*on average over time*) also higher on Y than other people (*on average over time*), such that the person mean of the TV predictor accounts for level-2 random intercept variance (τ_{U0}^2)?
 - This would be indicated by a significant fixed effect of **PMx_i**
 - Note: this is NOT controlling for the absolute value of x_{ti} at each occasion
- **Is the Within-Person (WP) effect significant?**
 - If you have higher predictor values than usual (*at this occasion*), do you also have higher outcomes values than usual (*at this occasion*), such that the within-person deviation of the TV predictor accounts for level-1 residual variance (σ_e^2)?
 - This would be indicated by a significant fixed effect of **WPx_{ti}**
 - Note: this is represented by the relative value of x_{ti} , NOT the absolute value of x_{ti}

3 Kinds of Effects for TV Predictors

- **What Person-Mean-Centering DOES NOT tell us directly:**
- **Are the **BP** and **WP** effects different sizes: Is there a **contextual** effect?**
 - After controlling for the absolute value of the TV predictor at each occasion, is there still an incremental contribution from having a higher person mean of the TV predictor (i.e., does one's general tendency predict τ_{U0}^2 above and beyond just the time-specific value of the predictor)?
 - If there is no contextual effect, then the BP and WP effects of the TV predictor show **convergence**, such that their effects are of equivalent magnitude
- **To answer this question about the **contextual effect for the incremental contribution of the person mean**, we have two options:**
 - Ask for the contextual effect via an ESTIMATE statement in SAS (or TEST in SPSS, or NEW in Mplus, or LINCOM in STATA): **WPx_{ti} -1 PMx_i 1**
 - Use **"grand-mean-centering"** for time-varying x_{ti} instead: **TVx_{ti} = $x_{ti} - C$**
→ **centered at a CONSTANT, NOT A LEVEL-2 VARIABLE**
 - Which constant only matters for what the reference point is; it could be the grand mean or other

Remember Regular Old Regression?

- In this model: $Y_i = \beta_0 + \beta_1(X_{1i}) + \beta_2(X_{2i}) + e_i$
 - If X_{1i} and X_{2i} **ARE NOT** correlated:
 - β_1 is **ALL the relationship** between X_{1i} and Y_i
 - β_2 is **ALL the relationship** between X_{2i} and Y_i
 - If X_{1i} and X_{2i} **ARE** correlated:
 - β_1 is **different than** the full relationship between X_{1i} and Y_i
 - “Unique” effect of X_{1i} *controlling for* X_{2i} or *holding* X_{2i} *constant*
 - β_2 is **different than** the full relationship between X_{2i} and Y_i
 - “Unique” effect of X_{2i} *controlling for* X_{1i} or *holding* X_{1i} *constant*
 - Hang onto that idea...

Person-MC vs. Grand-MC for Time-Varying Predictors

| Level 2 | | Original | Person-MC Level 1 | Grand-MC Level 1 |
|-------------|-------------------------|----------|---------------------------------|-------------------------|
| \bar{X}_i | $PMx_i = \bar{X}_i - 5$ | x_{ti} | $WPx_{ti} = x_{ti} - \bar{X}_i$ | $TVx_{ti} = x_{ti} - 5$ |
| 3 | -2 | 2 | -1 | -3 |
| 3 | -2 | 4 | 1 | -1 |
| 7 | 2 | 6 | -1 | 1 |
| 7 | 2 | 8 | 1 | 3 |

Same PMx_i goes into the model using either way of centering the level-1 variable x_{ti}

Using **Person-MC**, WPx_{ti} has NO level-2 BP variation, so it is not correlated with PMx_i

Using **Grand-MC**, TVx_{ti} STILL has level-2 BP variation, so it is STILL CORRELATED with PMx_i

So the effects of PMx_i and TVx_{ti} when included together under Grand-MC will be different than their effects would be if they were by themselves...

Within-Person Fluctuation Model with x_{ti} represented at Level 1 Only:

→ WP and BP Effects are **Smushed Together**

x_{ti} is grand-mean-centered into TVx_{ti} , **WITHOUT** PMx_i at L2:

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

$TVx_{ti} = x_{ti} - C \rightarrow$ it still has both Level-2 BP and Level-1 WP variation

Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$

$\beta_{1i} = \gamma_{10}$

γ_{10} = *smushed* WP and BP effects

Because TVx_{ti} still contains its original 2 different kinds of variation (BP and WP), its 1 fixed effect has to do the work of 2 predictors!

A *smushed* effect is also referred to as the *convergence, conflated, or composite effect*

Convergence (Smushed) Effect of a Time-Varying Predictor

$$\text{Convergence Effect: } \gamma_{\text{conv}} \approx \frac{\frac{\gamma_{BP}}{SE_{BP}^2} + \frac{\gamma_{WP}}{SE_{WP}^2}}{\frac{1}{SE_{BP}^2} + \frac{1}{SE_{WP}^2}}$$

Adapted from Raudenbush & Bryk (2002, p. 138)

- **The convergence effect will often be closer to the within-person effect** (due to larger level-1 sample size and thus smaller SE)
- **It is the rule, not the exception, that between and within effects differ** (Snijders & Bosker, 1999, p. 52-56, and personal experience!)
- However—when grand-mean-centering a time-varying predictor, **convergence is testable** by including a **contextual effect (carried by the person mean)** for how the **BP effect** differs from the **WP effect**...

Within-Person Fluctuation Model with Grand-Mean-Centered Level-1 x_{ti}

→ Model tests difference of WP vs. BP effects (It's been fixed!)

x_{ti} is grand-mean-centered into TVx_{ti} , WITH PMx_i at L2:

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

$TVx_{ti} = x_{ti} - C \rightarrow$ it still has both Level-2 BP and Level-1 WP variation

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

$PMx_i = \bar{x}_i - C \rightarrow$ it has only Level-2 BP variation

$\beta_{1i} = \gamma_{10}$

γ_{10} becomes the WP effect → unique level-1 effect after controlling for PMx_i

γ_{01} becomes the contextual effect that indicates how the BP effect differs from the WP effect → unique level-2 effect after controlling for TVx_{ti} → does usual level matter beyond current level?

Person-MC and Grand-MC Models are Equivalent Given a Fixed Level-1 Main Effect Only

Person-MC: $WPx_{ti} = x_{ti} - PMx_i$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PMx_i) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10}$

→ $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti}$

→ $y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti}$

Composite Model:
← In terms of P-MC
← In terms of G-MC

Grand-MC: $TVx_{ti} = x_{ti}$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti}$

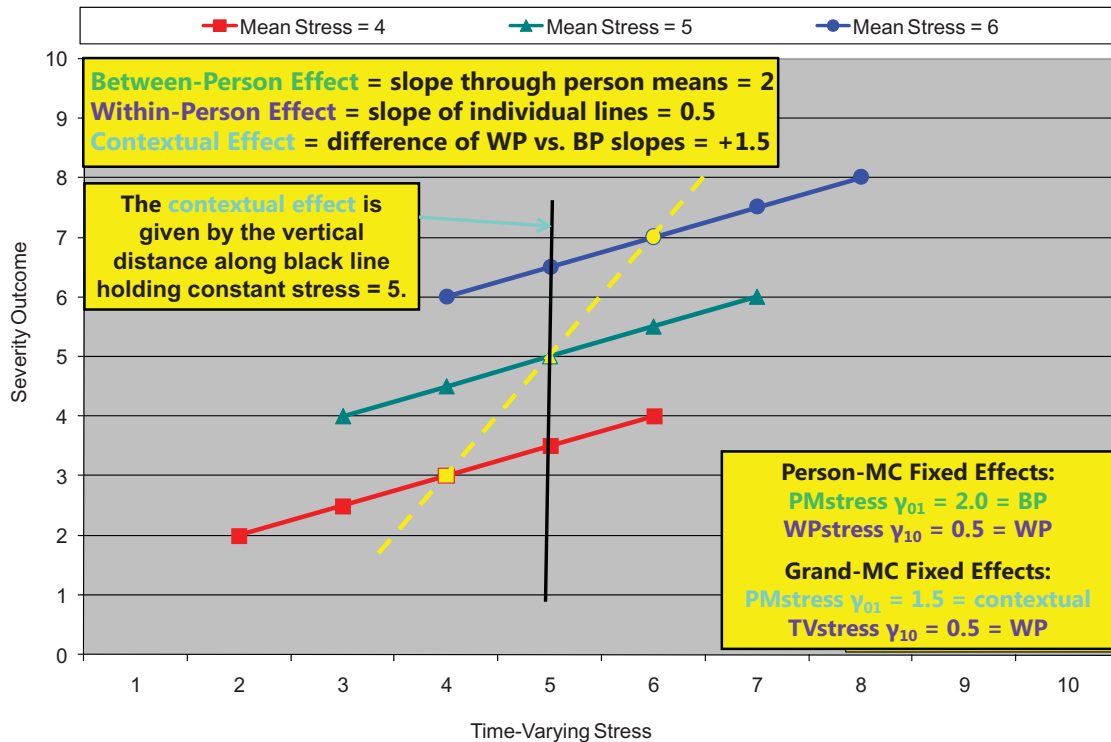
Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10}$

→ $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti}$

| Effect | P-MC | G-MC |
|------------|-----------------------------|-----------------------------|
| Intercept | γ_{00} | γ_{00} |
| WP Effect | γ_{10} | γ_{10} |
| Contextual | $\gamma_{01} - \gamma_{10}$ | γ_{01} |
| BP Effect | γ_{01} | $\gamma_{01} + \gamma_{10}$ |

P-MC vs. G-MC: Interpretation Example



Lecture 4

24

Summary: 3 Effects for TV Predictors

- **Is the Between-Person (BP) effect significant?**
 - Are people with higher predictor values than other people (on average over time) also higher on Y than other people (on average over time), such that the person mean of the TV predictor accounts for level-2 random intercept variance ($\tau_{U_0}^2$)?
 - Given directly by level-2 effect of PMx_i if using Person-MC for the level-1 predictor (or can be requested via ESTIMATE if using Grand-MC for the level-1 predictor)
- **Is the Within-Person (WP) effect significant?**
 - If you have higher predictor values than usual (at this occasion), do you also have higher outcomes values than usual (at this occasion), such that the within-person deviation of the TV predictor accounts for level-1 residual variance (σ_e^2)?
 - Given directly by the level-1 effect of WPx_{ti} if using Person-MC —OR— given directly by the level-1 effect of TVx_{ti} if using Grand-MC and including PMx_i at level 2 (without PMx_i , the level-1 effect of TVx_{ti} if using Grand-MC is the smushed effect)
- **Are the BP and WP Effects different sizes: Is there a contextual effect?**
 - After controlling for the absolute value of TV predictor value at each occasion, is there still an incremental contribution from having a higher person mean of the TV predictor (i.e., does one's general tendency predict $\tau_{U_0}^2$ above and beyond)?
 - Given directly by level-2 effect of PMx_i if using Grand-MC for the level-1 predictor (or can be requested via ESTIMATE if using Person-MC for the level-1 predictor)

Lecture 4

25

Variance Accounted For By Level-1 Predictors

- **Fixed effects of level 1 predictors *by themselves*:**
 - Level-1 (WP) main effects reduce Level-1 (WP) residual variance
 - Level-1 (WP) interactions also reduce Level-1 (WP) residual variance
- **What happens at level 2 depends on what kind of variance the level-1 predictor has:**
 - If the level-1 predictor ALSO has level-2 variance (e.g., Grand-MC predictors), then its level-2 variance will also likely reduce level-2 random intercept variance
 - If the level-1 predictor DOES NOT have level-2 variance (e.g., Person-MC predictors), then its reduction in the level-1 residual variance will cause an INCREASE in level-2 random intercept variance
 - Same thing happens with Grand-MC level-1 predictors, but you don't generally see it
 - It's just an artifact that the estimate of true random intercept variance is:
$$\text{True } \tau_{U_0}^2 = \text{observed } \tau_{U_0}^2 - \frac{\sigma_e^2}{n} \rightarrow \text{so if only } \sigma_e^2 \text{ decreases, } \tau_{U_0}^2 \text{ increases}$$

Time-Varying Predictors in Longitudinal Models

- Topics:
 - Time-varying predictors that fluctuate over time
 - Person-Mean-Centering (PMC)
 - Grand-Mean-Centering (GMC)
 - **Model extensions under Person-MC vs. Grand-MC**
 - Time-varying predictors that change over time

The Joy of Interactions Involving Time-Varying Predictors

- Must consider interactions with both its BP and WP parts:
- Example: Does time-varying stress (x_{ti}) interact with sex (Sex_i)?
- Person-Mean-Centering:
 - $WPx_{ti} * Sex_i \rightarrow$ Does the WP stress effect differ between men and women?
 - $PMx_i * Sex_i \rightarrow$ Does the BP stress effect differ between men and women?
 - Not controlling for current levels of stress
 - If forgotten, then Sex_i moderates the stress effect only at level 1 (WP, not BP)
- Grand-Mean-Centering:
 - $TVx_{ti} * Sex_i \rightarrow$ Does the WP stress effect differ between men and women?
 - $PMx_i * Sex_i \rightarrow$ Does the *contextual* stress effect differ b/t men and women?
 - Incremental BP stress effect *after controlling for current levels of stress*
 - If forgotten, then although the level-1 main effect of stress has been un-smushed via the main effect of PMx_i , the interaction of $TVx_{ti} * Sex_i$ would still be smushed

Interactions with Time-Varying Predictors: Example: TV Stress (x_{ti}) by Gender (Sex_i)

Person-MC: $WPx_{ti} = x_{ti} - PMx_i$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PMx_i) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(Sex_i)$

Composite: $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti} + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti} - PMx_i)$

Grand-MC: $TVx_{ti} = x_{ti}$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(Sex_i)$

Composite: $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti})$

Interactions Involving Time-Varying Predictors Belong at Both Levels of the Model

On the left below → Person-MC: $WP_{x_{ti}} = x_{ti} - PM_{x_i}$

$$y_{ti} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + \gamma_{10}(x_{ti} - PM_{x_i}) + U_{0i} + e_{ti} \\ + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PM_{x_i}) + \gamma_{11}(Sex_i)(x_{ti} - PM_{x_i})$$

← Composite model written as Person-MC

$$y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PM_{x_i}) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \\ + \gamma_{02}(Sex_i) + (\gamma_{03} - \gamma_{11})(Sex_i)(PM_{x_i}) + \gamma_{11}(Sex_i)(x_{ti})$$

← Composite model written as Grand-MC

On the right below → Grand-MC: $TV_{x_{ti}} = x_{ti}$

$$y_{ti} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \\ + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PM_{x_i}) + \gamma_{11}(Sex_i)(x_{ti})$$

After adding an interaction for Sex_i with stress at both levels, then the Person-MC and Grand-MC models are equivalent

Intercept: $\gamma_{00} = \gamma_{00}$

BP Effect: $\gamma_{01} = \gamma_{01} + \gamma_{10}$

Contextual: $\gamma_{01} = \gamma_{01} - \gamma_{10}$

WP Effect: $\gamma_{10} = \gamma_{10}$

BP*Sex Effect: $\gamma_{03} = \gamma_{03} + \gamma_{11}$

Contextual*Sex: $\gamma_{03} = \gamma_{03} - \gamma_{11}$

Sex Effect: $\gamma_{20} = \gamma_{20}$

BP*WP or Contextual*WP is the same: $\gamma_{11} = \gamma_{11}$

Intra-variable Interactions

- Still must consider interactions with both its BP and WP parts!
- Example: Interaction of TV stress (x_{ti}) with person mean stress (PM_{x_i})
- Person-Mean-Centering:
 - $WP_{x_{ti}} * PM_{x_i} \rightarrow$ Does the WP stress effect differ by overall stress level?
 - $PM_{x_i} * PM_{x_i} \rightarrow$ Does the BP stress effect differ by overall stress level?
 - Not controlling for current levels of stress
 - If forgotten, then PM_{x_i} moderates the stress effect only at level 1 (WP, not BP)
- Grand-Mean-Centering:
 - $TV_{x_{ti}} * PM_{x_i} \rightarrow$ Does the WP stress effect differ by overall stress level?
 - $PM_{x_i} * PM_{x_i} \rightarrow$ Does the *contextual* stress effect differ by overall stress?
 - Incremental BP stress effect *after controlling for current levels of stress*
 - If forgotten, then although the level-1 main effect of stress has been un-smushed via the main effect of $TV_{x_{ti}}$, the interaction of $TV_{x_{ti}} * PM_{x_i}$ would still be smushed

Intra-variable Interactions: Example: TV Stress (x_{ti}) by Person Mean Stress (PMx_i)

Person-MC: $WPx_{ti} = x_{ti} - PMx_i$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PMx_i) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(PMx_i)(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(PMx_i)$

Composite: $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti}$
 $+ \gamma_{02}(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti} - PMx_i)$

Grand-MC: $TVx_{ti} = x_{ti}$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(PMx_i)(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(PMx_i)$

Composite: $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti}$
 $+ \gamma_{02}(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti})$

Intra-variable Interactions: Example: TV Stress (x_{ti}) by Person Mean Stress (PMx_i)

On the left below → Person-MC: $WPx_{ti} = x_{ti} - PMx_i$

$y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti}$
 $+ \gamma_{02}(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti} - PMx_i)$

$y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti}$
 $+ (\gamma_{02} - \gamma_{11})(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti})$

← Written as
Person-MC

← Written as
Grand-MC

On the right below → Grand-MC: $TVx_{ti} = x_{ti}$

$y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti}$
 $+ \gamma_{02}(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti})$

After adding an interaction for PMx_i with stress at both levels, then the Person-MC and Grand-MC models are equivalent

Intercept: $\gamma_{00} = \gamma_{00}$

BP Effect: $\gamma_{01} = \gamma_{01} + \gamma_{10}$

Contextual: $\gamma_{01} = \gamma_{01} - \gamma_{10}$

WP Effect: $\gamma_{10} = \gamma_{10}$

BP² Effect: $\gamma_{02} = \gamma_{02} + \gamma_{11}$

Contextual²: $\gamma_{02} = \gamma_{02} - \gamma_{11}$

BP*WP or Contextual*WP is the same: $\gamma_{11} = \gamma_{11}$

When Person-MC \neq Grand-MC: Random Effects of TV Predictors

Person-MC: $WP_{x_{ti}} = x_{ti} - PM_{x_i}$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PM_{x_i}) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + U_{0i}$

$\beta_{1i} = \gamma_{10} + U_{1i}$

Variance due to PM_{x_i} is removed from the random slope in Person-MC.

$\rightarrow y_{ti} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + \gamma_{10}(x_{ti} - PM_{x_i}) + U_{0i} + U_{1i}(x_{ti} - PM_{x_i}) + e_{ti}$

Grand-MC: $TV_{x_{ti}} = x_{ti}$

Level-1: $y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + U_{0i}$

$\beta_{1i} = \gamma_{10} + U_{1i}$

Variance due to PM_{x_i} is still part of the random slope in Grand-MC. So these models cannot be made equivalent.

$\rightarrow y_{ti} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + \gamma_{10}(x_{ti}) + U_{0i} + U_{1i}(x_{ti}) + e_{ti}$

Random Effects of TV Predictors

- **Random intercepts** mean different things under each model:
 - **Person-MC** \rightarrow Individual differences at $WP_{x_{ti}} = 0$ (that everyone has)
 - **Grand-MC** \rightarrow Individual differences at $TV_{x_{ti}} = 0$ (that not everyone has)
- **Differential shrinkage of the random intercepts** results from differential reliability of the intercept data across models:
 - Person-MC \rightarrow Won't affect shrinkage of slopes unless highly correlated
 - Grand-MC \rightarrow Will affect shrinkage of slopes due to forced extrapolation
- As a result, the **random slope variance may be too small** when using Grand-MC rather than Person-MC
 - Problem worsens with greater ICC of TV Predictor (more extrapolation)
 - Anecdotal example using clustered data was presented in Raudenbush & Bryk (2002; chapter 6)

Modeling Time-Varying Categorical Predictors

- Person-MC and Grand-MC really only apply to *continuous* TV predictors, but the need to consider BP and WP effects applies to *categorical* TV predictors too
- Binary level-1 predictors do not lend themselves to Person-MC
 - e.g., $x_{ti} = 0$ or 1 per occasion, person mean = $.50$ across occasions → impossible values
 - If $x_{ti} = 0$, then $WPx_{ti} = 0 - .50 = -0.50$; If $x_{ti} = 1$, then $WPx_{ti} = 1 - .50 = 0.50$
 - Better: Leave x_{ti} uncentered and include person mean as level-2 predictor (results ~ Grand-MC)
- For >2 categories, person means of multiple dummy codes starts to break down, but we can think about types of people, and code BP effects accordingly
- Example: Dementia present/not at each time point?
 - **BP effects** → Ever diagnosed with dementia (no, yes)?
 - People who will eventually be diagnosed may differ prior to diagnosis (a BP effect)
 - **TV effect** → Diagnosed with dementia at each time point (no, yes)?
 - Acute differences of before/after diagnosis logically can only exist in the “ever” people
- Other examples: Mentor status, father absence, type of shift work (AM/PM)

Wrapping Up: Person-MC vs. Grand-MC

- Time-varying predictors carry at least two potential effects:
 - Some people are higher/lower than other people → BP, level-2 effect
 - Some occasions are higher/lower than usual → WP, level-1 effect
- BP and WP effects almost always need to be represented by two or more model parameters, using either:
 - *Person-mean-centering* (WPx_{ti} and PMx_i): $WP \neq 0?$, $BP \neq 0?$
 - *Grand-mean-centering* (TVx_{ti} and PMx_i): $WP \neq 0?$, $BP \neq WP?$
 - Both yield equivalent models if the level-1 WP effect is fixed, but not if the level-1 WP effect is random
 - Grand MC → *absolute* effect of x_{ti} varies randomly over people
 - Person MC → *relative* effect of x_{ti} varies randomly over people
 - Use prior theory and empirical data (ML AIC, BIC) to decide

Time-Varying Predictors in Longitudinal Models

- Topics:
 - Time-varying predictors that fluctuate over time
 - Person-Mean-Centering (PMC)
 - Grand-Mean-Centering (GMC)
 - Model extensions under Person-MC vs. Grand-MC
 - **Time-varying predictors that change over time**

Baseline Centering for Time-Varying Predictors that Change over Time

- Although using the person mean of the time-varying predictor at level-2 (PMx_i) is the most common way to represent the effect of between-person differences, there are other options that sometimes can be more useful
- **Level-2 \rightarrow X at centering point of time (e.g., x_{ti} at time 0)**
 - Useful if x_{ti} at specific time point conveys useful information, such as baseline level of a covariate in an intervention
 - Useful if x_{ti} is expected to change systematically over time, too
- Create predictors using a variant of PMC \rightarrow **baseline centering**:
 - Level 1 = $stress_{ti} - stressTime0_i \rightarrow$ longitudinal effect
 - L1 represents *change from baseline*, not deviation from own mean
 - Level 2 = $stressTime0_i - C \rightarrow$ cross-sectional effect
 - L2 represents effect of *baseline level*, not effect of mean level averaged over time

Baseline Centering: Caveats

- In using baseline centering instead of person-mean-centering, a complete separation of the BP and WP variance in the time-varying predictor is not obtained:
 - If the time-varying predictor shows change, you are not fitting a model for that change—no separation of true change from error
 - The level-1 predictor for “WP change in X” is both individual differences in change (U_{1i}) and residual deviations from change (e_{ti}), which should each really have their own relationship to the outcome
 - Therefore, there may be systematic BP differences with regard to the individual slope still contained in the WP change in X predictor (which may be related to BP differences in level at time 0)
- A better option is to use a multivariate model instead, in which a model for change X is fitted for both X and Y
 - Can examine relationships between intercepts, slopes, and residuals as separate model parameters
 - Can be done in MLM programs, but more flexibility in SEM programs

Multivariate Models via M-SEM

- Person-MC (or baseline centering) is the poor man’s version of a model-based decomposition of BP and WP variance, which is necessary when X is treated as a predictor in MLM programs
- Through Multilevel Structural Equation Modeling (M-SEM), it is possible to fit a model for X along with the model for Y
 - It’s called SEM because random effects = latent variables, but there is no latent variable measurement model as in traditional uses of SEM
 - Person mean = random intercept variance, WP deviation = residual variance, but can also include random slopes for change over time in X
 - Can directly assess multilevel mediation through simultaneous analysis
 - Some evidence that level-2 effects are less biased (because person mean is not perfectly reliable), but more imprecise (more parameters to estimate)
- What could go wrong? No REML! Good luck fitting interactions!
 - Those involving level-2 effects are modeled as latent variable interactions
 - This requires numeric integration, a very computationally intense way of getting parameter estimates in ML, which may not be possible in all data

Two-Level Models for Clustered* Data

- Topics:
 - **Fixed vs. random effects for modeling clustered data**
 - **ICC and design effects in clustered data**
 - Group-Mean-Centering vs. Grand-Mean Centering
 - Model extensions under Group-MC and Grand-MC

* *Clustering = Nesting = Grouping...*

MLM for Clustered Data

- So far we've built models to account for dependency created by repeated measures (time within person)
- Now we examine two-level models for more general examples of nesting/clustering/grouping:
 - Students within schools, athletes within teams
 - Siblings within families, partners within dyads
 - Employees within businesses, patients within doctors
- Residuals of people from same group are likely to be correlated due to group differences (e.g., purposeful grouping or shared experiences create dependency)
- **Recurring theme: You still have to care about group-level variation, even if that's not the point of your study**

2 Options for Differences Across Groups

Represent Group Differences as Fixed Effects

- Include (#groups-1) contrasts for group membership in the **model for the means** (via CLASS) → so group is NOT another “level”
- Permits inference about differences between specific groups, but you cannot include between-group predictors (group is saturated)
- Snijders & Bosker (1999) ch. 4, p. 44 recommend if #groups < 10ish

Represent Group Differences as a Random Effect

- Include a **random intercept variance in the model for the variance**, such that group differences become another “level”
- Permits inference about differences across groups more generally, for which you can test effects of between-group predictors
- Better if #groups > 10ish and you want to **predict** group differences

Empty Means, Random Intercept Model

MLM for Clustered Data:

- Change in notation:
 - i = level 1, j = level 2

- Level 1:

$$y_{ij} = \beta_{0j} + e_{ij}$$

- Level 2:

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

Fixed Intercept
= grand mean
(because no
predictors yet)

Random Intercept
= group-specific
deviation from
predicted intercept

Residual = person-specific deviation
from group's predicted outcome

3 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

- Level-1 Variance of $e_{ij} \rightarrow \sigma_e^2$
- Level-2 Variance of $U_{0j} \rightarrow \tau_{U_0}^2$

Composite equation:

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

Matrices in a Random Intercept Model

Total predicted data matrix is called **V matrix**, created from the **G [TYPE=UN]** and **R [TYPE=VC]** matrices as follows:

$$\mathbf{V} = \mathbf{Z} * \mathbf{G} * \mathbf{Z}^T + \mathbf{R} = \mathbf{V}$$

$$\mathbf{V} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \tau_{U_0}^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix} = \begin{bmatrix} \tau_{U_0}^2 + \sigma_e^2 & \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 \\ \tau_{U_0}^2 & \tau_{U_0}^2 + \sigma_e^2 & \tau_{U_0}^2 & \tau_{U_0}^2 \\ \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 + \sigma_e^2 & \tau_{U_0}^2 \\ \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 & \tau_{U_0}^2 + \sigma_e^2 \end{bmatrix}$$

VCORR then provides the intraclass correlation, calculated as:

$$\text{ICC} = \tau_{U_0}^2 / (\tau_{U_0}^2 + \sigma_e^2)$$

$$\begin{bmatrix} 1 & \text{ICC} & \text{ICC} & \text{ICC} \\ \text{ICC} & 1 & \text{ICC} & \text{ICC} \\ \text{ICC} & \text{ICC} & 1 & \text{ICC} \\ \text{ICC} & \text{ICC} & \text{ICC} & 1 \end{bmatrix} \text{ assumes a constant correlation over time}$$

The **G**, **Z**, and **R** matrices still get combined to create the **V** matrix, except that they are now per group. **R** and **V** have n rows by n columns, in which $n = \# \text{ level-1 units}$, which is now people, not time. Thus, no type of **R** matrix other than VC will be used, and REPEATED is not needed.

Intraclass Correlation (ICC)

$$\text{ICC} = \frac{\text{BG}}{\text{BG} + \text{WG}} = \frac{\text{Intercept Variance}}{\text{Intercept Variance} + \text{Residual Variance}}$$

$$= \frac{\tau_{U_0}^2}{\tau_{U_0}^2 + \sigma_e^2}$$

$\tau_{U_0}^2 \rightarrow$ Why don't all groups have the same mean?
 $\sigma_e^2 \rightarrow$ Why don't all people from the same group have the same outcome?

- ICC = Proportion of total variance that is between groups
- ICC = Average correlation among persons from same group
- ICC is a standardized way of expressing how much we need to worry about *dependency due to group mean differences*
(i.e., ICC is an effect size for constant group dependency)
 - Dependency of other kinds can still be created by differences between groups in the effects of predictors (stay tuned)

Effects of Clustering on Effective N

- **Design Effect** expresses how much effective sample size needs to be adjusted due to clustering/grouping
- **Design Effect** = ratio of the variance obtained with the given sampling design to the variance obtained for a simple random sample from the same population, given the same total sample size either way

$n = \# \text{ level-1 units}$

- Design Effect = $1 + [(n - 1) * ICC]$
- Effective sample size $\rightarrow N_{\text{effective}} = \frac{\# \text{ Total Observations}}{\text{Design Effect}}$
- As ICC goes UP and cluster size goes UP, the effective sample size goes DOWN
 - See Snijders & Bosker (1999) ch. 3, p. 22-24 for more info

Design Effects in 2-Level Nesting

- Design Effect = $1 + [(n - 1) * ICC]$
- Effective sample size $\rightarrow N_{\text{effective}} = \frac{\# \text{ Total Observations}}{\text{Design Effect}}$
- $n=5$ patients from each of 100 doctors, $ICC = .30$?
 - Patients Design Effect = $1 + (4 * .30) = 2.20$
 - $N_{\text{effective}} = 500 / 2.20 = \mathbf{227}$ (not 500)
- $n=20$ students from each of 50 schools, $ICC = .05$?
 - Students Design Effect = $1 + (19 * .05) = 1.95$
 - $N_{\text{effective}} = 1000 / 1.95 = \mathbf{513}$ (not 1000)

Does a non-significant ICC mean you can ignore groups and just do a regression?

- 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 “safe” to ignore, not even 0!
 - An ICC=0 in an *empty (unconditional)* model can become ICC>0 after adding level-1 predictors, because reducing the residual variance leads to an increase in the random intercept variance (→ *conditional* ICC > 0)
- So just do a multilevel analysis anyway...
 - Even if “that’s not your question”... because people come from groups, you still have to model group dependency appropriately because of:
 - Effect of clustering on level-1 fixed effect SE’s → biased SEs
 - Potential for contextual effects of level-1 predictors

Predictors in MLM for Clustered Data Example: Achievement in Students nested in Schools

- Level-2 predictors now refer to Group-Level Variables
 - Can only have fixed or systematically varying effects (level-2 predictors cannot have random effects in a two-level model, same as before)
 - e.g., Does mean school achievement differ b/t rural and urban schools?
- Level-1 predictors now refer to Person-Level Variables
 - Can have fixed, systematically varying, or random effects over groups
 - e.g., Does student achievement differ between boys and girls?
 - Fixed effect: Is there a gender difference in achievement, period?
 - Systematically varying effect: Does the gender effect differ b/t rural and urban schools? (but the gender effect is the same within rural and within urban schools)
 - Random effect: Does the gender effect differ *randomly* across schools?
 - We can skip all the steps for building models for “time” and head straight to predictors (given that level-1 units are exchangeable here)

Two-Level Models for Clustered* Data

- Topics:
 - Fixed vs. random effects for modeling clustered data
 - ICC and design effects in clustered data
 - **Group-Mean-Centering vs. Grand-Mean Centering**
 - Model extensions under Group-MC and Grand-MC

* *Clustering = Nesting = Grouping...*

Predictors in MLM for Clustered Data

- BUT we still need to distinguish level-2 BG effects from level-1 WG effects of level-1 predictors: NO SMUSHING ALLOWED
- Options for representing level-2 BG variance as a predictor:
 - Use **obtained** group mean of level-1 x_{ij} from your sample (labeled as **GM x_j** or \bar{X}_j), centered at a constant so that 0 is a meaningful value
 - Use **actual** group mean of level-1 x_{ij} from outside data (also centered so 0 is meaningful) → better if your sample is not the full population
- Can use either **Group-MC** or **Grand-MC** for level-1 predictors (where Group-MC is like Person-MC in longitudinal models)
 - Level-1 Group-MC → center at a VARIABLE: **WG $x_{ij} = x_{ij} - \bar{X}_j$**
 - Level-1 Grand-MC → center at a CONSTANT: **L1 $x_{ij} = x_{ij} - C$**
 - Use L1 x_{ij} when including the actual group mean instead of sample group mean

3 Kinds of Effects for Level-1 Predictors

- **Is the Between-Group (BG) effect significant?**
 - Are groups with higher predictor values than other groups also higher on Y than other groups, such that the group mean of the person-level predictor GMx_j accounts for level-2 random intercept variance ($\tau_{U_0}^2$)?
- **Is the Within-Group (WG) effect significant?**
 - If you have higher predictor values than others in your group, do you also have higher outcomes values than others in your group, such that the within-group deviation WGx_{ij} accounts for level-1 residual variance (σ_e^2)?
- **Are the BG and WG effects different sizes: Is there a contextual effect?**
 - After controlling for the absolute value of level-1 predictor for each person, is there still an incremental contribution from having a higher group mean of the predictor (i.e., does a group's general tendency predict $\tau_{U_0}^2$ above and beyond)?
 - If there is no contextual effect, then the BG and WG effects of the level-1 predictor show convergence, such that their effects are of equivalent magnitude

Clustered Data Model with Group-Mean-Centered Level-1 x_{ij}

→ WG and BG Effects directly through separate parameters

x_{ij} is group-mean-centered into WGx_{ij} , with GMx_j at L2:

$$\text{Level 1: } y_{ij} = \beta_{0j} + \beta_{1j}(WGx_{ij}) + e_{ij}$$

$WGx_{ij} = x_{ij} - \bar{X}_j \rightarrow$ it has only Level-1 WG variation

$$\text{Level 2: } \beta_{0j} = Y_{00} + Y_{01}(GMx_j) + U_{0j}$$

$$\beta_{1j} = Y_{10}$$

$GMx_j = \bar{X}_j - C \rightarrow$ it has only Level-2 BG variation

Y_{10} = WG main effect of having more x_{ij} than others in your group

Y_{01} = BG main effect of having more \bar{X}_j than other groups

Because WGx_{ij} and GMx_j are uncorrelated, each gets the total effect for its level (WG=L1, BG=L2)

3 Kinds of Effects for Level-1 Predictors

- **What Group-Mean-Centering tells us directly:**
- **Is the Between-Group (BG) effect significant?**
 - Are groups with higher predictor values than other groups also higher on Y than other groups, such that the group mean of the person-level predictor \mathbf{GMx}_j accounts for level-2 random intercept variance ($\tau_{U_0}^2$)?
 - This would be indicated by a significant fixed effect of \mathbf{GMx}_j
 - Note: this is NOT controlling for the absolute value of x_{ij} for each person
- **Is the Within-Group (WG) effect significant?**
 - If you have higher predictor values than others in your group, do you also have higher outcomes values than others in your group, such that the within-group deviation \mathbf{WGx}_{ij} accounts for level-1 residual variance (σ_e^2)?
 - This would be indicated by a significant fixed effect of \mathbf{WGx}_{ij}
 - Note: this is represented by the relative value of x_{ij} , NOT the absolute value of x_{ij}

3 Kinds of Effects for Level-1 Predictors

- **What Group-Mean-Centering DOES NOT tell us directly:**
- **Are the BG and WG effects different sizes: Is there a contextual effect?**
 - After controlling for the absolute value of the level-1 predictor for each person, is there still an incremental contribution from the group mean of the predictor (i.e., does a group's general tendency predict $\tau_{U_0}^2$ above and beyond just the person-specific value of the predictor)?
 - In clustered data, the contextual effect is phrased as "after controlling for the individual, what is the additional contribution of the group"?
- **To answer this question about the contextual effect for the incremental contribution of the group mean, we have two options:**
 - Ask for the contextual effect via an ESTIMATE statement in SAS (or TEST in SPSS, or NEW in Mplus, or LINCOM in STATA): $\mathbf{WGx} \quad -1 \quad \mathbf{GMx} \quad 1$
 - Use "**grand-mean-centering**" for level-1 x_{ij} instead: $\mathbf{L1x}_{ij} = x_{ij} - C$
→ **centered at a CONSTANT, NOT A LEVEL-2 VARIABLE**
 - Which constant only matters for what the reference point is; it could be the grand mean or other

Group-MC vs. Grand-MC for Level-1 Predictors

| Level 2 | | Original | Group-MC Level 1 | Grand-MC Level 1 |
|-------------|-------------------------|----------|---------------------------------|-------------------------|
| \bar{X}_j | $GMx_j = \bar{X}_j - 5$ | x_{ij} | $WGx_{ij} = x_{ij} - \bar{X}_j$ | $L1x_{ij} = x_{ij} - 5$ |
| 3 | -2 | 2 | -1 | -3 |
| 3 | -2 | 4 | 1 | -1 |
| 7 | 2 | 6 | -1 | 1 |
| 7 | 2 | 8 | 1 | 3 |

Same GMx_j goes into the model using either way of centering the level-1 variable x_{ij}

Using **Group-MC**, WGx_{ij} has NO level-2 BG variation, so it is not correlated with GMx_j

Using **Grand-MC**, $L1x_{ij}$ STILL has level-2 BG variation, so it is STILL CORRELATED with GMx_j

So the effects of GMx_j and $L1x_{ij}$ when included together under Grand-MC will be different than their effects would be if they were by themselves...

Clustered Data Model with x_{ij} represented at Level 1 Only: → WG and BG Effects are Smushed Together

x_{ij} is grand-mean-centered into $L1x_{ij}$, WITHOUT GMx_j at L2:

Level 1: $y_{ij} = \beta_{0j} + \beta_{1j}(L1x_{ij}) + e_{ij}$

$L1x_{ij} = x_{ij} - C \rightarrow$ it still has both Level-2 BG and Level-1 WG variation

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

$\beta_{1j} = Y_{10}$

Y_{10} = *smushed* WG and BG effects

Because $L1x_{ij}$ still contains its original 2 different kinds of variation (BG and WG), its 1 fixed effect has to do the work of 2 predictors!

A *smushed* effect is also referred to as the *convergence, conflated, or composite* effect

Convergence (Smushed) Effect of a Level-1 Predictor

$$\text{Convergence Effect: } \gamma_{\text{conv}} \approx \frac{\frac{\gamma_{\text{BG}}}{\text{SE}_{\text{BG}}^2} + \frac{\gamma_{\text{WG}}}{\text{SE}_{\text{WG}}^2}}{\frac{1}{\text{SE}_{\text{BG}}^2} + \frac{1}{\text{SE}_{\text{WG}}^2}}$$

Adapted from
Raudenbush & Bryk
(2002, p. 138)

- **The convergence effect will often be closer to the within-group effect** (due to larger level-1 sample size and thus smaller SE)
- **It is the rule, not the exception, that between and within effects differ** (Snijders & Bosker, 1999, p. 52-56, and personal experience!)
- However—when grand-mean-centering a level-1 predictor, **convergence is testable** by including a **contextual effect (carried by the group mean)** for how the **BG effect** differs from the **WG effect**...

Clustered Data Model with Grand-Mean-Centered Level-1 x_{ij}

→ Model tests difference of WG vs. BG effects (It's been fixed!)

x_{ij} is grand-mean-centered into $L1x_{ij}$, WITH GMx_j at L2:

Level 1: $y_{ij} = \beta_{0j} + \beta_{1j}(L1x_{ij}) + e_{ij}$

$L1x_{ij} = x_{ij} - C \rightarrow$ it still
has both Level-2 BG and
Level-1 WG variation

Level 2: $\beta_{0j} = Y_{00} + Y_{01}(GMx_j) + U_{0j}$

$\beta_{1j} = Y_{10}$

$GMx_j = \bar{x}_j - C \rightarrow$ it has
only Level-2 BG variation

Y_{10} becomes the **WG effect** → **unique level-1 effect** after controlling for GMx_j

Y_{01} becomes the **contextual effect** that indicates how the BG effect differs from the WG effect → **unique level-2 effect** after controlling for $L1x_{ij}$ → does group matter beyond individuals?

Group-MC and Grand-MC Models are Equivalent Given a Fixed Level-1 Main Effect Only

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10}$

$\rightarrow y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$

$\rightarrow y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$

Composite Model:

← As Group-MC

← As Grand-MC

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10}$

$\rightarrow y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$

| Effect | Group-MC | Grand-MC |
|------------|-----------------------------|-----------------------------|
| Intercept | γ_{00} | γ_{00} |
| WG Effect | γ_{10} | γ_{10} |
| Contextual | $\gamma_{01} - \gamma_{10}$ | γ_{01} |
| BG Effect | γ_{01} | $\gamma_{01} + \gamma_{10}$ |

Contextual Effects in Clustered Data

- Group-MC is equivalent to Grand-MC if the group mean of the level-1 predictor is included and the level-1 effect is not random
- Grand-MC may be more convenient in clustered data due to its ability to directly provide contextual effects
- Example: Effect of SES for students (nested in schools) on achievement:
 - Group-MC** of level-1 student SES_{ij} , school mean \overline{SES}_j included at level 2
 - Level-1 **WG** effect: Effect of being rich kid relative to your school (is already purely WG because of centering around \overline{SES}_j)
 - Level-2 **BG** effect: Effect of going to a rich school NOT controlling for kid SES_{ij}
 - Grand-MC** of level-1 student SES_{ij} , school mean \overline{SES}_j included at level 2
 - Level-1 **WG** effect: Effect of being rich kid relative to your school (is purely WG after *statistically* controlling for \overline{SES}_j)
 - Level-2 **Contextual** effect: Incremental effect of going to a rich school (after *statistically* controlling for student SES)

3 Kinds of Effects for Level-1 Predictors

- **Is the Between-Group (BG) effect significant?**

- Are groups with higher predictor values than other groups also higher on Y than other groups, such that the group mean of the person-level predictor GMx_j accounts for level-2 random intercept variance (τ_{U0}^2)?
- Given directly by level-2 effect of GMx_j if using Group-MC for the level-1 predictor (or can be requested via ESTIMATE if using Grand-MC for the level-1 predictor)

- **Is the Within-Group (WG) effect significant?**

- If you have higher predictor values than others in your group, do you also have higher outcomes values than others in your group, such that the within-group deviation WGx_{ij} accounts for level-1 residual variance (σ_e^2)?
- Given directly by the level-1 effect of WGx_{ij} if using Group-MC —OR— given directly by the level-1 effect of $L1x_{ij}$ if using Grand-MC and including GMx_j at level 2 (without GMx_j , the level-1 effect of $L1x_{ij}$ if using Grand-MC is the smushed effect)

- **Are the BG and WG effects different sizes: Is there a contextual effect?**

- After controlling for the absolute value of the level-1 predictor for each person, is there still an incremental contribution from the group mean of the predictor (i.e., does a group's general tendency predict τ_{U0}^2 above and beyond the person-specific predictor value)?
- Given directly by level-2 effect of GMx_j if using Grand-MC for the level-1 predictor (or can be requested via ESTIMATE if using Group-MC for the level-1 predictor)

Variance Accounted For By Level-2 Predictors

- **Fixed effects of level 2 predictors *by themselves*:**

- Level-2 (BG) main effects reduce level-2 (BG) random intercept variance
- Level-2 (BG) interactions also reduce level-2 (BG) random intercept variance

- **Fixed effects of *cross-level interactions* (level 1* level 2):**

- If the interacting level-1 predictor is random, any cross-level interaction with it will reduce its corresponding level-2 BG random slope variance (that line's U)
- If the interacting level-1 predictor not random, any cross-level interaction with it will reduce the level-1 WG residual variance instead
 - This is because the level-2 BG random slope variance would have been created by decomposing the level-1 residual variance in the first place
 - The level-1 effect would then be called "**systematically varying**" to reflect a compromise between "fixed" (all the same) and "random" (all different)—it's not that each group needs their own slope, but that the slope varies systematically across groups as a function of a known group predictor (and not otherwise)

Variance Accounted For By Level-1 Predictors

- **Fixed effects of level 1 predictors *by themselves*:**
 - Level-1 (WG) main effects reduce Level-1 (WG) residual variance
 - Level-1 (WG) interactions also reduce Level-1 (WG) residual variance
- **What happens at level 2 depends on what kind of variance the level-1 predictor has:**
 - If the level-1 predictor ALSO has level-2 variance (e.g., Grand-MC predictors), then its level-2 variance will also likely reduce level-2 random intercept variance
 - If the level-1 predictor DOES NOT have level-2 variance (e.g., Group-MC predictors), then its reduction in the level-1 residual variance will cause an INCREASE in level-2 random intercept variance
 - Same thing happens with Grand-MC level-1 predictors, but you don't generally see it
 - It's just an artifact that the estimate of true random intercept variance is:
$$\text{True } \tau_{U_0}^2 = \text{observed } \tau_{U_0}^2 - \frac{\sigma_e^2}{n} \rightarrow \text{so if only } \sigma_e^2 \text{ decreases, } \tau_{U_0}^2 \text{ increases}$$

Two-Level Models for Clustered* Data

- Topics:
 - Fixed vs. random effects for modeling clustered data
 - ICC and design effects in clustered data
 - Group-Mean-Centering vs. Grand-Mean Centering
 - **Model extensions under Group-MC and Grand-MC**

* *Clustering = Nesting = Grouping...*

The Joy of Interactions Involving Level-1 Predictors

- Must consider interactions with both its BG and WG parts:
- Example: Does the effect of employee motivation (x_{ij}) on employee performance interact with type of business (for profit or non-profit; $Type_j$)?
- Group-Mean-Centering:
 - $WGx_{ij} * Type_j \rightarrow$ Does the WG motivation effect differ between business types?
 - $GMx_j * Type_j \rightarrow$ Does the BG motivation effect differ between business types?
 - Moderation of total group motivation effect (not controlling for individual motivation)
 - If forgotten, then $Type_j$ moderates the motivation effect only at level 1 (WG, not BG)
- Grand-Mean-Centering:
 - $L1x_{ij} * Type_j \rightarrow$ Does the WG motivation effect differ between business types?
 - $GMx_j * Type_j \rightarrow$ Does the *contextual* motivation effect differ b/t business types?
 - Moderation of incremental group motivation effect controlling for employee motivation (moderation of the "boost" in group performance from working with motivated people)
 - If forgotten, then although the level-1 main effect of motivation has been un-smushed via the main effect of GMx_j , the interaction of $L1x_{ij} * Type_j$ would still be smushed

Interactions with Level-1 Predictors: Example: Employee Motivation (x_{ij}) by Business Type ($Type_j$)

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

$$\text{Level-1: } y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$$

$$\text{Level-2: } \beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(Sex_i)$$

$$\text{Composite: } y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij} \\ + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij} - GMx_j)$$

Grand-MC: $L1x_{ij} = x_{ij}$

$$\text{Level-1: } y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$$

$$\text{Level-2: } \beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(Type_j)$$

$$\text{Composite: } y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij} \\ + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij})$$

Interactions Involving Level-1 Predictors Belong at Both Levels of the Model

On the left below → Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

$$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij} + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij} - GMx_j)$$

← As Group-MC

$$y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij} + \gamma_{02}(Type_j) + (\gamma_{03} - \gamma_{11})(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij})$$

← As Grand-MC

On the right below → Grand-MC: $L1x_{ij} = x_{ij}$

$$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij} + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij})$$

After adding an interaction for $Type_j$ with x_{ij} at both levels, then the Group-MC and Grand-MC models are equivalent

Intercept: $\gamma_{00} = \gamma_{00}$

BG Effect: $\gamma_{01} = \gamma_{01} + \gamma_{10}$

Contextual: $\gamma_{01} = \gamma_{01} - \gamma_{10}$

WG Effect: $\gamma_{10} = \gamma_{10}$

BG*Type Effect: $\gamma_{03} = \gamma_{03} + \gamma_{11}$

Contextual*Type: $\gamma_{03} = \gamma_{03} - \gamma_{11}$

Type Effect: $\gamma_{20} = \gamma_{20}$

BG*WG or Contextual*WG is the same: $\gamma_{11} = \gamma_{11}$

Intra-variable Interactions

- Still must consider interactions with both its BG and WG parts!
- Example: Does the effect of employee motivation (x_{ij}) on employee performance interact with business group mean motivation (GMx_j)?
- Group-Mean-Centering:
 - $WGx_{ij} * GMx_j \rightarrow$ Does the WG motivation effect differ by group motivation?
 - $GMx_j * GMx_j \rightarrow$ Does the BG motivation effect differ by group motivation?
 - Moderation of total group motivation effect (not controlling for individual motivation)
 - If forgotten, then GMx_j moderates the motivation effect only at level 1 (WG, not BG)
- Grand-Mean-Centering:
 - $L1x_{ij} * GMx_j \rightarrow$ Does the WG motivation effect differ by group motivation?
 - $GMx_j * GMx_j \rightarrow$ Does the *contextual* motivation effect differ by group motiv.?
 - Moderation of incremental group motivation effect controlling for employee motivation (moderation of the boost in group performance from working with motivated people)
 - If forgotten, then although the level-1 main effect of motivation has been un-smushed via the main effect of GMx_j , the interaction of $L1x_{ij} * GMx_j$ would still be smushed

Intra-variable Interactions:

Example: Employee Motivation (x_{ij}) by Business Mean Motivation (GMx_j)

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(GMx_j)(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(GMx_j)$

Composite: $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$
 $+ \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij} - GMx_j)$

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(GMx_j)(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(GMx_j)$

Composite: $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
 $+ \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij})$

Intra-variable Interactions:

Example: Employee Motivation (x_{ij}) by Business Mean Motivation (GMx_j)

On the left below → Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$
 $+ \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij} - GMx_j)$

$y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
 $+ (\gamma_{02} - \gamma_{11})(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij})$

← As Group-MC

← As Grand-MC

On the right below → Grand-MC: $L1x_{ij} = x_{ij}$

$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
 $+ \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij})$

After adding an interaction for **Type_j** with x_{ij} at both levels, then the Group-MC and Grand-MC models are equivalent

Intercept: $\gamma_{00} = \gamma_{00}$

BG Effect: $\gamma_{01} = \gamma_{01} + \gamma_{10}$

Contextual: $\gamma_{01} = \gamma_{01} - \gamma_{10}$

WG Effect: $\gamma_{10} = \gamma_{10}$

BG² Effect: $\gamma_{02} = \gamma_{02} + \gamma_{11}$

Contextual²: $\gamma_{02} = \gamma_{02} - \gamma_{11}$

BG*WG or Contextual*WG is the same: $\gamma_{11} = \gamma_{11}$

When Group-MC \neq Grand-MC: Random Effects of Level-1 Predictors

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + U_{1j}$

Variance due to GMx_j is removed from the random slope in Group-MC.

$\rightarrow y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + U_{1j}(x_{ij} - GMx_j) + e_{ij}$

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + U_{1j}$

Variance due to GMx_j is still part of the random slope in Grand-MC. So these models cannot be made equivalent.

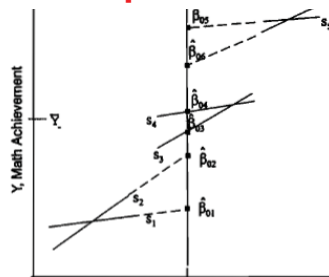
$\rightarrow y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + U_{1j}(x_{ij}) + e_{ij}$

Random Effects of Level-1 Predictors

- **Random intercepts** mean different things under each model:
 - **Group-MC** \rightarrow Group differences at $WGx_{ij} = 0$ (that every group has)
 - **Grand-MC** \rightarrow Group differences at $L1x_{ij} = 0$ (that not every group will have)
- **Differential shrinkage of the random intercepts** results from differential reliability of the intercept data across models:
 - Group-MC \rightarrow Won't affect shrinkage of slopes unless highly correlated
 - Grand-MC \rightarrow Will affect shrinkage of slopes due to forced extrapolation
- As a result, the **random slope variance may be smaller** under Grand-MC than under Group-MC
 - Problem worsens with greater ICC of level-1 predictor (more extrapolation)
 - Anecdotal example was presented in Raudenbush & Bryk (2002; chapter 5)

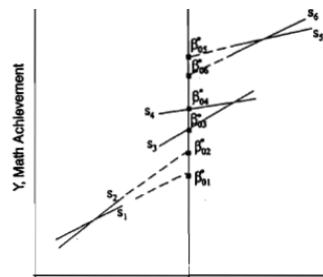
Bias in Random Slope Variance

OLS Per-Group Estimates



Level-1 X

EB Shrunk Estimates



Level-1 X

Top right: Intercepts and slopes are homogenized in Grand-MC because of intercept extrapolation

Bottom: Downwardly-biased random slope variance in Grand-MC relative to Group-MC

| Unconditional Results | Conditional Results |
|---|---|
| Group-MC | |
| $\hat{\mathbf{T}} = \begin{bmatrix} 8.68 & 0.05 \\ 0.05 & 0.68 \end{bmatrix}$ | $\hat{\mathbf{T}} = \begin{bmatrix} 2.38 & 0.19 \\ 0.19 & 0.15 \end{bmatrix}$ |
| $\hat{\sigma}^2 = 36.70$ | $\hat{\sigma}^2 = 36.70$ |
| Grand-MC | |
| $\hat{\mathbf{T}} = \begin{bmatrix} 4.83 & -0.15 \\ -0.15 & 0.42 \end{bmatrix}$ | $\hat{\mathbf{T}} = \begin{bmatrix} 2.41 & 0.19 \\ 0.19 & 0.06 \end{bmatrix}$ |
| $\hat{\sigma}^2 = 36.83$ | $\hat{\sigma}^2 = 36.74$ |

MLM for Clustered Data: Summary

- Models now come in only two kinds: "empty" and "conditional"
 - The lack of a comparable dimension to "time" simplifies things greatly!
- L2 = Between-Group, L1 = Within-Group (between-person)
 - Level-2 predictors are group variables: can have fixed or systematically varying effects (but not random effects in two-level models)
 - Level-1 predictors are person variables: can have fixed, random, or systematically varying effects
- No smushing main effects or interactions of level-1 predictors:
 - Group-MC at Level 1: Get L1=WG and L2=BG effects directly
 - Grand-MC at Level 1: Get L1=WG and L2=contextual effects directly
 - As long as some representation of the L1 effect is included in L2; otherwise, the L1 effect (and any interactions thereof) will be smushed

Three-Level Models for Clustered Longitudinal Data

- 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 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: $\delta_{00j} = Y_{000} + V_{00j}$

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 Y_{00}

Model for the Variance (2):

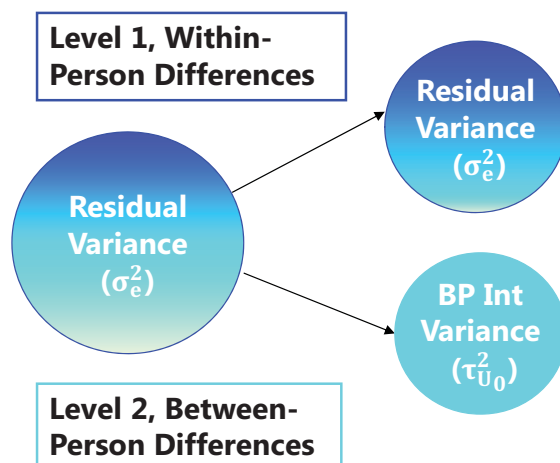
- Level-1 Variance of $e_{tij} \rightarrow \sigma_e^2$
- Level-2 Variance of $U_{0ij} \rightarrow \tau_{U_0}^2$
- Level-3 Variance of $V_{00j} \rightarrow \tau_{V_{00}}^2$

Composite equation:

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

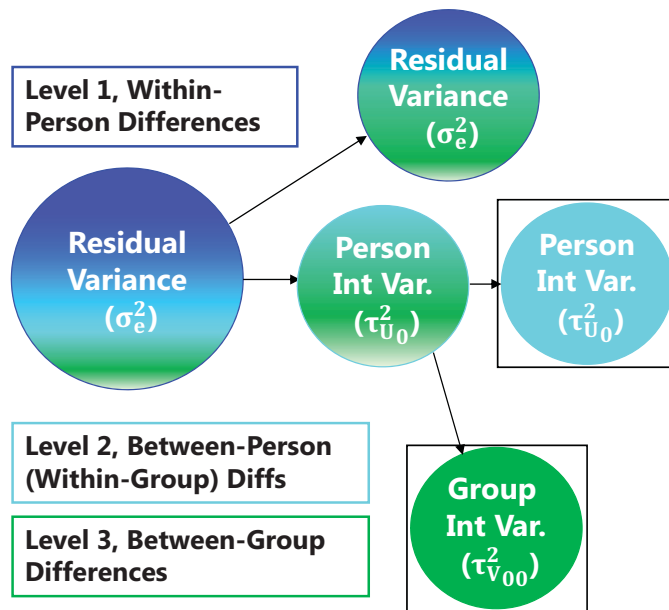
2-Level Random Intercept Model

- 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:



Lecture 6

5

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:**

$$ICC_{L2} = \frac{\text{Between-Person}}{\text{Total}} = \frac{L3+L2}{L3+L2+L1} = \frac{\tau_{V00}^2 + \tau_{U0}^2}{\tau_{V00}^2 + \tau_{U0}^2 + \sigma_e^2}$$

→ 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):**

$$ICC_{L3} = \frac{\text{Between-Group}}{\text{Between-Person}} = \frac{L3}{L3+L2} = \frac{\tau_{V00}^2}{\tau_{V00}^2 + \tau_{U0}^2}$$

→ 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?**

Lecture 6

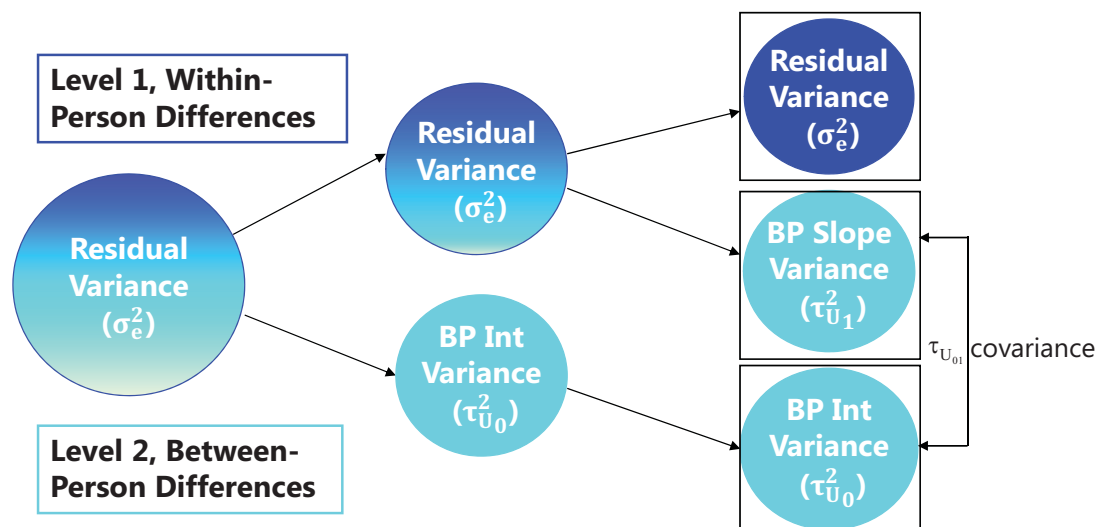
6

Three-Level Models for Clustered Longitudinal Data

- 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 designs

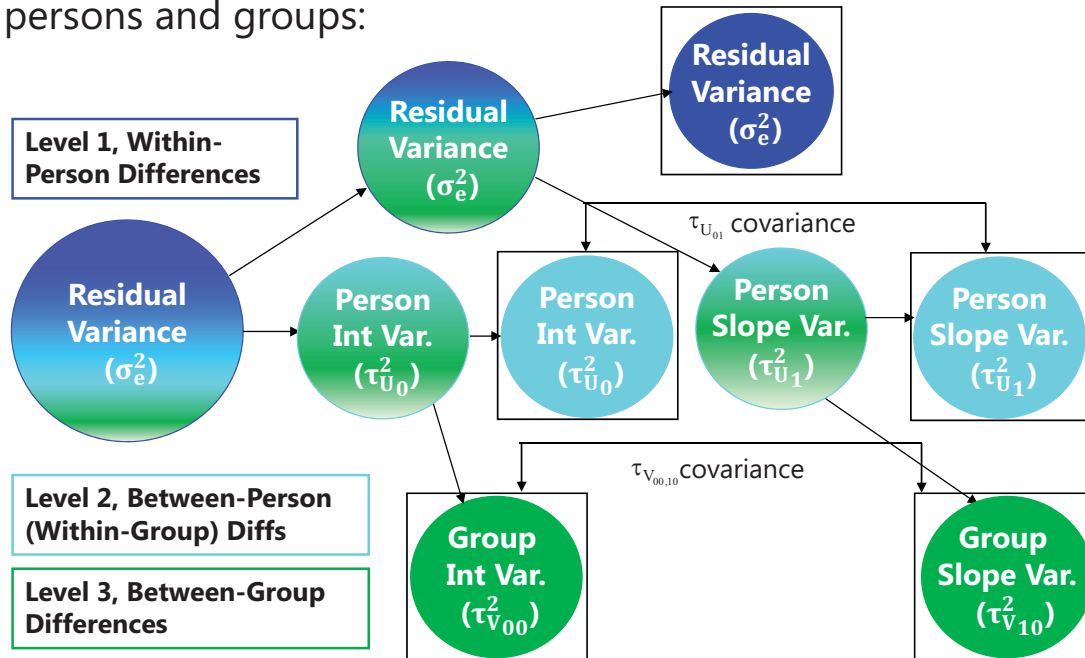
2-Level Random Slope Model

- 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

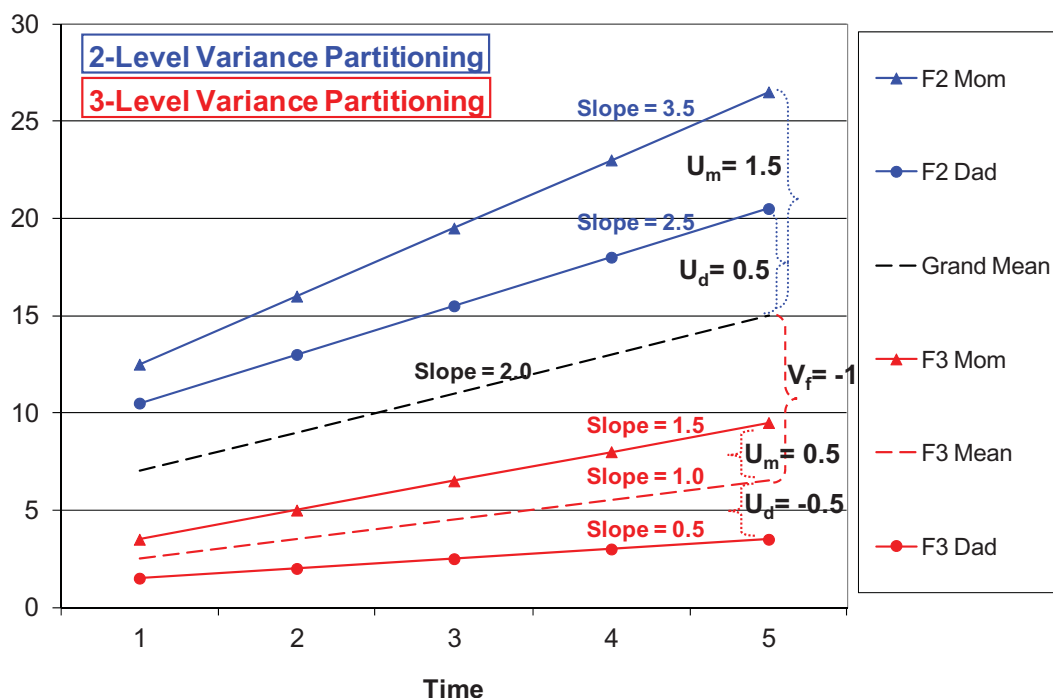
- In a 3-level model, we can have random effects of time over persons and groups:



Lecture 6

9

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



Lecture 6

10

3-Level Random Time Slope Model

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

Level 1: $y_{tij} = \beta_{0ij} + \beta_{1ij}(\text{Time}_{tij}) + e_{tij}$ ← **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 ($\tau_{U0}^2, \tau_{U1}^2, \tau_{U01}$)**

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 ($\tau_{V00}^2, \tau_{V10}^2, \tau_{V00,10}$)**

**Fixed Intercept,
Fixed Linear
Time Slope**

Composite equation (9 parameters):

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

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)

$$ICC_{Int} = \frac{\text{Between} - \text{Group}}{\text{Between} - \text{Person}} = \frac{\text{L3 Int}}{\text{L3 Int} + \text{L2 Int}} = \frac{\tau_{V00}^2}{\tau_{V00}^2 + \tau_{U0}^2}$$

$$ICC_{Slope} = \frac{\text{Between} - \text{Group}}{\text{Between} - \text{Person}} = \frac{\text{L3 Slope}}{\text{L3 Slope} + \text{L2 Slope}} = \frac{\tau_{V10}^2}{\tau_{V10}^2 + \tau_{U1}^2}$$

- 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")

Three-Level Models for Clustered Longitudinal Data

- 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 designs

Conditional Model Specification

- Remember separating between- and within-person effects?
Now there are 3 potential effects for any level-1 predictor!
 - 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
- 2 potential effects for any level-2 predictor, also
 - 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*
 - $WPstress_{tij} = Stress_{tij} - PersonMeanStress_{ij}$
 - Direct tests if within-person effect $\neq 0$?
 - **Total** within-person effect of having more stress **than usual** $\neq 0$?
- **Level 2 (Person):** *Person mean stress relative to family*
 - $WFstress_{ij} = PersonMeanStress_{ij} - FamilyMeanStress_j$
 - Direct tests if within-family effect $\neq 0$?
 - **Total** effect of having more stress **than other family members** $\neq 0$?
- **Level 3 (Family):** *Family mean stress relative to all families (from constant)*
 - $BFstress_j = FamilyMeanStress_j - C$
 - Direct tests if between-family effect $\neq 0$?
 - **Total** effect of having more stress **than other families** $\neq 0$?

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: $y_{tij} = \beta_{0ij} + \beta_{1ij}(\text{Time}_{tij}) + \beta_{2ij}(\text{Stress}_{tij} - \text{PMstress}_{ij}) + e_{tij}$

Level 2: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(\text{PMstress}_{ij} - \text{FMstress}_j) + U_{0ij}$

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

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

Level 3: $\delta_{00j} = Y_{000} + Y_{001}(\text{FMstress}_j - C) + V_{00j}$

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

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

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

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)*
 - $\text{TVstress}_{tij} = \text{Stress}_{tij} - C$
 - Direct tests if within-person effect $\neq 0$?
 - **Total** within-person effect of having more stress **than usual** $\neq 0$?
- **Level 2 (Person):** *Person mean stress (relative to sample constant)*
 - $\text{BPstress}_{ij} = \text{PersonMeanStress}_{ij} - C$
 - Direct tests if within-person and within-family effects $\neq ?$
 - **Contextual** effect of having more stress **than other family members** $\neq 0$?
- **Level 3 (Family):** *Family mean stress relative to all families (from constant)*
 - $\text{BFstress}_j = \text{FamilyMeanStress}_j - C$
 - Direct tests if within-family and between-family effects $\neq ?$
 - **Contextual** effect of having more stress **than other families** $\neq 0$?

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: $y_{tij} = \beta_{0ij} + \beta_{1ij}(\text{Time}_{tij}) + \beta_{2ij}(\text{Stress}_{tij} - C) + e_{tij}$

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

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

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

Level 3: $\delta_{00j} = Y_{000} + Y_{001}(\text{FMstress}_j - C) + V_{00j}$

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

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

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

Fixed intercept,
Contextual family
stress main effect

Contextual 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 does not exist (= 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 equivalent to 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: $BFstress_j * BFcope_j$, $(BFstress_j * WFcope_{ij})$

Lecture 6

21

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: $y_{tij} = \beta_{0ij} + \beta_{1ij}(\text{Time}_{tij}) + \beta_{2ij}(\text{Stress}_{tij} - \text{PMstress}_{ij}) + e_{tij}$

Level 2: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(\text{PMstress}_{ij} - \text{FMstress}_j) + \delta_{02j}(\text{Cope}_{ij} - \text{FMcope}_j) + \delta_{03j}(\text{PMstress}_{ij} - \text{FMstress}_j)(\text{Cope}_{ij} - \text{FMcope}_j) + U_{0ij}$
 $\beta_{1ij} = \delta_{10j} + U_{1ij}$
 $\beta_{2ij} = \delta_{20j} + \delta_{21j}(\text{Cope}_{ij} - \text{FMcope}_j) + (U_{2ij})$

Level 3: $\delta_{00j} = Y_{000} + Y_{001}(\text{FMstress}_j - C) + Y_{002}(\text{FMcope}_j - C) + Y_{003}(\text{FMstress}_j - C)(\text{FMcope}_j - C) + V_{00j}$
 $\delta_{01j} = Y_{010} + (V_{01j})$ $\delta_{02j} = Y_{020} + (V_{02j})$ $\delta_{03j} = Y_{030} + (V_{03j})$
 $\delta_{10j} = Y_{100} + V_{10j}$
 $\delta_{20j} = Y_{200} + Y_{202}(\text{FMcope}_j - C) + (V_{20j})$ $\delta_{21j} = Y_{210} + (V_{21j})$

Lecture 6

22

Summary: Clustered Longitudinal Models

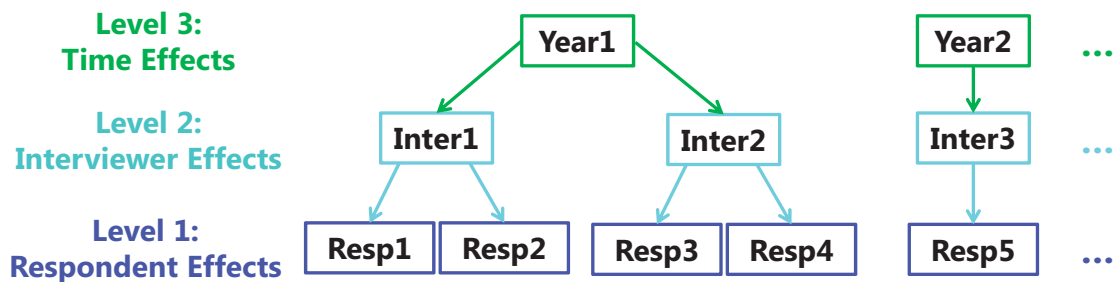
- Estimating 3-level models requires no new concepts, but everything is just at an order of complexity higher:
 - Proportioning variance over 3 levels instead of 2 → 2+ ICCs
 - Random slope variance will come from term directly beneath:
 - Level-2 random slope comes from level-1 residual
 - Level-3 random 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....

Three-Level Models for Clustered Longitudinal Data

- 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 designs**

Other 3-Level Designs

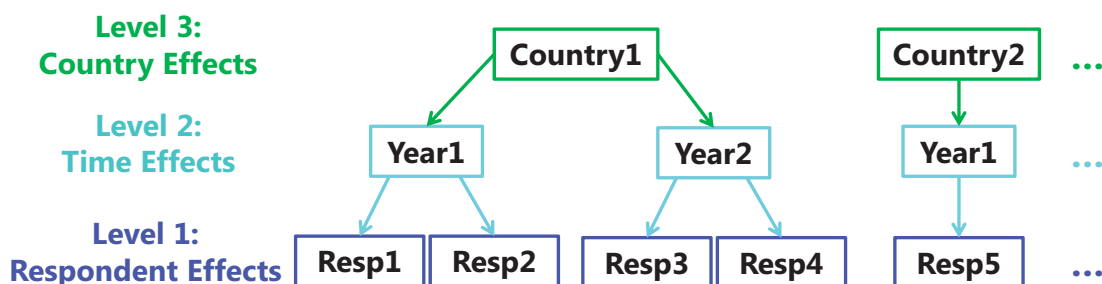
- The sampling design for the outcome (not the predictors) dictates what your levels will be, **so time may not always be level 1**
- Example: Predicting answer compliance in respondents nested in interviewers, collected over several years (all different people)



- 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

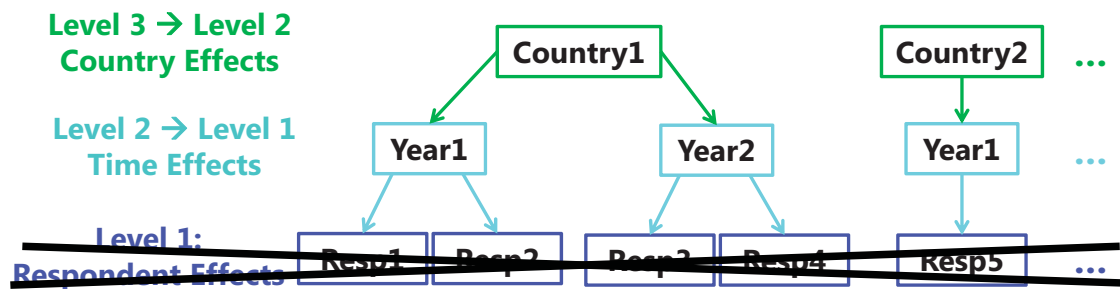
- Another example: Predicting **time-specific respondent outcomes** for people nested in countries, collected over several years (all different people, but the same countries measured over time)



- 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**?

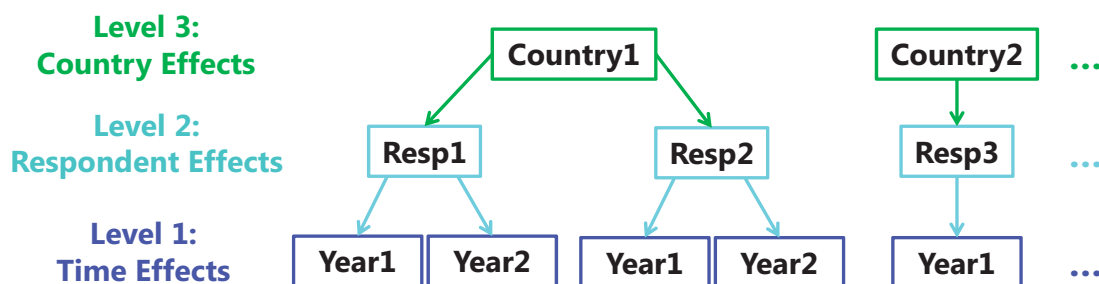


Because the outcome was measured at level 2 (country per time):

- 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 (all same people and same countries 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 time-varying predictors?
 - For People: effects should be included at all 3 levels (+random over 2 and 3)
 - For Countries: effects are only possible at levels 1 and 3 (+random over 3)

Appendix A: Overview of Multilevel Modeling Texts and Suggested Readings

Textbooks for Multilevel Modeling

These texts cover multilevel modeling within the context of clustered (nested) observations primarily. They are ordered in terms of my opinion of their accessibility (most to least).

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Heck, R. H., & Thomas, S. L. (2008). *An introduction to multilevel modeling techniques* (2nd ed.). New York: Routledge

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.

Snijders, T. A. B., & Bosker, R. (1999 1st ed.; 2011 2nd ed.). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd Ed.). Thousand Oaks, CA: Sage.

These texts cover multilevel modeling within the context of longitudinal observations primarily. They are ordered in terms of my opinion of their accessibility (most to least).

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Fitzmaurice, G., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. New York: Wiley.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. New York: Wiley.

Verbeke, G., & Molenberghs, G. (2001). *Linear mixed models for longitudinal data*. New York: Springer-Verlag.

Diggle, P. J., Heagerty, P. J., Liang, K. Y., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). New York: Oxford University Press.

These texts cover longitudinal models within the context of structural equation modeling.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling. Quantitative applications in the social sciences, #157*. Thousand Oaks, CA: Sage.

Bollen, K. A., & Curran, P. J. (2005). *Latent curve models: A structural equation perspective*. New York: Wiley.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Erlbaum.

The latter chapters in this ANOVA text introduce MLM from the ANOVA perspective.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data*. Mahwah, NJ: Erlbaum.

Suggested Readings by Topic

Lectures 1 and 2. Introduction to MLM

- Snijders & Bosker ch. 1-2
- Singer & Willett ch. 1-2
- Raudenbush & Bryk ch. 2
- Hoffman ch. 1

Lecture 1. Review of General Linear Models and Repeated Measures ANOVA

- Hedeker & Gibbons ch. 1-3
- Fitzmaurice, Laird, & Ware ch. 5-6
- Hoffman ch. 2-3

Lecture 2. Fixed and Random Effects of Time

- Singer & Willett ch. 3-6
- Hedeker & Gibbons ch. 4
- Willett, J.B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement*, 49, 587-602.
- Rovine, M. J., & Molenaar, P. C. M. (1998). The covariance between level and shape in the latent growth curve model with estimated basis vector coefficients. *Methods of Psychological Research Online*, 3(2), 95-107.
- Snijders & Bosker ch. 4, 12
- Hox ch. 5
- Raudenbush & Bryk ch. 6
- Hoffman ch. 4-6
- Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, 58, 615-637.
- Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling*, 16, 676-701.

Lecture 2. Fun with Model Comparisons

- Singer & Willett ch. 4
- Snijders & Bosker ch. 6-7
- Raudenbush & Bryk ch. 3
- Stoel, R. D., Garre, F. G., Dolan, C., & van den Wittenboer, G. (2006). On the likelihood ratio test in structural equation modeling when parameters are subject to boundary constraints. *Psychological Methods*, 11(4), 439-455.
- Verbeke & Molenberghs ch. 5-6
- Hoffman ch. 3 and 5

Lecture 3. Time-Invariant Predictors

Lecture 4. Time-Varying Predictors and Centering Decisions

- Hoffman, L., & Stawski, R. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Development*, 6(2-3), 97-100. Available at: <http://digitalcommons.unl.edu/psychfacpub/415/>.
- Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 24(5), 623-641.
- Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, 30(1), 1-21.
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13(3), 203-229.
- Singer & Willett ch. 5
- Snijders & Bosker ch. 3-5
- Raudenbush & Bryk ch. 5
- Fitzmaurice, Laird, & Ware ch. 15
- Hedeker & Gibbons ch. 4
- Hoffman ch. 7-8 (chapter 9 not drafted yet)

Lecture 5. Two-Level Models for Clustered Observations

Lecture 6. Three-Level Models for Clustered Longitudinal Observations

- Raudenbush & Bryk ch. 5, 8
- Snijders & Bosker ch. 4-5
- Hedeker & Gibbons ch. 13

SAS, SPSS, and STATA Multilevel Modeling Syntax Guides

```

/*****
*****          SAS MULTILEVEL MODELING SYNTAX          *****
*****
*****/

/* PROC MIXED STATEMENT:
    DATA=:      File to use - default is last accessed
    NOCLPRINT:   Do not print class variable values
    NOITPRINT:   Do not print iteration history
    COVTEST:     Print SEs and p-values for significance test of variance estimates
    NAMELEN=:    # characters printed in fixed effects tables (default=20)
    METHOD=:      select REML or ML estimator - REML is default
    Other options...
    IC:          Print other information criteria and associated df
    MAXITER=:    # iterations (default=50)
    EMPIRICAL:   adjust SEs for non-normality of residuals (sandwich estimator) --
                not available with Satterthwaite or KR DDFM (use BW instead) */
PROC MIXED DATA=work.datafile NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;

/* CLASS statement makes SAS dummy code categorical variables (highest value is reference)
    Also include ID variables and REPEATED variables on CLASS statement */
CLASS IDvar catvar1 catvar2 cattime;

/* MODEL STATEMENT:
    Predict DV from time, catvar1, catvar2, contvar, 3 example interactions
    Don't need to construct interaction terms as variables a priori, use * between variables
    Options after the / ...
        NOINT:      Remove fixed intercept (is included by default)
        SOLUTION:   Print fixed effects solution (not included by default)
        DDFM=:      Change denominator degrees of freedom
                    Choose from Satterthwaite, KR, BW
    Other options...
        OUTP=:      Save predicted values from fixed+random effects to =dataset
        OUTPM=:     Save predicted values from fixed effects only to =dataset */
MODEL DV = time catvar1 catvar2 contvar catvar1*catvar2 time*catvar1 time*catvar2
/ SOLUTION DDFM=Satterthwaite;

/* RANDOM STATEMENT: Random Intercept must be listed if needed, also list any random slopes
    Options after the / ...
        G GCORR:    Print covariance and correlation matrices for random effects
        V VCORR:    Print covariance and correlation matrices for total outcome
                    V=1 VCORR=1 prints for 1st case (is default - can change if needed)
        TYPE=UN:    UN for unstructured to allow random effect covariances (VC=default)
        SUBJECT:    ID variable to identify nesting per level
    Other options...
        SOLUTION:   Print solution of random effects (the U's) - will take long time
        GROUP=:     =groupvar by which to get separate G matrices */
RANDOM INTERCEPT time / G GCORR V VCORR TYPE=UN SUBJECT=IDvar;

/* REPEATED STATEMENT: is always there even if not listed
    List variables repeated over (e.g., cattime) here and on CLASS statement
    Options after the / ...
        R RCORR:    Print covariance and correlation matrices for residuals
                    R=1 RCORR=1 prints for 1st case (is default - can change if needed)
        TYPE=:      VC for variance components (diagonal) by default
                    many, many other types available, such as AR(1), TOEP(n)
        SUBJECT:    ID variable to identify nesting per level (what is repeated over)
    Other options...
        GROUP=:     =groupvar by which to get separate R matrices
        LOCAL=:     EXP(predictor) for predictors of log of residual variance */
REPEATED cattime / R RCORR TYPE=VC SUBJECT=IDvar;

```

```

/* Execute PROC MIXED */ RUN;
/*****      OTHER OPTIONAL SAS PROC MIXED STATEMENTS      *****/

/* PARMS statement used to provide start values for variance components
   Must list as many ( ) as there are variance components in the model
   List in order of appearance in CovParms table
   Options after the / ...
       HOLD=:      means fix those components (here, #3 is fixed to 1)
       NOBOUND:    allows variances to go negative (useful for troubleshooting) */
PARMS (5) (3) (2) (1) / HOLD=3 NOBOUND;

/*      LSMEANS Generates means & tests for specified variables
       SLICE Tests effect of catvar1 at each level of catvar2
       if the time variable is at 1, uses Tukey Adjustment
       DIFF=ALL requests all possible pairwise comparisons */
LSMEANS catvar1*catvar2 / SLICE=catvar2 AT time=1 ADJUST=TUKEY DIFF=ALL;

/*      ESTIMATE is used for specific hypothesis tests
       "LABEL" is first, followed by effect being estimated.
       Below we ask for group differences in main effect of catvar1
       and group differences in the linear slope of time */
ESTIMATE "L vs. H Catvar1 for Main Effect"    catvar1 -1 0 1 ;
ESTIMATE "M vs. H Catvar1 for Main Effect"    catvar1 0 -1 1 ;
ESTIMATE "L vs. M Catvar1 for Main Effect"    catvar1 -1 1 0 ;
ESTIMATE "L vs. H Catvar1 for Linear Slope"    time*catvar1 -1 0 1 ;
ESTIMATE "M vs. H Catvar1 for Linear Slope"    time*catvar1 0 -1 1 ;
ESTIMATE "L vs. M Catvar1 for Linear Slope"    time*catvar1 -1 1 0 ;

/*      Below we ask for simple effects at the interacting variable=3 */
ESTIMATE "Simple Effect of X if Z=3"    xvar 1    xvar*zvar 3;
ESTIMATE "Simple Effect of Z if X=3"    zvar 1    xvar*zvar 3;

/*      ODS OUTPUT is used to save output tables to SAS datasets
       SolutionF=:      Save fixed effects to =dataset
       SolutionR=:      Save random effects to =dataset
       CovParms=:       Save covariance parameters to =dataset
       FitStatistics=:   Save fit statistics to =dataset
       Estimates=:       Save requested estimates to =dataset */
ODS OUTPUT    SolutionF=work.FixedEffects
               CovParms=work.CovarianceParameters
               FitStatistics=work.FitStats;

*****
*****      SPSS MULTILEVEL MODELING SYNTAX      *****
*****

* MIXED STATEMENT: DV is listed first
  BY:    list categorical predictors (main effects only)
  WITH:  list continuous predictors (main effects only).

* /METHOD=REML: used to select estimator (REML is default, ML is also available).

* /PRINT=: used to request specific output
  SOLUTION:      Print fixed effects solution (not default)
  TESTCOV:       Print SEs and p-values for significance tests for variances (not default)
  G:             Covariance matrix of random effects (no G correlation matrix available)
  R:             Covariance matrix of residuals (no R correlation matrix available)
  CPS:          Case processing summary: factor values, repeated measures variables,
               repeated measure subjects, random effects subjects & frequencies
  DESCRIPTIVES: Sample sizes, means, SD of DV & covariates for each combination of factors
  HISTORY(1):    Iteration History (1=print every iteration).

```



```

* /FIXED=: used to specify fixed effects (intercept included by default)
    Don't have to define interaction terms ahead of time, can do so via * between variables
    Options after the | :
        SSTYPE(3): Sums of Squares Type (3=default, also 1 available)
        NOTINT: To remove fixed intercept.

* /RANDOM=: used to specify random effects (intercept NOT included by default)
    Options after the | :
        COVTYPE(UN)=: UN for unstructure to allow random effects covariances (not default)
        SUBJECT(IDvar): ID variable to identify nesting per level.

* /REPEATED=: is always there even if not listed
    List variables repeated over (e.g., cattime) here and on BY statement
    Options after the | :
        COVTYPE(DIAG)=: DIAG for diagonal (default)
        SUBJECT(IDvar): ID variable to identify nesting per level (what repeated over).

***** OTHER OPTIONAL MIXED STATEMENTS *****

* /CRITERIA: used to change estimation options (leave the convergence ones alone)
    MXITER(100): Change number of iterations (100=default).

* /SAVE=: used to save predicted values to dataset
    FIXPRED: Save estimates for predicted values from fixed effects only
    SEFIXP: Save SE for predicted values from fixed effects only
    DFFIXP: Save Satterthwaite DDFM for predicted values from fixed effects only
    PRED: Save estimates for predicted values from fixed+random effects
    SEPRED: Save SE for predicted values from fixed+random effects
    DFPRED: Save Satterthwaite DDFM for predicted values from fixed+random effects
    RESID: Save residuals from fixed+random effects.

* /EMMEANS: used to request means for categorical predictors
    TABLES: list variables to get means for (unique combination)
    WITH: values of continuous predictors to be evaluated at
    COMPARE(var): variable to be compared per level of interacting variable
    ADJ(LSD): pairwise comparison (LSD=no adjustment, also Bonferroni, Sideak).

* /TEST is used for specific hypothesis tests
    "LABEL" is first, followed by effect being estimated
    Below we ask for group differences in main effect of catvar1
    and group differences in the linear slope of time
    Next we ask for simple effects at the interacting variable=3.

```

```

MIXED DV BY IDvar catvar1 catvar2 WITH time contvar
/METHOD = REML
/PRINT = SOLUTION TESTCOV G R
/FIXED = catvar1 catvar2 contvar catvar1*catvar2 time*catvar1 time*catvar2
/RANDOM = intercept time | COVTYPE(UN) SUBJECT(IDvar)
/REPEATED = cattime | COVTYPE(DIAG) SUBJECT(IDvar)
/* Other optional commands would follow */
/EMMEANS TABLES(catvar1*catvar2) WITH(time=0) COMPARE(catvar1) ADJ(LSD)
/* Examples of TEST commands */
/TEST = "L vs. H Catvar1 for Main Effect" catvar1 -1 0 1
/TEST = "M vs. H Catvar1 for Main Effect" catvar1 0 -1 1
/TEST = "L vs. M Catvar1 for Main Effect" catvar1 -1 1 0
/TEST = "L vs. H Catvar1 for Linear Slope" time*catvar1 -1 0 1
/TEST = "M vs. H Catvar1 for Linear Slope" time*catvar1 0 -1 1
/TEST = "L vs. M Catvar1 for Linear Slope" time*catvar1 -1 1 0
/TEST = "Simple Effect of X if Z=3" xvar 1 xvar*zvar 3
/TEST = "Simple Effect of Z if X=3" zvar 1 xvar*zvar 3.

```

```

*****
*****          STATA MULTILEVEL MODELING SYNTAX          *****
*****

* GENERIC EXAMPLE SYNTAX FOR XT MIXED:
  xtmixed DV fixed effects, FE options || Level2ID: random effects, RE options ///
    variance reml/mle covariance(Gmatrixtype) residuals(Rmatrixtype, t(Level1ID)),
  estat ic n(#persons),

* Fixed effects (FE) options:
  * noconstant to remove fixed intercept (included by default)
  * i. indicates categorical predictors (reference is first by default)
  * c. indicates continuous predictors (default if not specified)
  * can fit interactions on the fly
    * c.age#c.age creates quadratic age slope
    * i.group#c.age creates group by age interaction
    * i.program##i.day creates program by day categorical interaction

* Random effects (RE) options:
  * noconstant to remove random intercept (included by default)
  * covariance(Unstructured) is for G matrix unstructured
  * estat recovariance --> display G matrix
    * options: , level(levelvar) correlation
    * levelvar says at what level, correlation prints GCORR
  * Can add another level as || Level3ID: random effects, RE options
  * Can add group predictors of random effects heterogeneity
    * gen: boyXage = boy*age
    * ID: boy boyXage, no constant --> separate int and age slope per gender
  * Add R. to indicate categorical variables as random effects
  * Add _all: instead of ID: to indicate no nesting
  * Can do crossed models
    * || _all: R.id || _all: R.week --> persons by weeks as crossed
    * More efficient and equivalent: || _all: R.week || id:

* Residual options:
  * R matrix via residual(form,options)
    * independent-->VC, exchangeable-->CS, unstructured, toeplitz #-->TOEPn
    * AR #, exponential --> AR for unbalanced time
  * option by(varname) allows heterogeneous residual variance
  * option t(varname) is level-1 ID variable (i.e., for time)

* General options:
  * mle for ML, reml for REML is default, variance asks for variances rather than SD
  * noretale for no random effects solution, nogroup for no table summarizing groups
  * noheader suppresses output header, estat ic prints AIC and BIC (#parms = #total parms)

*** Other options

* estimates --> Can store results and do LR test comparisons
* lrtest --> LR test for models listed that have been saved
  estimates store bigmodel, estimates store smallmodel
  lrtest bigmodel smallmodel

* predict --> predicted estimates, linear predictor and SE from fixed effects
  predict xb

* lincom --> point estimates and SEs for linear combinations (like ESTIMATE)
  lincom 1*xvar + 3*xvar*zvar // Simple effect of X if Z=3
  lincom 1*zvar + 3*xvar*zvar // Simple effect of Z if X=3

* margins --> marginal means (LSMEANS)

* Estimating group means at first and last occasions
  margins ib(last).catvar1, at(c.time=(0) c.timesq=(0 ))
  margins ib(last).catvar1, at(c.time=(5) c.timesq=(25))

* test --> Wald test of simple and composite linear hypotheses

* Example contrasts between groups on intercept and linear time slopes
  test 1.catvar1=3.catvar1 // Low vs. High: Intercept
  test 2.catvar1=3.catvar1 // Med vs. High: Intercept
  test 1.catvar1=2.catvar1 // Low vs. Med: Intercept
  test 1.catvar1#time=3.catvar1#time // Low vs. High: Linear Slope
  test 2.catvar1#time=3.catvar1#time // Med vs. High: Linear Slope
  test 1.catvar1#time=2.catvar1#time // Low vs. Med: Linear Slope

```

Appendix C. Chi-square values from regular and mixture distributions.

Critical Values for Regular Chi-Square Distribution

| df | Significance Level | | | | |
|----|--------------------|--------|--------|--------|--------|
| | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
| 1 | 2.706 | 3.842 | 5.024 | 6.635 | 7.879 |
| 2 | 4.605 | 5.992 | 7.378 | 9.210 | 10.597 |
| 3 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 |
| 4 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 |
| 5 | 9.236 | 11.071 | 12.833 | 15.086 | 16.750 |
| 6 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 |
| 7 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 |
| 8 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 |
| 9 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 |
| 10 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 |
| 11 | 17.275 | 19.675 | 21.920 | 24.725 | 26.757 |

A critical value of .05 is recommended when comparing models differing in fixed effects.

A critical value of .10 is recommended when comparing models differing in random intercepts.

Critical Values for 50:50 Mixture of Chi-Square Distributions

| df (q) | Significance Level | | | | |
|-----------|--------------------|-------|-------|-------|-------|
| | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
| 0 vs. 1 | 1.64 | 2.71 | 3.84 | 5.41 | 6.63 |
| 1 vs. 2 | 3.81 | 5.14 | 6.48 | 8.27 | 9.63 |
| 2 vs. 3 | 5.53 | 7.05 | 8.54 | 10.50 | 11.97 |
| 3 vs. 4 | 7.09 | 8.76 | 10.38 | 12.48 | 14.04 |
| 4 vs. 5 | 8.57 | 10.37 | 12.10 | 14.32 | 15.97 |
| 5 vs. 6 | 10.00 | 11.91 | 13.74 | 16.07 | 17.79 |
| 6 vs. 7 | 11.38 | 13.40 | 15.32 | 17.76 | 19.54 |
| 7 vs. 8 | 12.74 | 14.85 | 16.86 | 19.38 | 21.23 |
| 8 vs. 9 | 14.07 | 16.27 | 18.35 | 20.97 | 22.88 |
| 9 vs. 10 | 15.38 | 17.67 | 19.82 | 22.52 | 24.49 |
| 10 vs. 11 | 16.67 | 19.04 | 21.27 | 24.05 | 26.07 |

Critical values such that the right-hand tail probability = $0.5 \times \Pr(\chi^2_q > c) + 0.5 \times \Pr(\chi^2_{q+1} > c)$

Source: Appendix C (p. 484) from:

Fitzmaurice, Laird, & Ware (2004). *Applied Longitudinal Analysis*. Hoboken, NJ: Wiley

Example 2: Unconditional Polynomial Models for Change in Number Match 3 Response Time (complete data, syntax, and output available for SAS, SPSS, and STATA electronically)

These data (in "Example23" data files) come from a short-term longitudinal study of 6 observations over 2 weeks for 101 adults age 65–80. The goal is to see how performance on this processing speed task ("number match 3"), as measured by response time in milliseconds, declines over the 6 practice sessions.

SAS Code for Data Manipulation:

```
* SAS code to import data, center time for polynomial models;
DATA work.example23; SET filepath.example23;
    clsess = session - 1; LABEL clsess = "clsess: Session Centered at 1";
RUN;
```

SPSS Code for Data Manipulation:

```
* SPSS code to import data, center time for polynomial models.
GET FILE = "example/Example23.sav".
DATASET NAME example23 WINDOW=FRONT.
COMPUTE clsess = session - 1.
VARIABLE LABELS clsess "clsess: Session Centered at 1".
```

STATA Code for Data Manipulation:

```
* STATA code to center time for polynomial models (and make quadratic version)
gen clsess = session - 1
gen clsess2 = clsess * clsess
label variable clsess "clsess: Session Centered at 1"
label variable clsess2 "clsess2: Quadratic Session Centered at 1"
```

Model 1a. Most Conservative Baseline: Empty Means, Random Intercept

$$\text{Level 1: } y_{ti} = \beta_{0i} + e_{ti}$$

$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$

TITLE1 "SAS Model 1a: Empty Means, Random Intercept Only";

```
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT
    COVTEST NAMELEN=100 METHOD=REML;
    CLASS ID session;
    MODEL nm3rt = / SOLUTION DDFM=Satterthwaite;
    RANDOM INTERCEPT / G V V CORR TYPE=UN SUBJECT=ID;
    REPEATED session / R TYPE=VC SUBJECT=ID; RUN;
```

METHOD = ML or REML (default)
CLASS = categorical predictors, nesting
MODEL dv = fixed effects / print solution
RANDOM = person variances in **G**
REPEATED = residuals in **R** matrix

TITLE "SPSS Model 1a: Empty Means, Random Intercept".

```
MIXED nm3rt BY ID session
    /METHOD = REML
    /PRINT = SOLUTION TESTCOV G R
    /FIXED =
    /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
    /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

MIXED dv BY categorical predictors
WITH continuous predictors
/METHOD = REML or ML
/PRINT = regression solution
/FIXED = predictors for means model
/RANDOM = person variances in **G**

* STATA Model 1a: Empty Means, Random Intercept

```
xtmixed nm3rt , || id: , ///
    variance reml covariance(unstructured) residuals(independent,t(session)),
    estat ic, n(101)
    estat recovariance, level(id)
```

DV = nm3rt, random part after ||
Level 2 ID is PersonID, random intercept by default
Print variances instead of SD, use reml
covariance(unstructured) refers to **G** matrix
residuals(independent) → refers to **R** matrix by session
estat ic → Print IC given N = 101 persons

STATA output:

```

Mixed-effects REML regression              Number of obs   =       606
Group variable: id                        Number of groups  =       101

                                         Obs per group: min =        6
                                         avg   =       6.0
                                         max   =        6

                                         Wald chi2(0)      =        .
                                         Prob > chi2       =        .

```

NOTE: LL is given rather than -2LL

Log restricted-likelihood = -4268.4304

```

-----+-----
      nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _cons |    1770.701   45.42063   38.98   0.000    1681.679    1859.724
-----+-----

```

This is the fixed intercept
(just grand mean so far).

```

-----+-----
Random-effects Parameters |   Estimate  Std. Err.      [95% Conf. Interval]
-----+-----
id: Identity
      var(_cons) |       200883   29471.23     150683.2     267806.8
-----+-----
      var(Residual) |    44899.96   2825.63     39689.76     50794.13
-----+-----

```

LR test vs. linear regression: chibar2(01) = 691.74 Prob >= chibar2 = 0.0000

. estat ic, n(101)

```

-----+-----
      Model |   Obs   ll(null)   ll(model)    df       AIC       BIC
-----+-----
      .    |   101         .   -4268.43      3   8542.861   8550.706
-----+-----

```

Note: N=101 used in calculating BIC

```

.      estat recovariance, level(id)
Random-effects covariance matrix for level id

```

```

-----+-----
      |      _cons
-----+-----
      |
      _cons |      200883
-----+-----

```

This is the level-2 **G** matrix, just a
random intercept variance so far.

**Calculate the ICC for the
Number Match 3 outcome:**

$$ICC = \frac{200883}{200883 + 44900} = .82$$

This LR test tells us that the
random intercept variance is
significantly greater than 0,
and thus so is the ICC.

REML-based AIC and BIC are
calculated differently in
STATA (they count fixed
effects), so they won't match
the values in other programs.

Extra SAS output not provided by STATA:**Estimated R Matrix for ID 101**

```

Row      Col1      Col2      Col3      Col4      Col5      Col6
  1      44900
  2
  3
  4
  5
  6

```

This level-1 **R** matrix (with
equal variance over time, no
covariance of any kind, known
as VC or independence) will
be used repeatedly as we add
fixed and random effects.

Estimated G Matrix

```

Participant
Row      Effect      ID      Col1
  1      Intercept      101      200883

```

This is the level-2 **G** matrix, just a
random intercept variance so far.

Estimated V Matrix for ID 101

```

Row      Col1      Col2      Col3      Col4      Col5      Col6
  1      245783      200883      200883      200883      200883      200883
  2      200883      245783      200883      200883      200883      200883
  3      200883      200883      245783      200883      200883      200883
  4      200883      200883      200883      245783      200883      200883
  5      200883      200883      200883      200883      245783      200883
  6      200883      200883      200883      200883      200883      245783

```

The **V** matrix is the total
variance-covariance matrix
after combining the level-2
G and level-1 **R** matrices.

Model 1b. Most Liberal Baseline – Saturated Means, Unstructured Variances (Model Answer Key)

```

TITLE1 "SAS Model 1b: Saturated Means, Unstructured Variances";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS ID session;
  MODEL nm3rt = session / SOLUTION DDFM=Satterthwaite;
  REPEATED session / R RCORR TYPE=UN SUBJECT=ID;
  LSMEANS session /; RUN;

```

Placing *session* on the CLASS/BY statements and in the FIXED/MODEL statements treats it as a categorical predictor. So this is an ANOVA means model. No RANDOM statements mean no random effects.

```

TITLE "SPSS Model 1b: Saturated Means, Unstructured Variances".
MIXED nm3rt BY ID session
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV R
  /FIXED = session
  /REPEATED = session | SUBJECT(ID) COVTYPE(UN)
  /EMMEANS = TABLES(session).

```

```

* STATA Model 1b: Saturated Means, Unstructured Variances
xtmixed nm3rt ib(last).session, || id: , noconstant ///
  variance reml residuals(unstructured, t(session)),
  estat ic, n(101),
  contrast session,          // omnibus test of mean differences
  margins i.session,         // observed means per session
  marginsplot name(observed_means, replace) // plot observed means

```

i. indicates categorical predictor of *session* (ref=last to match others) noconstant = no random intercept (just **R** matrix)

STATA output:

```

Mixed-effects REML regression
Group variable: id

```

| | | |
|--------------------|---|--------|
| Number of obs | = | 606 |
| Number of groups | = | 101 |
| Obs per group: min | = | 6 |
| avg | = | 6.0 |
| max | = | 6 |
| Wald chi2(5) | = | 83.60 |
| Prob > chi2 | = | 0.0000 |

Log restricted-likelihood = -4114.8942

This is the multivariate Wald test for all the fixed effects simultaneously (5 mean differences from the fixed intercept here).

| nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|-------------|----------|-----------|-------|-------|----------------------|----------|
| -----+----- | | | | | | |
| session | | | | | | |
| 1 | 289.7574 | 32.69997 | 8.86 | 0.000 | 225.6666 | 353.8481 |
| 2 | 143.0364 | 26.20308 | 5.46 | 0.000 | 91.67927 | 194.3935 |
| 3 | 77.89864 | 22.8842 | 3.40 | 0.001 | 33.04642 | 122.7509 |
| 4 | 45.66045 | 20.78533 | 2.20 | 0.028 | 4.921952 | 86.39894 |
| 5 | 35.03972 | 18.11681 | 1.93 | 0.053 | -.468579 | 70.54802 |
| -----+----- | | | | | | |
| _cons | 1672.136 | 44.13439 | 37.89 | 0.000 | 1585.634 | 1758.638 |

Mean diffs relative to session 6

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|----------|-----------|----------------------|----------|
| -----+----- | | | | |
| id: | (empty) | | | |
| -----+----- | | | | |
| Residual: Unstructured | | | | |
| var(e1) | 301983.1 | 42696.65 | 228893.4 | 398411.6 |
| var(e2) | 259148.8 | 36635.7 | 196433.4 | 341887.4 |
| var(e3) | 233366.9 | 32990.48 | 176891.6 | 307872.9 |
| var(e4) | 217542.8 | 30753.82 | 164896.4 | 286997.6 |
| var(e5) | 212096.8 | 29984.63 | 160767.3 | 279814.7 |
| var(e6) | 196732.3 | 27812.21 | 149121.6 | 259543.9 |
| cov(e1,e2) | 235657.1 | 36563.79 | 163993.4 | 307320.8 |

These are the total variances at each occasion...

```

cov(e1,e3) | 217992.5 34336.3 150694.6 285290.4
cov(e1,e4) | 202605.4 32657.69 138597.5 266613.2
cov(e1,e5) | 192152.4 31762.13 129899.7 254405
cov(e1,e6) | 195358.7 31224.07 134160.6 256556.7
cov(e2,e3) | 230215.6 33672.52 164218.7 296212.5
cov(e2,e4) | 213230.6 31899.27 150709.2 275752
cov(e2,e5) | 202091 30938.65 141452.3 262729.6
cov(e2,e6) | 193267.2 29707.66 135041.2 251493.1
cov(e3,e4) | 205208 30462.92 145501.8 264914.2
cov(e3,e5) | 196917.7 29697.89 138710.9 255124.5
cov(e3,e6) | 188603.5 28532.39 132681 244525.9
cov(e4,e5) | 193674.7 28910.48 137011.2 250338.2
cov(e4,e6) | 185320 27762.64 130906.2 239733.8
cov(e5,e6) | 187839.5 27739.82 133470.5 242208.6

```

And these are
the total
covariances
across
occasions...

```

-----
LR test vs. linear regression:      chi2(20) = 925.64   Prob > chi2 = 0.0000

```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

This is the LRT of whether the
unstructured R model fits better
than the e-only R model...

```

.      estat ic, n(101),
-----
Model | Obs   ll(null)   ll(model)   df       AIC       BIC
-----+-----
. | 101          . -4114.894    27   8283.788   8354.397
-----

```

Note: N=101 used in calculating BIC

```

.      contrast session, // omnibus test of mean differences

```

Contrasts of marginal linear predictions

Margins : asbalanced

```

-----
          |      df      chi2      P>chi2
-----+-----
nm3rt     |
session   |      5      83.60      0.0000
-----

```

This is the omnibus test of mean
differences across 6 sessions.

```

.      margins i.session, // observed means per session

```

Adjusted predictions Number of obs = 606

Expression : Linear prediction, fixed portion, predict()

THESE ARE THE SATURATED MEANS THE FIXED EFFECTS WILL BE TRYING TO REPRODUCE.

```

-----
          |      Delta-method
          |      Margin   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
session |
1 | 1961.893   54.68027   35.88   0.000   1854.722   2069.065
2 | 1815.172   50.65402   35.83   0.000   1715.892   1914.452
3 | 1750.035   48.06832   36.41   0.000   1655.822   1844.247
4 | 1717.796   46.41001   37.01   0.000   1626.835   1808.758
5 | 1707.176   45.82541   37.25   0.000   1617.36   1796.992
6 | 1672.136   44.13439   37.89   0.000   1585.634   1758.638
-----

```


Extra SAS output not provided by STATA:

Estimated R Matrix for ID 101

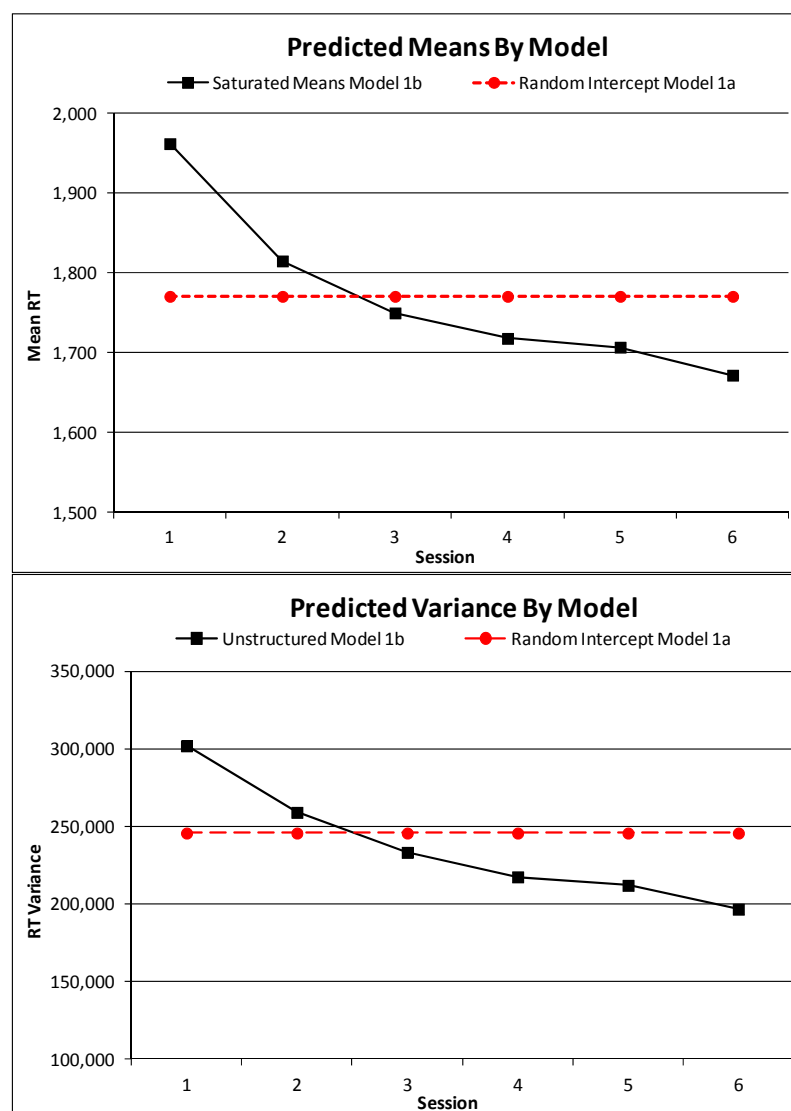
| Row | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 |
|-----|---------------|---------------|---------------|---------------|---------------|---------------|
| 1 | 301985 | 235659 | 217994 | 202607 | 192154 | 195360 |
| 2 | 235659 | 259150 | 230217 | 213232 | 202092 | 193268 |
| 3 | 217994 | 230217 | 233368 | 205209 | 196919 | 188604 |
| 4 | 202607 | 213232 | 205209 | 217544 | 193676 | 185321 |
| 5 | 192154 | 202092 | 196919 | 193676 | 212098 | 187840 |
| 6 | 195360 | 193268 | 188604 | 185321 | 187840 | 196733 |

This Unstructured **R** matrix estimates all variances and covariances separately. THIS IS THE DATA we are trying to duplicate with our model for the variances.

Estimated R Correlation Matrix for ID 101

| Row | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 |
|-----|--------|--------|--------|--------|--------|--------|
| 1 | 1.0000 | 0.8424 | 0.8212 | 0.7905 | 0.7593 | 0.8015 |
| 2 | 0.8424 | 1.0000 | 0.9361 | 0.8981 | 0.8620 | 0.8559 |
| 3 | 0.8212 | 0.9361 | 1.0000 | 0.9108 | 0.8851 | 0.8802 |
| 4 | 0.7905 | 0.8981 | 0.9108 | 1.0000 | 0.9016 | 0.8958 |
| 5 | 0.7593 | 0.8620 | 0.8851 | 0.9016 | 1.0000 | 0.9196 |
| 6 | 0.8015 | 0.8559 | 0.8802 | 0.8958 | 0.9196 | 1.0000 |

So here is what are we trying to model—means and variances, where model 1b is the data:



Model 2a. Fixed Linear Time, Random Intercept

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + e_{ti}$$

$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$

$$\text{Linear Session: } \beta_{1i} = \gamma_{10}$$

```
TITLE1 "SAS Model 2a: Fixed Linear Time, Random Intercept";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS ID session;
  MODEL nm3rt = c1sess / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / G V VCORR TYPE=UN SUBJECT=ID;
  REPEATED session / R TYPE=VC SUBJECT=ID; RUN;
```

The predictor of *c1sess* will be treated as continuous given that it is not on the CLASS statement (SAS) and it is on WITH (SPSS).

```
TITLE "SPSS Model 2a: Fixed Linear Time, Random Intercept".
MIXED nm3rt BY ID session WITH c1sess
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV G R
  /FIXED = c1sess
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

```
* STATA Model 2a: Fixed Linear Time, Random Intercept
xtmixed nm3rt c.c1sess, || id: , ///
  variance reml covariance(un) residuals(independent,t(session)),
  estat ic, n(101),
  estimates store FixLin
```

DV = nm3rt, c. means continuous fixed slope for *c1sess*
 Level 2 ID is id, random intercept by default
 estimates → save results as "FixLin" for next LRT

STATA output:

```
Mixed-effects REML regression
Group variable: id

Number of obs      =      606
Number of groups   =      101
Obs per group: min =         6
                  avg =        6.0
                  max =         6

Wald chi2(1)       =     131.82
Prob > chi2        =      0.0000

Log restricted-likelihood = -4207.344
```

| nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|--------|-----------|-----------|--------|-------|----------------------|
| c1sess | -51.57185 | 4.491815 | -11.48 | 0.000 | -60.37565 -42.76806 |
| _cons | 1899.631 | 46.7882 | 40.60 | 0.000 | 1807.928 1991.334 |

The fixed linear effect of *c1sess* is significant according to the Wald test (*p*-value for fixed effect).

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|----------|-----------|----------------------|
| id: Identity | | | |
| var(_cons) | 202422.7 | 29469.85 | 152172.6 269266.3 |
| var(Residual) | 35661.79 | 2246.481 | 31519.73 40348.16 |

Relative to the empty means, random intercept model 1a, the fixed linear effect of session explained ~21% of the residual variance (which made the random intercept variance increase due to its residual variance correction factor).

```
LR test vs. linear regression: chibar2(01) = 787.61 Prob >= chibar2 = 0.0000
```

```
. estat ic, n(101),

Model | Obs   ll(null)   ll(model)   df       AIC       BIC
-----+-----
. | 101           . -4207.344     4   8422.688   8433.149
```

Note: N=101 used in calculating BIC

Model 2b. Random Linear Time

```

TITLE1 "SAS Model 2b: Random Linear Time";
PROC MIXED DATA=work.example23 NOCLPRINT
NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
    CLASS ID session;
    MODEL nm3rt = c1sess / SOLUTION DDFM=Satterthwaite;
    RANDOM INTERCEPT c1sess / G V VCORR TYPE=UN SUBJECT=ID;
    REPEATED session / R TYPE=VC SUBJECT=ID; RUN;

```

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i} (\text{Session}_{ti} - 1) + e_{ti}$$

$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$

$$\text{Linear Session: } \beta_{1i} = \gamma_{10} + U_{1i}$$

```

TITLE "SPSS Model 2b: Random Linear Time".
MIXED nm3rt BY ID session WITH c1sess
    /METHOD = REML
    /PRINT = SOLUTION TESTCOV G R
    /FIXED = c1sess
    /RANDOM = INTERCEPT c1sess | SUBJECT(ID) COVTYPE(UN)
    /REPEATED = session | SUBJECT(ID) COVTYPE(ID).

```

Now there are 2 random effects: intercept and linear slope, given by c1sess on the RANDOM statements.

```

* STATA Model 2b: Random Linear Time
xtmixed nm3rt c.c1sess, || id: c1sess, ///
    variance reml covariance(un) residuals(independent,t(session)),
    estat ic, n(101),
    estat recovariance, level(id),
    estimates store RandLin,
    lrtest RandLin FixLin

```

DV = nm3rt, c. means continuous fixed slope for c1sess
 Level 2 ID is id, random intercept and c1sess now
 estimates → save results as "RandLin" for LRT

STATA output:

```

Mixed-effects REML regression
Group variable: id
Number of obs      =      606
Number of groups   =      101
Obs per group: min =        6
                  avg =       6.0
                  max =        6

```

```

Log restricted-likelihood = -4186.0512
Wald chi2(1)           =      70.17
Prob > chi2            =      0.0000

```

| nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|--------|-----------|-----------|-------|-------|----------------------|
| c1sess | -51.57185 | 6.156722 | -8.38 | 0.000 | -63.63881 -39.5049 |
| _cons | 1899.631 | 51.4998 | 36.89 | 0.000 | 1798.693 2000.569 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|-----------|-----------|----------------------|
| id: Unstructured | | | |
| var(c1sess) | 2233.833 | 552.9239 | 1375.178 3628.626 |
| var(_cons) | 253258 | 37897.26 | 188881.9 339575.3 |
| cov(c1sess,_cons) | -12700.79 | 3621.977 | -19799.74 -5601.848 |
| var(Residual) | 27905.42 | 1963.419 | 24310.74 32031.62 |

```

LR test vs. linear regression:      chi2(3) = 830.20  Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.

```

```

.      estat ic, n(101),

```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 101 | . | -4186.051 | 6 | 8384.102 | 8399.793 |

Note: N=101 used in calculating BIC

```
.          estat recovariance, level(id),

Random-effects covariance matrix for level id
              |      c1sess      _cons
-----+-----
c1sess |      2233.833
_cons |     -12700.79      253258

.          estimates store RandLin,
.          lrtest RandLin FixLin
```

Is the random linear time model (2b) better than the fixed linear time, random intercept model (2a)?

Yep, $-2\Delta LL = 43$, which is bigger than the critical value of 5.99ish on $df \approx 2$ ish

```
Likelihood-ratio test                LR chi2(2)  =      42.59
(Assumption: FixLin nested in RandLin) Prob > chi2 =      0.0000
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative. Note: LR tests based on REML are valid only when the fixed-effects specification is identical for both models.

Extra SAS output not provided by STATA:

```
Estimated R Matrix for ID 101
Row      Col1      Col2      Col3      Col4      Col5      Col6
1      27905
2
3
4
5
6
```

```
Estimated G Matrix
Participant
Row      Effect      ID      Col1      Col2
1      Intercept      101      253258      -12701
2      C1sess          101      -12701      2233.83
```

```
Estimated V Matrix for ID 101
Row      Col1      Col2      Col3      Col4      Col5      Col6
1      281163      240557      227856      215155      202455      189754
2      240557      257995      219623      209156      198689      188222
3      227856      219623      239295      203157      194924      186691
4      215155      209156      203157      225063      191158      185159
5      202455      198689      194924      191158      215298      183627
6      189754      188222      186691      185159      183627      210001
```

The **V** matrix is the total variance-covariance matrix after combining the level-2 **G** and level-1 **R** matrices. Now the variances and covariances are predicted to change based on time.

```
Estimated V Correlation Matrix for ID 101
Row      Col1      Col2      Col3      Col4      Col5      Col6
1      1.0000      0.8932      0.8784      0.8553      0.8229      0.7809
2      0.8932      1.0000      0.8839      0.8680      0.8430      0.8086
3      0.8784      0.8839      1.0000      0.8754      0.8588      0.8328
4      0.8553      0.8680      0.8754      1.0000      0.8684      0.8517
5      0.8229      0.8430      0.8588      0.8684      1.0000      0.8636
6      0.7809      0.8086      0.8328      0.8517      0.8636      1.0000
```

The **VCORR** matrix is the correlation version. The ICC is now predicted to change over time, too (and conditional on linear time).

How the V matrix variances and covariances get calculated in a random linear time model:

$$\mathbf{V}_i \text{ matrix: Variance}[y_{\text{time}}] = \tau_{U_0}^2 + \left[(\text{Session} - 1)^2 \tau_{U_1}^2 \right] + \left[2(\text{Session} - 1) \tau_{U_{01}} \right] + \sigma_e^2$$

$$\mathbf{V}_i \text{ matrix: Covariance}[y_A, y_B] = \tau_{U_0}^2 + \left[(A + B) \tau_{U_{01}} \right] + \left[(AB) \tau_{U_1}^2 \right]$$

Model 3a. Fixed Quadratic, Random Linear Time

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$

Level 2: Intercept: $\beta_{0i} = \gamma_{00} + U_{0i}$

Linear Session: $\beta_{1i} = \gamma_{10} + U_{1i}$

Quadratic Session: $\beta_{2i} = \gamma_{20}$

```
TITLE1 "SAS Model 3a: Fixed Quadratic, Random Linear Time";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS ID session;
  MODEL nm3rt = c1sess c1sess*c1sess / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT c1sess / G V V CORR TYPE=UN SUBJECT=ID;
  REPEATED session / R TYPE=VC SUBJECT=ID; RUN;
```

```
TITLE "SPSS Model 3a: Fixed Quadratic, Random Linear Time".
MIXED nm3rt BY ID session WITH c1sess
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV G R
  /FIXED = c1sess c1sess*c1sess
  /RANDOM = INTERCEPT c1sess | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

Interactions can be defined on the fly in SAS and SPSS using *, or in STATA using # (but only for fixed effects in STATA).

```
* STATA Model 3a: Fixed Quadratic, Random Linear Time
xtmixed nm3rt c.c1sess c.c1sess#c.c1sess, || id: c1sess, ///
  variance reml covariance(un) residuals(independent,t(session)),
  estat ic, n(101),
  estat recovariance, level(id),
  estimates store FixQuad
```

STATA output:

```
Mixed-effects REML regression
Group variable: id
Number of obs      =      606
Number of groups   =      101
Obs per group: min =         6
                  avg =        6.0
                  max =         6
Wald chi2(2)       =      97.86
Prob > chi2        =      0.0000
Log restricted-likelihood = -4170.7386
```

| nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-------------------|-----------|-----------|-------|-------|----------------------|
| c1sess | -120.8999 | 14.54147 | -8.31 | 0.000 | -149.4007 -92.39917 |
| c.c1sess#c.c1sess | 13.86561 | 2.634761 | 5.26 | 0.000 | 8.701578 19.02965 |
| _cons | 1945.85 | 52.2433 | 37.25 | 0.000 | 1843.455 2048.245 |

The fixed quadratic effect of c1sess is significant according to the Wald test (p -value for fixed effect).

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|-----------|-----------|----------------------|
| id: Unstructured | | | |
| var(c1sess) | 2332.667 | 551.5799 | 1467.501 3707.891 |
| var(_cons) | 254164 | 37895.62 | 189758.3 340429.7 |
| cov(c1sess,_cons) | -12947.88 | 3620.697 | -20044.31 -5851.442 |
| var(Residual) | 26175.83 | 1844.008 | 22800.05 30051.42 |

Relative to the random linear time model 2b, the fixed quadratic effect of session explained another ~6% of the residual variance (which made the random intercept variance increase due to its residual variance correction factor).

LR test vs. linear regression: $\chi^2(3) = 851.78$ Prob > $\chi^2 = 0.0000$

Note: LR test is conservative and provided only for reference.

```
.      estat ic, n(101),
```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 101 | . | -4170.739 | 7 | 8355.477 | 8373.783 |

Note: N=101 used in calculating BIC

Model 3b. Random Quadratic Time (and an example of ESTIMATE/TEST/MARGINS statements)

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$$

$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$

$$\text{Linear Session: } \beta_{1i} = \gamma_{10} + U_{1i}$$

$$\text{Quadratic Session: } \beta_{2i} = \gamma_{20} + U_{2i}$$

```
TITLE1 "SAS Model 3b: Random Quadratic Time";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS ID session;
  MODEL nm3rt = clsess clsess*clsess / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT clsess clsess*clsess / G V V CORR TYPE=UN SUBJECT=ID;
  REPEATED session / R TYPE=VC SUBJECT=ID;
  ESTIMATE "Intercept at Session 1"    intercept 1 clsess 0    clsess*clsess 0;
  ESTIMATE "Intercept at Session 2"    intercept 1 clsess 1    clsess*clsess 1;
  ESTIMATE "Intercept at Session 3"    intercept 1 clsess 2    clsess*clsess 4;
  ESTIMATE "Intercept at Session 4"    intercept 1 clsess 3    clsess*clsess 9;
  ESTIMATE "Intercept at Session 5"    intercept 1 clsess 4    clsess*clsess 16;
  ESTIMATE "Intercept at Session 6"    intercept 1 clsess 5    clsess*clsess 25;
  * Predicting linear rate of change at each session (linear changes by 2*quad);
  ESTIMATE "Linear Slope at Session 1"  clsess 1    clsess*clsess 0;
  ESTIMATE "Linear Slope at Session 2"  clsess 1    clsess*clsess 2;
  ESTIMATE "Linear Slope at Session 3"  clsess 1    clsess*clsess 4;
  ESTIMATE "Linear Slope at Session 4"  clsess 1    clsess*clsess 6;
  ESTIMATE "Linear Slope at Session 5"  clsess 1    clsess*clsess 8;
  ESTIMATE "Linear Slope at Session 6"  clsess 1    clsess*clsess 10; RUN;

TITLE "SPSS Model 3b: Random Quadratic Time".
MIXED nm3rt BY ID session WITH clsess
  /METHOD = REML
  /PRINT  = SOLUTION TESTCOV G R
  /FIXED  = clsess clsess*clsess
  /RANDOM  = INTERCEPT clsess clsess*clsess | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID)
  /TEST = "Intercept at Session 1"    intercept 1 clsess 0    clsess*clsess 0
  /TEST = "Intercept at Session 2"    intercept 1 clsess 1    clsess*clsess 1
  /TEST = "Intercept at Session 3"    intercept 1 clsess 2    clsess*clsess 4
  /TEST = "Intercept at Session 4"    intercept 1 clsess 3    clsess*clsess 9
  /TEST = "Intercept at Session 5"    intercept 1 clsess 4    clsess*clsess 16
  /TEST = "Intercept at Session 6"    intercept 1 clsess 5    clsess*clsess 25
  /TEST = "Linear Slope at Session 1"  clsess 1    clsess*clsess 0
  /TEST = "Linear Slope at Session 2"  clsess 1    clsess*clsess 2
  /TEST = "Linear Slope at Session 3"  clsess 1    clsess*clsess 4
  /TEST = "Linear Slope at Session 4"  clsess 1    clsess*clsess 6
  /TEST = "Linear Slope at Session 5"  clsess 1    clsess*clsess 8
  /TEST = "Linear Slope at Session 6"  clsess 1    clsess*clsess 10.
```

Because twice the quadratic slope is how the linear slope changes per unit time, the value for *clsess* used in estimating the linear slope per session gets multiplied by 2.

*** STATA Model 3b: Random Quadratic Time**

```
xtmixed nm3rt c.c1sess c.c1sess#c.c1sess, || id: c1sess c1sess2, ///
    variance reml covariance(un) residuals(independent,t(session)),
    estat ic, n(101),
    estat recovariance, level(id),
    estimates store RandQuad,
    lrtest RandQuad FixQuad,
    margins, at(c.c1sess=(0(1)5)) vsquish // intercepts per session
    marginsplot, name(predicted_means, replace) // plot intercepts
    margins, at(c.c1sess=(0(1)5)) dydx(c.c1sess) vsquish // linear slope per session
    marginsplot, name(change_in_linear_slope, replace) // plot quadratic effect
```

The random statement will not accept interaction terms, so we are using the c1sess2 created manually before.

STATA output:

```
Mixed-effects REML regression
Group variable: id
Number of obs      =      606
Number of groups   =      101
Obs per group: min =         6
                  avg =        6.0
                  max =         6
Wald chi2(2)       =      71.74
Prob > chi2        =      0.0000
Log restricted-likelihood = -4151.3728
```

| | nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-------------------|-------|-----------|-----------|-------|-------|----------------------|
| c1sess | | -120.8999 | 20.04752 | -6.03 | 0.000 | -160.1923 -81.6075 |
| c.c1sess#c.c1sess | | 13.86561 | 3.41541 | 4.06 | 0.000 | 7.171534 20.55969 |
| _cons | | 1945.85 | 53.84993 | 36.13 | 0.000 | 1840.306 2051.394 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|-----------|-----------|----------------------|
| id: Unstructured | | | |
| var(c1sess) | 25839.79 | 5864.685 | 16561.42 40316.29 |
| var(c1sess2) | 634.4659 | 172.375 | 372.5198 1080.605 |
| var(_cons) | 276207.8 | 41445.59 | 205831.2 370647.1 |
| cov(c1sess,c1sess2) | -3903.291 | 982.6248 | -5829.2 -1977.381 |
| cov(c1sess,_cons) | -35734.05 | 11947.96 | -59151.62 -12316.48 |
| cov(c1sess2,_cons) | 3901.974 | 1950.304 | 79.44722 7724.5 |
| var(Residual) | 20298.19 | 1649.117 | 17310.19 23801.96 |

```
LR test vs. linear regression:      chi2(6) = 890.51 Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

```
.      estat ic, n(101),
-----+-----
Model | Obs   ll(null)   ll(model)   df       AIC       BIC
-----+-----
.      | 101      .      -4151.373    10      8322.746    8348.897
```

Note: N=101 used in calculating BIC

```
.      estimates store RandQuad,
.      lrtest RandQuad FixQuad,
```

```
Likelihood-ratio test
(Assumption: FixQuad nested in RandQuad)
```

Is the random quadratic model (3b) better than the fixed quadratic, random linear model (3a)?

Yep, $-2\Delta LL = 39$, which is bigger than the critical value of 7.82ish on $df \sim 3$ ish

```
LR chi2(3) = 38.73
Prob > chi2 = 0.0000
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative. Note: LR tests based on REML are valid only when the fixed-effects specification is identical for both models.


```
. margins, at(c1sess=(0(1)5)) vsquish // intercepts per session
Adjusted predictions      Number of obs   =      606
Expression   : Linear prediction, fixed portion, predict()
1._at       : c1sess      =      0
2._at       : c1sess      =      1
3._at       : c1sess      =      2
4._at       : c1sess      =      3
5._at       : c1sess      =      4
6._at       : c1sess      =      5
```

These are the quadratic-model-predicted means per session.

| | | Delta-method | | | | |
|-----|--|--------------|-----------|-------|-------|----------------------|
| | | Margin | Std. Err. | z | P> z | [95% Conf. Interval] |
| _at | | | | | | |
| 1 | | 1945.85 | 53.84993 | 36.13 | 0.000 | 1840.306 2051.394 |
| 2 | | 1838.815 | 48.48658 | 37.92 | 0.000 | 1743.784 1933.847 |
| 3 | | 1759.512 | 46.99744 | 37.44 | 0.000 | 1667.399 1851.626 |
| 4 | | 1707.941 | 45.89598 | 37.21 | 0.000 | 1617.986 1797.895 |
| 5 | | 1684.1 | 44.23964 | 38.07 | 0.000 | 1597.392 1770.808 |
| 6 | | 1687.991 | 44.20394 | 38.19 | 0.000 | 1601.352 1774.629 |

```
. marginsplot, name(predicted_means, replace) // plot intercepts
Variables that uniquely identify margins: c1sess
. margins, at(c1sess=(0(1)5)) dydx(c1sess) vsquish // linear slope per session
Conditional marginal effects      Number of obs   =      606
```

```
Expression   : Linear prediction, fixed portion, predict()
dy/dx w.r.t. : c1sess
```

```
1._at       : c1sess      =      0
2._at       : c1sess      =      1
3._at       : c1sess      =      2
4._at       : c1sess      =      3
5._at       : c1sess      =      4
6._at       : c1sess      =      5
```

These are the instantaneous linear slopes at each session. Note how the SEs narrow towards the middle of the data.

| | | Delta-method | | | | |
|--------|-----|--------------|-----------|-------|-------|----------------------|
| | | dy/dx | Std. Err. | z | P> z | [95% Conf. Interval] |
| c1sess | _at | | | | | |
| | 1 | -120.8999 | 20.04752 | -6.03 | 0.000 | -160.1923 -81.6075 |
| | 2 | -93.1687 | 13.64968 | -6.83 | 0.000 | -119.9216 -66.4158 |
| | 3 | -65.43747 | 8.002796 | -8.18 | 0.000 | -81.12266 -49.75228 |
| | 4 | -37.70624 | 5.92417 | -6.36 | 0.000 | -49.3174 -26.09508 |
| | 5 | -9.975015 | 9.973315 | -1.00 | 0.317 | -29.52235 9.572324 |
| | 6 | 17.75621 | 16.03616 | 1.11 | 0.268 | -13.67408 49.18651 |

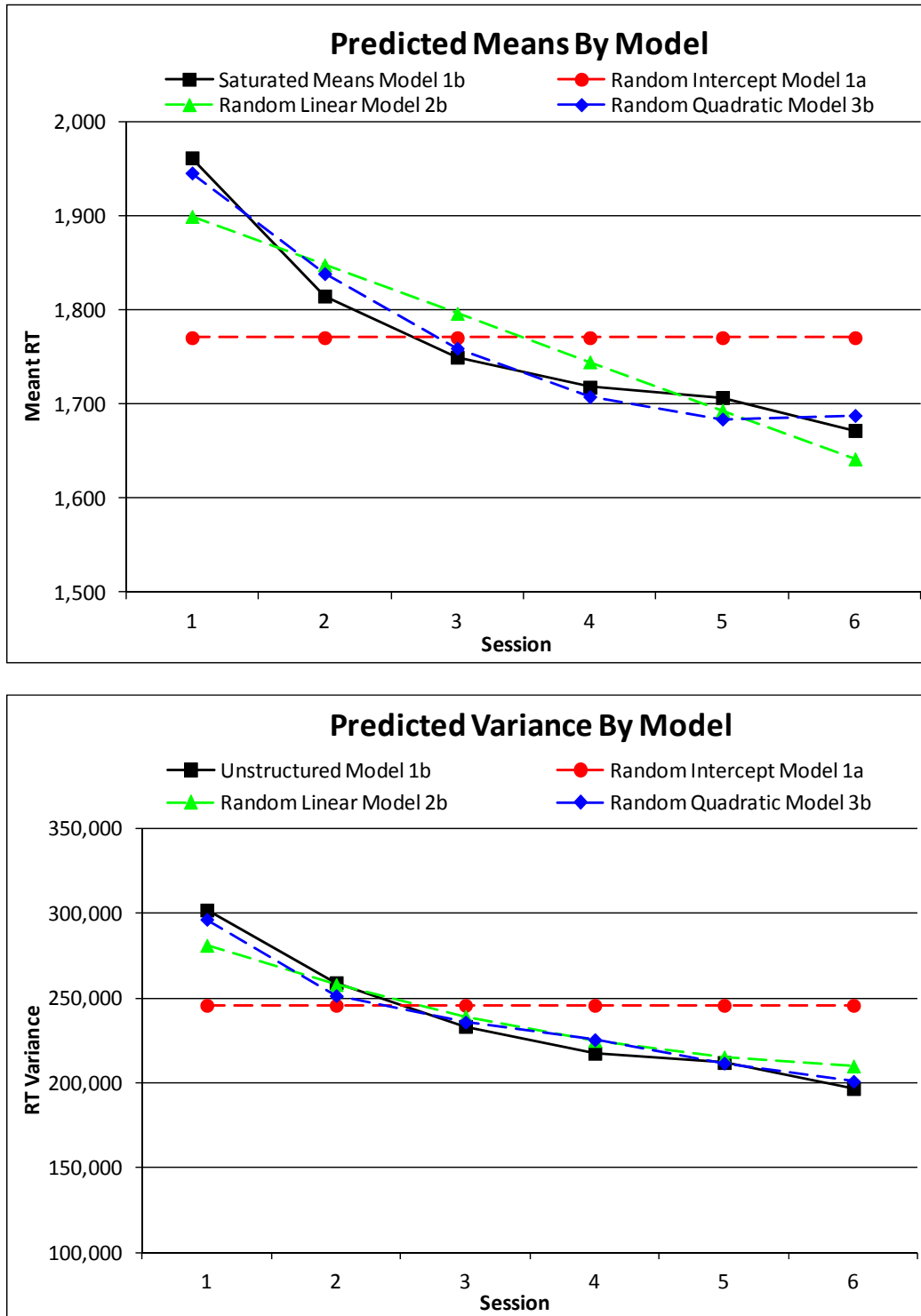
How well do the predicted means, variances, and covariances from the random quadratic model (3b) match the original means, variances, and covariances from the saturated means model (1b)?

Extra SAS output not provided by STATA:

| Estimated V Matrix for ID 101 | | | | | | |
|-------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Row | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 |
| 1 | 296504 | 244374 | 220346 | 204122 | 195702 | 195085 |
| 2 | 244374 | 251508 | 219312 | 208680 | 199315 | 191215 |
| 3 | 220346 | 219312 | 235842 | 209043 | 199808 | 187840 |
| 4 | 204122 | 208680 | 209043 | 225508 | 197182 | 184958 |
| 5 | 195702 | 199315 | 199808 | 197182 | 211735 | 182571 |
| 6 | 195085 | 191215 | 187840 | 184958 | 182571 | 200977 |

The **V** matrix is the total variance-covariance matrix after combining the level-2 **G** and level-1 **R** matrices. The variances and covariances are predicted to change based on time, but differently.

Figure 1



How the V matrix variances and covariances get calculated in a random quadratic time model:

Predicted Variance at Time T :

$$\text{Var}(y_T) = \sigma_e^2 + \tau_{U_0}^2 + 2*T*\tau_{U_{01}} + T^2*\tau_{U_1}^2 + 2*T^2*\tau_{U_{02}} + 2*T^3*\tau_{U_{12}} + T^4*\tau_{U_2}^2$$

Predicted Covariance between Time A and B :

$$\text{Cov}(y_A, y_B) = \tau_{U_0}^2 + (A+B)*\tau_{U_{01}} + (AB)*\tau_{U_1}^2 + (A^2+B^2)*\tau_{U_{02}} + (AB^2)+(A^2B)*\tau_{U_{12}} + (A^2B^2)*\tau_{U_2}^2$$

Simple Processing Speed – Example Unconditional Models of Change Results

Model Specification

Linear mixed models were estimated using restricted maximum likelihood (REML) in order to examine the overall pattern of and individual differences in response time over six sessions for a simple processing speed test (number match three). The significance of new fixed effects were evaluated using Wald tests, whereas the significance of new random effects was evaluated using likelihood ratio tests (i.e., $-2\Delta LL$), with degrees of freedom equal to the number of new random effects variances and covariances. The 95% confidence interval (CI) for random variation around each fixed effect was calculated as ± 1.96 standard deviations of its accompanying random variance term.

Although the six sessions were held over a period of 6–10 days, given that experience to the test (and not *time* per se) was the most likely reason for changes in response time, session was used as the metric of time (i.e., as opposed to age or day). Session was centered at the first occasion, such that the intercept represented initial status in all models. Observed mean response times (in milliseconds) estimated from a saturated means model (i.e., multivariate analysis of variance) are shown in Figure 1. The intraclass correlation from the unconditional means model (i.e., empty model; random intercept only) was calculated as .82, indicating that over 80% of the variance in number match 3 across sessions occurred between persons in mean RT. Polynomial models were then estimated to approximate the effects of practice across the six sessions, as presented below.

Polynomial Models

Polynomial models were first specified with a random intercept only. A fixed linear effect of session was significant ($p < .001$), such that average response time declined across sessions. The addition of a random linear slope (as well as a covariance between the random intercept and random linear slope) resulted in a significant improvement to the model, $-2\Delta LL(2) = 43$, $p < .001$. However, the magnitude of this linear decline was reduced in later sessions, as indicated by a significant fixed quadratic effect of session (i.e., a decelerating negative trend; $p < .001$). The addition of a random quadratic slope (and its two accompanying covariances with the random intercept and random linear slope) also resulted in a significant improvement in model fit, $-2\Delta LL(3) = 39$, $p < .001$.

The predicted means from the unconditional random quadratic polynomial model for session (i.e., without predictors) are shown in Figure 1, and model parameters using REML estimation are given in Table 1. As shown, the mean predicted response time at session 1 was 1946 ms, with a 95% CI of 916 to 2976 ms. The mean instantaneous linear rate of change at session 1 was -121 ms per session, with a 95% CI of -436 to 194 ms, indicating that not all participants were predicted to improve as evaluated at session 1. Half the mean deceleration in linear rate of change was 14 ms per session, such that the linear rate of change became less negative by 28 ms with each session. The 95% CI for the quadratic effect was of -36 to 63 ms, indicating that not all participants were predicted to decelerate in their rate of improvement across sessions.

Computing random effects confidence intervals for each random effect:

$$\text{Random Effect 95\% CI} = \text{fixed effect} \pm \left(1.96 * \sqrt{\text{Random Variance}}\right)$$

$$\text{Intercept 95\% CI} = \gamma_{00} \pm \left(1.96 * \sqrt{\tau_{U_0}^2}\right) \rightarrow 1,945.9 \pm \left(1.96 * \sqrt{276,209}\right) = 916 \text{ to } 2,976$$

$$\text{Linear Time Slope 95\% CI} = \gamma_{10} \pm \left(1.96 * \sqrt{\tau_{U_1}^2}\right) \rightarrow -120.9 \pm \left(1.96 * \sqrt{25,840}\right) = -436 \text{ to } 194$$

$$\text{Quadratic Time Slope 95\% CI} = \gamma_{20} \pm \left(1.96 * \sqrt{\tau_{U_2}^2}\right) \rightarrow 13.9 \pm \left(1.96 * \sqrt{634}\right) = -36 \text{ to } 63$$

Example 3: Time-Invariant Predictors of Practice Effects (uses same data as Example 2)

In this example we will examine time-invariant predictors of individual differences in intercepts, linear slopes, and quadratic slopes representing improvement in RT (in msec) across six practice sessions. We will examine age, abstract reasoning, and education in sequential conditional (predictor) models.

SAS Code for Data Manipulation:

```
* Centering level-2 predictor variables for analysis;
DATA work.example23; SET work.example23;
    age80 = age - 80;          * Convenient value;
    reas22 = absreas - 22;     * Near sample mean;
    LABEL age80 = "age80: Age Centered (0=80)"
           reas22 = "reas22: Abstract Reasoning Centered (0=22)";
    * Make education a grouping variable for purpose of demonstration only;
    IF educyrs LE 12          THEN educgrp=1;
    ELSE IF educyrs GT 12 AND EducYrs LE 16 THEN educgrp=2;
    ELSE IF educyrs GT 16     THEN educgrp=3;
    ELSE IF educyrs = .       THEN educgrp=. ;
    LABEL educgrp = "educgrp: Education Group (1=HS, 2=BA, 3=GRAD)";

* Removing cases with missing predictors;
IF NMISS(age80, reas22, educgrp)>0 THEN DELETE;
RUN;
```

SPSS Code for Data Manipulation:

```
* Centering level-2 predictor variables for analysis.
DATASET ACTIVATE example23 WINDOW=FRONT.
COMPUTE age80 = age - 80.
COMPUTE reas22 = absreas - 22.
VARIABLE LABELS
    age80 "age80: Age Centered (0=80)"
    reas22 "reas22: Abstract Reasoning Centered (0=22)".
* Make education a grouping variable for purpose of demonstration only.
IF educyrs LE 12          educgrp=1.
IF educyrs GT 12 AND educyrs LE 16 educgrp=2.
IF educyrs GT 16         educgrp=3.
VARIABLE LABELS educgrp "educgrp: Education Group (1=HS, 2=BA, 3=GRAD)".

* Removing cases with missing predictors.
SELECT IF (NVALID(age80, reas22, educgrp)=3).
EXECUTE.
```

STATA Code for Data Manipulation:

```
* centering level-2 predictor variables for analysis
gen age80 = age - 80
gen reas22 = absreas - 22
label variable age80 "age80: Age Centered (0=80 years)"
label variable reas22 "reas22: Abstract Reasoning Centered (0=22)"
* make education a grouping variable for purpose of demonstration only
gen educgrp=.
replace educgrp=1 if (educyrs <= 12)
replace educgrp=2 if (educyrs > 12 & educyrs <= 16)
replace educgrp=3 if (educyrs > 16)
label variable educgrp "educgrp: Education Group (1=HS, 2=BA, 3=GRAD)"

* create new variable to hold number of missing cases
* then drop cases with incomplete predictors
egen nummiss = rowmiss(age80 reas22 educgrp)
drop if nummiss>0
```

Model 3b. Random Quadratic Time Baseline (in ML now)

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$

Level 2:

Intercept: $\beta_{0i} = \gamma_{00} + U_{0i}$

Linear: $\beta_{1i} = \gamma_{10} + U_{1i}$

Quadratic: $\beta_{2i} = \gamma_{20} + U_{2i}$

```
TITLE1 "SAS Model 3b: Random Quadratic Time Baseline in ML";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS ID session;
  MODEL nm3rt = c1sess c1sess*c1sess / SOLUTION DDFM=Satterthwaite OUTPM=work.TimePred;
  RANDOM INTERCEPT c1sess c1sess*c1sess / G GCORR V VCORR TYPE=UN SUBJECT=ID;
  REPEATED session / R TYPE=VC SUBJECT=ID; RUN;
PROC CORR NOSIMPLE DATA=work.TimePred; VAR nm3rt pred; RUN;
```

```
TITLE "SPSS Model 3b: Random Quadratic Time Baseline in ML".
MIXED nm3rt BY ID session WITH c1sess
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV G R
  /FIXED = c1sess c1sess*c1sess
  /RANDOM = INTERCEPT c1sess c1sess*c1sess | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID)
  /SAVE = FIXPRED (predtime).
CORRELATIONS nm3rt predtime.
```

```
* STATA Model 3b: Random Quadratic Time Baseline in ML
xtmixed nm3rt c.c1sess c.c1sess#c.c1sess, || id: c1sess c1sess2, ///
  variance ml covariance(un) residuals(independent,t(session)),
  estat ic, n(101),
  estat recovariance, level(id),
  estimates store Baseline, // save LL for LRT
  predict predtime // save fixed-effect predicted outcomes
corr nm3rt predtime // get total r to make r2
```

STATA output:

| | | | |
|-----------------------------|--------------------|---|--------|
| Mixed-effects ML regression | Number of obs | = | 606 |
| Group variable: id | Number of groups | = | 101 |
| | Obs per group: min | = | 6 |
| | avg | = | 6.0 |
| | max | = | 6 |
| | Wald chi2(2) | = | 72.45 |
| Log likelihood = -4160.8833 | Prob > chi2 | = | 0.0000 |

| nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-------------------|-----------|-----------|-------|-------|----------------------|
| c1sess | -120.8999 | 19.94803 | -6.06 | 0.000 | -159.9973 -81.80251 |
| c.c1sess#c.c1sess | 13.86561 | 3.398459 | 4.08 | 0.000 | 7.204756 20.52647 |
| _cons | 1945.85 | 53.58259 | 36.31 | 0.000 | 1840.83 2050.87 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|-----------|-----------|----------------------|
| id: Unstructured | | | |
| var(c1sess) | 25437.86 | 5781.419 | 16293.81 39713.52 |
| var(c1sess2) | 622.8 | 169.99 | 364.7687 1063.358 |
| var(_cons) | 273306.9 | 40831.76 | 203930.4 366285.1 |
| cov(c1sess,c1sess2) | -3837.723 | 968.8047 | -5736.545 -1938.9 |
| cov(c1sess,_cons) | -35261.67 | 11771.5 | -58333.38 -12189.95 |
| cov(c1sess2,_cons) | 3845.378 | 1921.468 | 79.37031 7611.386 |

```

var(Residual) |      20298.2    1649.119        17310.2    23801.98
-----
LR test vs. linear regression:      chi2(6) =    891.99    Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.

```

```

.      estat ic, n(101),
-----
Model |      Obs      ll(null)      ll(model)      df      AIC      BIC
-----+-----
. |      101      .      -4160.883      10      8341.767      8367.918
-----
Note: N=101 used in calculating BIC

```

In ML, the #parms is ALL
parms (both sides of model).
So STATA's versions should
agree with other programs.

```

      |      nm3rt preptime
-----+-----
nm3rt |      1.0000
preptime |      0.1917      1.0000

```

R = .1917, so R² for time = .0367

The model for the means (fixed linear and quadratic session
effects so far) accounted for ~4% of the variance in RT.

Model 4a. Age as Predictor of Intercept, Linear, and Quadratic Time Slopes

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$

Level 2:

Intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Age}_i - 80) + U_{0i}$

Linear: $\beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Age}_i - 80) + U_{1i}$

Quadratic: $\beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Age}_i - 80) + U_{2i}$

```

TITLE1 "SAS Model 4a: Age as Predictor of Intercept, Linear, and Quadratic";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS ID session;
  MODEL nm3rt = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
    / SOLUTION DDFM=Satterthwaite OUTPM=work.AgePred;
  RANDOM INTERCEPT clsess clsess*clsess / G GCORR TYPE=UN SUBJECT=ID;
  REPEATED session / TYPE=VC SUBJECT=ID;
  * Requesting additional effects for age;
  ESTIMATE "Age Effect at Session 1" age80 1 clsess*age80 0 clsess*clsess*age80 0;
  ESTIMATE "Age Effect at Session 2" age80 1 clsess*age80 1 clsess*clsess*age80 1;
  ESTIMATE "Age Effect at Session 3" age80 1 clsess*age80 2 clsess*clsess*age80 4;
  ESTIMATE "Age Effect at Session 4" age80 1 clsess*age80 3 clsess*clsess*age80 9;
  ESTIMATE "Age Effect at Session 5" age80 1 clsess*age80 4 clsess*clsess*age80 16;
  ESTIMATE "Age Effect at Session 6" age80 1 clsess*age80 5 clsess*clsess*age80 25;
RUN; PROC CORR NOSIMPLE DATA=work.AgePred; VAR nm3rt pred; RUN;

```

```

TITLE "SPSS Model 4a: Age as Predictor of Intercept, Linear, and Quadratic".
MIXED nm3rt BY ID session WITH clsess age80
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV G R
  /FIXED = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
  /RANDOM = INTERCEPT clsess clsess*clsess | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID)
  /SAVE = FIXPRED (predage)
  /TEST = "Age Effect at Session 1" age80 1 clsess*age80 0 clsess*clsess*age80 0
  /TEST = "Age Effect at Session 2" age80 1 clsess*age80 1 clsess*clsess*age80 1
  /TEST = "Age Effect at Session 3" age80 1 clsess*age80 2 clsess*clsess*age80 4
  /TEST = "Age Effect at Session 4" age80 1 clsess*age80 3 clsess*clsess*age80 9
  /TEST = "Age Effect at Session 5" age80 1 clsess*age80 4 clsess*clsess*age80 16
  /TEST = "Age Effect at Session 6" age80 1 clsess*age80 5 clsess*clsess*age80 25.
CORRELATIONS nm3rt predage.

```

```

* STATA Model 4a: Age as Predictor of Intercept, Linear, and Quadratic
xtmixed nm3rt c.c1sess c.c1sess#c.c1sess          ///
      c.age80 c.age80#c.c1sess c.age80#c.c1sess#c.c1sess,  ///
      || id: c1sess c1sess2,          ///
      variance ml covariance(un) residuals(independent,t(session)),
      estat ic, n(101),
      estat recovariance, level(id),
      estimates store age,          // save LL for LRT
      lrtest Age Baseline,          // LRT against non-age baseline
      predict predage          // save fixed-effect predicted outcomes
      margins, at(c.c1sess=(0(1)5)) dydx(c.age80) vsquish // age slope per session
      margins, at(c.c1sess=(0(1)5) c.age80=(-5 0 5)) vsquish // predictions per session
      marginsplot, name(predicted_age, replace) // plot age predictions
corr nm3rt predage          // get total r to make r2

```

STATA output:

```

Mixed-effects ML regression          Number of obs      =      606
Group variable: id                  Number of groups   =      101
                                   Obs per group: min =       6
                                   avg =      6.0
                                   max =       6
                                   Wald chi2(5)      =     88.55
Log likelihood = -4155.1009          Prob > chi2        =     0.0000

```

| | nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------------------------|-------|-----------|-----------|-------|-------|----------------------|-----------|
| c1sess | | -121.8325 | 19.66948 | -6.19 | 0.000 | -160.3839 | -83.28099 |
| c.c1sess#c.c1sess | | 13.97744 | 3.375686 | 4.14 | 0.000 | 7.361221 | 20.59367 |
| age80 | | 29.04954 | 8.377364 | 3.47 | 0.001 | 12.63021 | 45.46887 |
| c.age80#c.c1sess | | -5.594634 | 3.251901 | -1.72 | 0.085 | -11.96824 | .7789759 |
| c.age80#c.c1sess#c.c1sess | | .6709122 | .558093 | 1.20 | 0.229 | -.42293 | 1.764754 |
| _cons | | 1950.692 | 50.67139 | 38.50 | 0.000 | 1851.378 | 2050.006 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|-----------|-----------|----------------------|--------------------------------------|
| id: Unstructured | | | | |
| var(c1sess) | 24293.61 | 5623.947 | 15432.62 | 38242.33 → linear var down by 4.50% |
| var(c1sess2) | 606.3449 | 167.7546 | 352.5508 | 1042.84 → quad var down by 2.64% |
| var(_cons) | 242456.1 | 36492.45 | 180516.8 | 325648.3 → intercept var down 11.29% |
| cov(c1sess,c1sess2) | -3700.505 | 949.404 | -5561.302 | -1839.707 |
| cov(c1sess,_cons) | -29320.18 | 10868.45 | -50621.95 | -8018.411 |
| cov(c1sess2,_cons) | 3132.873 | 1793.883 | -383.0738 | 6648.819 |
| var(Residual) | 20298.2 | 1649.119 | 17310.2 | 23801.98 → residual var not reduced |

LR test vs. linear regression: chi2(6) = 857.76 Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.

```

.      estat ic, n(101),
-----+-----
Model |   Obs   ll(null)   ll(model)   df       AIC       BIC
-----+-----
.      |   101           .   -4155.101    13     8336.202    8370.198
-----+-----

```

Note: N=101 used in calculating BIC

```

.      estimates store Age,          // save LL for LRT
.      lrtest Age Baseline,          // LRT against non-age baseline

Likelihood-ratio test          LR chi2(3) =    11.56
(Assumption: Baseline nested in Age)   Prob > chi2 =    0.0090

```

Is the age model (4a) better than the baseline random quadratic model (3b)?

Yes, $-2\Delta LL = 11.6$ on $df=3$, $p=.009$


```
.      predict predage                                // save fixed-effect predicted outcomes
(option xb assumed)
.      margins, at(c1sess=(0(1)5)) dydx(age80) vsquish // age slope per session
```

Average marginal effects Number of obs = 606

Expression : Linear prediction, fixed portion, predict()

dy/dx w.r.t. : age80

```
1._at      : c1sess      =      0
2._at      : c1sess      =      1
3._at      : c1sess      =      2
4._at      : c1sess      =      3
5._at      : c1sess      =      4
6._at      : c1sess      =      5
```

These are the simple
slopes for age at
each session.

| | | Delta-method | | | | | |
|-------|-----|--------------|-----------|------|-------|----------------------|----------|
| | | dy/dx | Std. Err. | z | P> z | [95% Conf. Interval] | |
| age80 | _at | | | | | | |
| | 1 | 29.04954 | 8.377364 | 3.47 | 0.001 | 12.63021 | 45.46887 |
| | 2 | 24.12582 | 7.609705 | 3.17 | 0.002 | 9.211068 | 39.04056 |
| | 3 | 20.54392 | 7.459286 | 2.75 | 0.006 | 5.923987 | 35.16385 |
| | 4 | 18.30385 | 7.330177 | 2.50 | 0.013 | 3.936962 | 32.67073 |
| | 5 | 17.4056 | 7.071475 | 2.46 | 0.014 | 3.545761 | 31.26543 |
| | 6 | 17.84917 | 7.054461 | 2.53 | 0.011 | 4.022683 | 31.67566 |

```
.      margins, at(c1sess=(0(1)5) age80=(-5 0 5)) vsquish // predictions per session
```

Adjusted predictions Number of obs = 606

Expression : Linear prediction, fixed portion, predict()

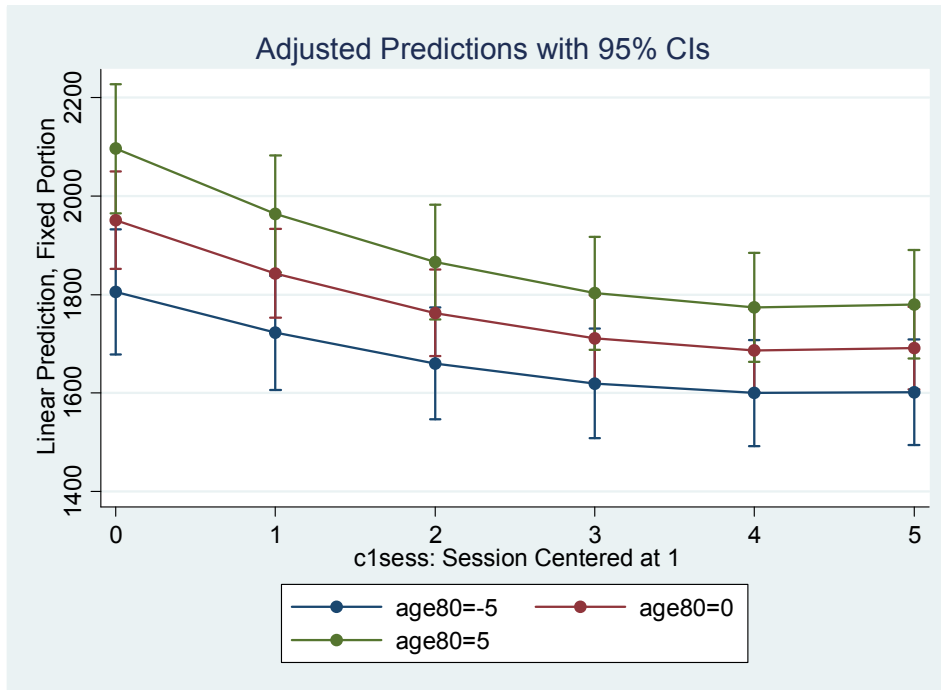
```
1._at      : c1sess      =      0
              age80      =     -5
2._at      : c1sess      =      0
              age80      =      0
3._at      : c1sess      =      0
              age80      =      5
```

(output continues for all other sessions)

| | | Delta-method | | | | | |
|--|-----|--------------|-----------|-------|-------|----------------------|----------|
| | | Margin | Std. Err. | z | P> z | [95% Conf. Interval] | |
| | _at | | | | | | |
| | 1 | 1805.444 | 64.84687 | 27.84 | 0.000 | 1678.347 | 1932.542 |
| | 2 | 1950.692 | 50.67139 | 38.50 | 0.000 | 1851.378 | 2050.006 |
| | 3 | 2095.94 | 66.62638 | 31.46 | 0.000 | 1965.354 | 2226.525 |
| | 4 | 1722.208 | 58.90463 | 29.24 | 0.000 | 1606.757 | 1837.659 |
| | 5 | 1842.837 | 46.02812 | 40.04 | 0.000 | 1752.623 | 1933.05 |
| | 6 | 1963.466 | 60.52108 | 32.44 | 0.000 | 1844.847 | 2082.085 |
| | 7 | 1660.217 | 57.74028 | 28.75 | 0.000 | 1547.048 | 1773.386 |
| | 8 | 1762.937 | 45.1183 | 39.07 | 0.000 | 1674.506 | 1851.367 |
| | 9 | 1865.656 | 59.32478 | 31.45 | 0.000 | 1749.382 | 1981.931 |
| | 10 | 1619.472 | 56.74089 | 28.54 | 0.000 | 1508.262 | 1730.682 |
| | 11 | 1710.991 | 44.33737 | 38.59 | 0.000 | 1624.092 | 1797.891 |
| | 12 | 1802.511 | 58.29796 | 30.92 | 0.000 | 1688.249 | 1916.773 |
| | 13 | 1599.973 | 54.73834 | 29.23 | 0.000 | 1492.688 | 1707.258 |
| | 14 | 1687.001 | 42.77258 | 39.44 | 0.000 | 1603.168 | 1770.834 |
| | 15 | 1774.029 | 56.24045 | 31.54 | 0.000 | 1663.8 | 1884.258 |
| | 16 | 1601.72 | 54.60664 | 29.33 | 0.000 | 1494.693 | 1708.747 |
| | 17 | 1690.966 | 42.66967 | 39.63 | 0.000 | 1607.335 | 1774.597 |
| | 18 | 1780.212 | 56.10514 | 31.73 | 0.000 | 1670.247 | 1890.176 |

The pattern of the interaction is shown by the simple effects of age at each session, graphed below.

```
. marginsplot, name(predicted_age, replace) // plot age predictions
```



Variables that uniquely identify margins: c1sess age80

```
. corr nm3rt predage // get total r to make r2
(obs=606)
```

| | nm3rt | predage |
|---------|--------|---------|
| nm3rt | 1.0000 | |
| predage | 0.3269 | 1.0000 |

R = .3269, so R^2 for time+age = .1069

The fixed effects of time before accounted for ~3.7% of the variance in RT, so there is a net increase of ~7% due to age.

Model 5a. +Abstract Reasoning as Predictor of Intercept, Linear, and Quadratic Time Slopes

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$$

Level 2:

$$\text{Intercept: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Age}_i - 80) + \gamma_{02}(\text{Reason}_i - 22) + U_{0i}$$

$$\text{Linear: } \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Age}_i - 80) + \gamma_{12}(\text{Reason}_i - 22) + U_{1i}$$

$$\text{Quadratic: } \beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Age}_i - 80) + \gamma_{22}(\text{Reason}_i - 22) + U_{2i}$$

```
TITLE1 "SAS Model 5a: +Reasoning as Predictor of Intercept, Linear, and Quadratic";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS ID session;
  MODEL nm3rt = c1sess c1sess*c1sess age80 c1sess*age80 c1sess*c1sess*age80
               reas22 c1sess*reas22 c1sess*c1sess*reas22
               / SOLUTION DDFM=Satterthwaite OUTPM=work.ReasPred;
  RANDOM INTERCEPT c1sess c1sess*c1sess / G GCORR TYPE=UN SUBJECT=ID;
  REPEATED session / TYPE=VC SUBJECT=ID;
RUN; PROC CORR NOSIMPLE DATA=work.ReasPred; VAR nm3rt pred; RUN;
```

TITLE "SPSS Model 5a: +Reasoning as Predictor of Intercept, Linear, and Quadratic".

```
MIXED nm3rt BY ID session WITH clsess age80 reas22
/METHOD = ML
/PRINT = SOLUTION TESTCOV G R
/FIXED = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
       reas22 clsess*reas22 clsess*clsess*reas22
/RANDOM = INTERCEPT clsess clsess*clsess | SUBJECT(ID) COVTYPE(UN)
/REPEATED = session | SUBJECT(ID) COVTYPE(ID)
/SAVE = FIXPRED (predreas).
CORRELATIONS nm3rt predreas.
```

*** STATA Model 5a: +Reasoning as Predictor of Intercept, Linear, and Quadratic**

```
xtmixed nm3rt c.clsess c.clsess#c.clsess ///
c.age80 c.age80#c.clsess c.age80#c.clsess#c.clsess ///
c.reas22 c.reas22#c.clsess c.reas22#c.clsess#c.clsess, ///
|| id: clsess clsess2, ///
variance ml covariance(un) residuals(independent,t(session)),
estat ic, n(101),
estat recovariance, level(id),
estimates store Reas, // save LL for LRT
lrtest Reas Age, // LRT against age baseline
predict predreas // save fixed-effect predicted outcomes
corr nm3rt predreas // get total r to make r2
```

STATA output:

```
Mixed-effects ML regression      Number of obs      =      606
Group variable: id              Number of groups   =      101
                                Obs per group: min =       6
                                avg =      6.0
                                max =       6
                                Wald chi2(8)    =     103.88
                                Prob > chi2     =      0.0000

Log likelihood = -4148.8645
```

| | nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|----------------------------|-------|-----------|-----------|-------|-------|----------------------|
| clsess | | -119.7417 | 19.77414 | -6.06 | 0.000 | -158.4983 -80.98505 |
| c.clsess#c.clsess | | 13.30362 | 3.36557 | 3.95 | 0.000 | 6.707229 19.90002 |
| age80 | | 22.27817 | 8.601751 | 2.59 | 0.010 | 5.419047 39.13729 |
| c.age80#c.clsess | | -6.492074 | 3.424732 | -1.90 | 0.058 | -13.20443 .2202772 |
| c.age80#c.clsess#c.clsess | | .9601368 | .5828914 | 1.65 | 0.100 | -.1823093 2.102583 |
| reas22 | | -27.10041 | 11.11411 | -2.44 | 0.015 | -48.88366 -5.317155 |
| c.reas22#c.clsess | | -3.591742 | 4.425011 | -0.81 | 0.417 | -12.2646 5.081121 |
| c.reas22#c.clsess#c.clsess | | 1.157537 | .7531395 | 1.54 | 0.124 | -.3185897 2.633663 |
| _cons | | 1966.467 | 49.66585 | 39.59 | 0.000 | 1869.124 2063.811 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|-----------|-----------|---|
| id: Unstructured | | | |
| var(clsess) | 24040.63 | 5589.24 | 15242.24 37917.78 → linear var down by 1.04% |
| var(clsess2) | 580.0652 | 164.1907 | 333.0729 1010.216 → quad var down by 4.33% |
| var(_cons) | 228049.3 | 34467.25 | 169581.6 306675.4 → intercept var down by 5.94% |
| cov(clsess,clsess2) | -3618.966 | 937.0759 | -5455.601 -1782.331 |
| cov(clsess,_cons) | -31229.47 | 10655.91 | -52114.67 -10344.27 |
| cov(clsess2,_cons) | 3748.206 | 1747.738 | 322.7024 7173.709 |
| var(Residual) | 20298.18 | 1649.113 | 17310.18 23801.94 → residual var not reduced |

LR test vs. linear regression: chi2(6) = 832.43 Prob > chi2 = 0.0000

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 101 | . | -4148.864 | 16 | 8329.729 | 8371.571 |

```
.      lrtest Reas Age,                // LRT against age baseline

Likelihood-ratio test                LR chi2(3) =      12.47
(Assumption: Age nested in Reas)    Prob > chi2 =      0.0059

. corr nm3rt predreas                // get total r to make r2
(obs=606)
```

Is the reasoning model (5a) better than the age model (4a)?

Yes, $-2\Delta LL = 12.5$ on $df=3$, $p=.0059$, so ΔR^2 is significant

```
-----+-----
      |      nm3rt predreas
nm3rt |      1.0000
predreas |      0.4011      1.0000
```

$R = .4011$, so R^2 for time+age+reas = .1609

The fixed effects of time and age before accounted for ~10.7% of the variance in RT, so there is a net increase of ~5.4% due to reasoning.

Model 5b. Abstract Reasoning on Intercept and Linear Time Slope Only

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$

Level 2:

Intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Age}_i - 80) + \gamma_{02}(\text{Reason}_i - 22) + U_{0i}$

Linear: $\beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Age}_i - 80) + \gamma_{12}(\text{Reason}_i - 22) + U_{1i}$

Quadratic: $\beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Age}_i - 80) + U_{2i}$

```
TITLE1 "SAS Model 5b: Reasoning on Intercept and Linear Time Slope Only";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS ID session;
  MODEL nm3rt = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
               reas22 clsess*reas22
               / SOLUTION DDFM=Satterthwaite OUTPM=work.ReasPred2;
  RANDOM INTERCEPT clsess clsess*clsess / G GCORR TYPE=UN SUBJECT=ID;
  REPEATED session / TYPE=VC SUBJECT=ID;
  * Requesting additional effects for reasoning instead;
  ESTIMATE "Reasoning Effect at Session 1" reas22 1 clsess*reas22 0;
  ESTIMATE "Reasoning Effect at Session 2" reas22 1 clsess*reas22 1;
  ESTIMATE "Reasoning Effect at Session 3" reas22 1 clsess*reas22 2;
  ESTIMATE "Reasoning Effect at Session 4" reas22 1 clsess*reas22 3;
  ESTIMATE "Reasoning Effect at Session 5" reas22 1 clsess*reas22 4;
  ESTIMATE "Reasoning Effect at Session 6" reas22 1 clsess*reas22 5;
RUN; PROC CORR NOSIMPLE DATA=work.ReasPred2; VAR nm3rt pred; RUN;
```

```
TITLE "SPSS Model 5b: +Reasoning as Predictor of Intercept, Linear Time Slope Only".
MIXED nm3rt BY ID session WITH clsess age80 reas22
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV G R
  /FIXED = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
           reas22 clsess*reas22
  /RANDOM = INTERCEPT clsess clsess*clsess | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID)
  /SAVE = FIXPRED (predreas2)
  /TEST = "Reasoning Effect at Session 1" reas22 1 clsess*reas22 0
  /TEST = "Reasoning Effect at Session 2" reas22 1 clsess*reas22 1
  /TEST = "Reasoning Effect at Session 3" reas22 1 clsess*reas22 2
  /TEST = "Reasoning Effect at Session 4" reas22 1 clsess*reas22 3
  /TEST = "Reasoning Effect at Session 5" reas22 1 clsess*reas22 4
  /TEST = "Reasoning Effect at Session 6" reas22 1 clsess*reas22 5.
CORRELATIONS nm3rt predreas2.
```

```
* STATA Model 5b: +Reasoning as Predictor of Intercept, Linear Time Slope Only
xtmixed nm3rt c.clsess c.clsess#c.clsess ///
  c.age80 c.age80#c.clsess c.age80#c.clsess#c.clsess ///
  c.reas22 c.reas22#c.clsess, || id: clsess clsess2, ///
  variance ml covariance(un) residuals(independent,t(session)),
  estat ic, n(101),
```

```

estat recovariance, level(id),
estimates store Reas2,          // save LL for LRT
lrtest Reas2 Age,              // LRT against age baseline
margins, at(c.c1sess=(0(1)5)) dydx(c.reas22) vsquish // reas slope per session
margins, at(c.c1sess=(0(1)5) c.reas22=(-5 0 5)) vsquish // predictions per session
marginsplot, name(predicted_reas, replace) // plot reas predictions
predict predreas2              // save fixed-effect predicted outcomes
corr nm3rt predreas2          // get total r to make r2

```

STATA output:

```

Mixed-effects ML regression              Number of obs   =       606
Group variable: id                     Number of groups =       101
                                      Obs per group: min =        6
                                      avg =       6.0
                                      max =        6
                                      Wald chi2(7)      =      101.09
                                      Prob > chi2       =       0.0000

Log likelihood = -4150.032

```

| | nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------------------------|-------|-----------|-----------|-------|-------|----------------------|-----------|
| c1sess | | -123.5416 | 19.82897 | -6.23 | 0.000 | -162.4057 | -84.67758 |
| c.c1sess#c.c1sess | | 13.97744 | 3.375697 | 4.14 | 0.000 | 7.3612 | 20.59369 |
| age80 | | 20.84705 | 8.561395 | 2.44 | 0.015 | 4.067021 | 37.62707 |
| c.age80#c.c1sess | | -4.860993 | 3.290742 | -1.48 | 0.140 | -11.31073 | 1.588743 |
| c.age80#c.c1sess#c.c1sess | | .6709122 | .5580948 | 1.20 | 0.229 | -.4229335 | 1.764758 |
| reas22 | | -32.82806 | 10.47071 | -3.14 | 0.002 | -53.35027 | -12.30585 |
| c.reas22#c.c1sess | | 2.93618 | 1.241355 | 2.37 | 0.018 | .5031693 | 5.369191 |
| _cons | | 1969.802 | 49.6827 | 39.65 | 0.000 | 1872.425 | 2067.178 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | | |
|---------------------------|-----------|-----------|----------------------|-----------|-------------------------------|
| | | | | | Relative to age-only model: |
| id: Unstructured | | | | | |
| var(c1sess) | 24876.74 | 5713.188 | 15860.12 | 39019.37 | → linear var up by -2.40% |
| var(c1sess2) | 606.3538 | 167.7562 | 352.557 | 1042.852 | → quad var not reduced |
| var(_cons) | 228693.2 | 34638.71 | 169952.4 | 307736.8 | → intercept var down by 5.68% |
| cov(c1sess,c1sess2) | -3767.223 | 957.7847 | -5644.446 | -1889.999 | |
| cov(c1sess,_cons) | -31963.12 | 10883.66 | -53294.71 | -10631.53 | |
| cov(c1sess2,_cons) | 3878.292 | 1787.961 | 373.9525 | 7382.631 | |
| | | | | | |
| var(Residual) | 20298.14 | 1649.108 | 17310.16 | 23801.9 | → residual var not reduced |

LR test vs. linear regression: chi2(6) = 830.58 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

```

. estat ic, n(101),

```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 101 | . | -4150.032 | 15 | 8330.064 | 8369.291 |

Note: N=101 used in calculating BIC

```

. estimates store Reas2, // save LL for LRT
. lrtest Reas2 Age, // LRT against age baseline

```

```

Likelihood-ratio test      LR chi2(2) =    10.14
(Assumption: Age nested in Reas2) Prob > chi2 =    0.0063

```

```

. margins, at(c1sess=(0(1)5)) dydx(reas22) vsquish // reas slope per session

```

```

Average marginal effects      Number of obs   =       606

```

```

Expression : Linear prediction, fixed portion, predict()

```

Is the revised reasoning model (5b) still better than the age model (4a)?

Yes, $-2\Delta LL = 10.1$ on $df=2$, $p=.006$ (so only 2.4 of the previous $-2\Delta LL$ was due to reason*quad)

```
1._at      : c1sess      =      0
2._at      : c1sess      =      1
3._at      : c1sess      =      2
4._at      : c1sess      =      3
5._at      : c1sess      =      4
6._at      : c1sess      =      5
```

These are the simple slopes for reasoning at each session.

| | | Delta-method | | | | | |
|--------|-----|--------------|-----------|-------|-------|----------------------|-----------|
| | | dy/dx | Std. Err. | z | P> z | [95% Conf. Interval] | |
| reas22 | _at | | | | | | |
| | 1 | -32.82806 | 10.47071 | -3.14 | 0.002 | -53.35027 | -12.30585 |
| | 2 | -29.89188 | 9.961528 | -3.00 | 0.003 | -49.41612 | -10.36764 |
| | 3 | -26.9557 | 9.586986 | -2.81 | 0.005 | -45.74585 | -8.165552 |
| | 4 | -24.01952 | 9.363252 | -2.57 | 0.010 | -42.37116 | -5.667884 |
| | 5 | -21.08334 | 9.301214 | -2.27 | 0.023 | -39.31338 | -2.853295 |
| | 6 | -18.14716 | 9.404074 | -1.93 | 0.054 | -36.57881 | .2844865 |

```
. margins, at(c1sess=(0(1)5) reas22=(-5 0 5)) vsquish // predictions per session
```

| | | | |
|--------------------|---------------|---|-----|
| Predictive margins | Number of obs | = | 606 |
|--------------------|---------------|---|-----|

Expression : Linear prediction, fixed portion, predict()

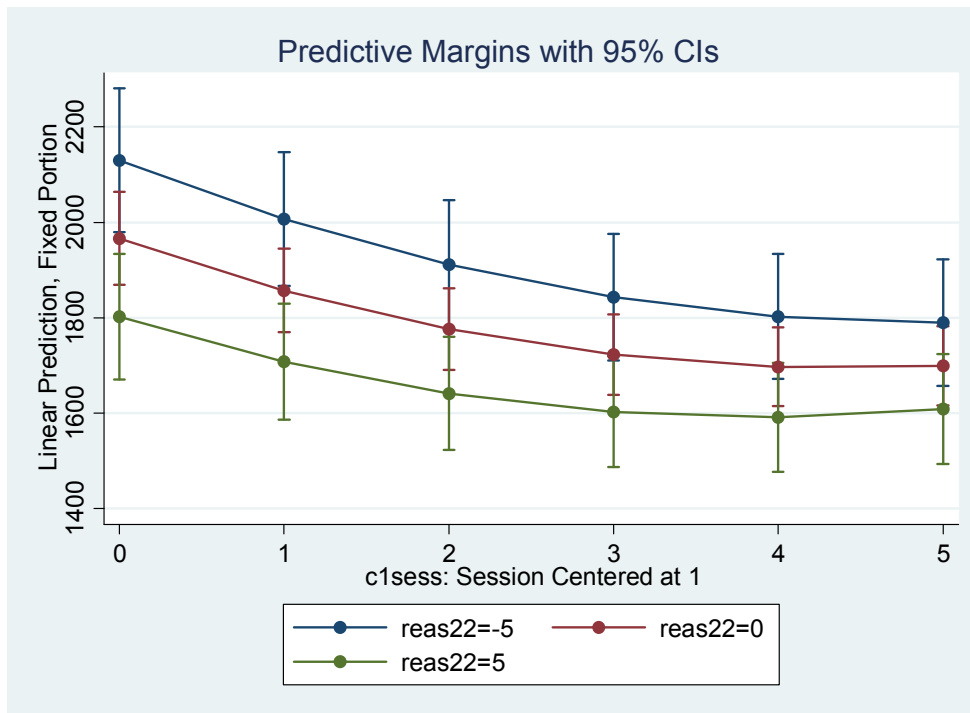
```

1._at      : c1sess      =      0
             reas22      =     -5
2._at      : c1sess      =      0
             reas22      =      0
3._at      : c1sess      =      0
             reas22      =      5

```

(output continues for all other sessions)

| | | Delta-method | | | | |
|----|--|--------------|-----------|-------|-------|----------------------|
| | | Margin | Std. Err. | z | P> z | [95% Conf. Interval] |
| at | | | | | | |
| 1 | | 2130.467 | 76.79056 | 27.74 | 0.000 | 1979.96 2280.974 |
| 2 | | 1966.327 | 49.71952 | 39.55 | 0.000 | 1868.878 2063.775 |
| 3 | | 1802.186 | 67.29827 | 26.78 | 0.000 | 1670.284 1934.089 |
| 4 | | 2006.92 | 71.31601 | 28.14 | 0.000 | 1867.143 2146.697 |
| 5 | | 1857.461 | 44.5668 | 41.68 | 0.000 | 1770.112 1944.81 |
| 6 | | 1708.002 | 62.03248 | 27.53 | 0.000 | 1586.42 1829.583 |
| 7 | | 1911.105 | 69.07653 | 27.67 | 0.000 | 1775.717 2046.492 |
| 8 | | 1776.326 | 43.59481 | 40.75 | 0.000 | 1690.882 1861.771 |
| 9 | | 1641.548 | 60.20767 | 27.26 | 0.000 | 1523.543 1759.553 |
| 10 | | 1843.021 | 67.77101 | 27.19 | 0.000 | 1710.192 1975.849 |
| 11 | | 1722.923 | 43.0615 | 40.01 | 0.000 | 1638.524 1807.322 |
| 12 | | 1602.825 | 59.15403 | 27.10 | 0.000 | 1486.886 1718.765 |
| 13 | | 1802.668 | 66.80286 | 26.98 | 0.000 | 1671.736 1933.599 |
| 14 | | 1697.251 | 41.95444 | 40.45 | 0.000 | 1615.022 1779.48 |
| 15 | | 1591.834 | 58.16663 | 27.37 | 0.000 | 1477.83 1705.839 |
| 16 | | 1790.046 | 67.57664 | 26.49 | 0.000 | 1657.598 1922.494 |
| 17 | | 1699.31 | 42.47414 | 40.01 | 0.000 | 1616.062 1782.558 |
| 18 | | 1608.574 | 58.8501 | 27.33 | 0.000 | 1493.23 1723.918 |



```
. marginsplot, name(predicted_reas, replace) // plot reas predictions
Variables that uniquely identify margins: c1sess reas22

. predict predreas2 // save fixed-effect predicted outcomes
(option xb assumed)

. corr nm3rt predreas2 // get total r to make r2
(obs=606)
```

| | nm3rt | predre~2 |
|-----------|--------|----------|
| nm3rt | 1.0000 | |
| predreas2 | 0.4001 | 1.0000 |

R = .4001, so R² for time+age+reas = .1601

So ~0.1% of the variance accounted for previously was due to reason*quad

Model 6a. +Education Group on Intercept, Linear, and Quadratic Time Slopes

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Session}_{ti} - 1) + \beta_{2i}(\text{Session}_{ti} - 1)^2 + e_{ti}$$

Level 2:

$$\text{Intercept: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Age}_i - 80) + \gamma_{02}(\text{Reason}_i - 22) + \gamma_{03}(\text{Highvs.LowEd}_i) + \gamma_{04}(\text{Highvs.MedEd}_i) + U_{0i}$$

$$\text{Linear: } \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Age}_i - 80) + \gamma_{12}(\text{Reason}_i - 22) + \gamma_{13}(\text{Highvs.LowEd}_i) + \gamma_{14}(\text{Highvs.MedEd}_i) + U_{1i}$$

$$\text{Quadratic: } \beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Age}_i - 80) + \gamma_{23}(\text{Highvs.LowEd}_i) + \gamma_{24}(\text{Highvs.MedEd}_i) + U_{2i}$$

Additional model-implied group differences:

$$\text{Medium vs. Low education intercept} = (\gamma_{00} + \gamma_{04}) - (\gamma_{00} + \gamma_{03}) = \gamma_{04} - \gamma_{03}$$

$$\text{Medium vs. Low education linear session} = (\gamma_{10} + \gamma_{14}) - (\gamma_{10} + \gamma_{13}) = \gamma_{14} - \gamma_{13}$$

$$\text{Medium vs. Low education quadratic session} = (\gamma_{20} + \gamma_{24}) - (\gamma_{20} + \gamma_{23}) = \gamma_{24} - \gamma_{23}$$


```

TITLE1 "SAS Model 6a: +Education Group on Intercept, Linear, and Quadratic";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS ID session educgrp;
  MODEL nm3rt = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
               reas22 clsess*reas22 educgrp clsess*educgrp clsess*clsess*educgrp
               / SOLUTION DDFM=Satterthwaite OUTPM=work.EducPred;
  RANDOM INTERCEPT clsess clsess*clsess / G GCORR TYPE=UN SUBJECT=ID;
  REPEATED session / TYPE=VC SUBJECT=ID;
  * Estimating group means at first and last sessions
  LSMEANS educgrp / AT (clsess) = (0) DIFF=ALL;
  LSMEANS educgrp / AT (clsess) = (5) DIFF=ALL;
  * Contrasts between groups on intercept, linear, and quadratic slopes
  ESTIMATE "L vs. H Educ for Intercept Main Effect" educgrp -1 0 1 ;
  ESTIMATE "M vs. H Educ for Intercept Main Effect" educgrp 0 -1 1 ;
  ESTIMATE "L vs. M Educ for Intercept Main Effect" educgrp -1 1 0 ;
  ESTIMATE "L vs. H Educ for Linear Session" clsess*educgrp -1 0 1 ;
  ESTIMATE "M vs. H Educ for Linear Session" clsess*educgrp 0 -1 1 ;
  ESTIMATE "L vs. M Educ for Linear Session" clsess*educgrp -1 1 0 ;
  ESTIMATE "L vs. H Educ for Quadratic Session" clsess*clsess*educgrp -1 0 1 ;
  ESTIMATE "M vs. H Educ for Quadratic Session" clsess*clsess*educgrp 0 -1 1 ;
  ESTIMATE "L vs. M Educ for Quadratic Session" clsess*clsess*educgrp -1 1 0 ;
RUN; PROC CORR NOSIMPLE DATA=work.EducPred; VAR nm3rt pred; RUN;

```

Think of the -1 as the
"0" and the "1" as the
"1" in a dummy code.

```

TITLE "SPSS Model 6a: +Education as Predictor of Intercept, Linear, and Quadratic".
MIXED nm3rt BY ID session educgrp WITH clsess age80 reas22
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV G R
  /FIXED = clsess clsess*clsess age80 clsess*age80 clsess*clsess*age80
           reas22 clsess*reas22 educgrp clsess*educgrp clsess*clsess*educgrp
  /RANDOM = INTERCEPT clsess clsess*clsess | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = session | SUBJECT(ID) COVTYPE(ID)
  /SAVE = FIXPRED (prededuc)
  /EMMEANS = TABLES(educgrp) WITH (clsess=0) COMPARE(educgrp)
  /EMMEANS = TABLES(educgrp) WITH (clsess=5) COMPARE(educgrp)
  /TEST = "L vs. H Educ for for Main Effect" educgrp -1 0 1
  /TEST = "M vs. H Educ for for Main Effect" educgrp 0 -1 1
  /TEST = "L vs. M Educ for for Main Effect" educgrp -1 1 0
  /TEST = "L vs. H Educ for for Linear Session" clsess*educgrp -1 0 1
  /TEST = "M vs. H Educ for for Linear Session" clsess*educgrp 0 -1 1
  /TEST = "L vs. M Educ for for Linear Session" clsess*educgrp -1 1 0
  /TEST = "L vs. H Educ for for Quadratic Session" clsess*clsess*educgrp -1 0 1
  /TEST = "M vs. H Educ for for Quadratic Session" clsess*clsess*educgrp 0 -1 1
  /TEST = "L vs. M Educ for for Quadratic Session" clsess*clsess*educgrp -1 1 0.
CORRELATIONS nm3rt prededuc.

```

```

* STATA Model 6a: +Education Group on Intercept, Linear, and Quadratic
xtmixed nm3rt c.clsess c.clsess#c.clsess ///
  c.age80 c.age80#c.clsess c.age80#c.clsess#c.clsess ///
  c.reas22 c.reas22#c.clsess ///
  b(last).educgrp ib(last).educgrp#c.clsess ///
  ib(last).educgrp#c.clsess#c.clsess, || id: clsess clsess2, ///
  variance ml covariance(un) residuals(independent,t(session)),
  estat ic, n(101),
  estat recovariance, level(id),
  estimates store Educ,
  lrtest Educ Reas2,
  * Estimating group means at first and last sessions
  margins ib(last).educgrp, at(c.clsess=(0 5))
  * Contrasts between groups on intercept, linear, and quadratic slopes
  test 1.educgrp=3.educgrp // Low vs. High: Intercept
  test 2.educgrp=3.educgrp // Med vs. High: Intercept
  test 1.educgrp=2.educgrp // Low vs. Med: Intercept
  test 1.educgrp#c.clsess=3.educgrp#c.clsess // Low vs. High: Linear
  test 2.educgrp#c.clsess=3.educgrp#c.clsess // Med vs. High: Linear
  test 1.educgrp#c.clsess=2.educgrp#c.clsess // Low vs. Med: Linear
  test 1.educgrp#c.clsess#c.clsess=3.educgrp#c.clsess#c.clsess // Low vs. High: Quad
  test 2.educgrp#c.clsess#c.clsess=3.educgrp#c.clsess#c.clsess // Med vs. High: Quad
  test 1.educgrp#c.clsess#c.clsess=2.educgrp#c.clsess#c.clsess // Low vs. Med: Quad

```

```

contrast educgrp, // omnibus group diff on intercept
contrast educgrp#c.c1sess, // omnibus group diff on linear
contrast educgrp#c.c1sess#c.c1sess, // omnibus group diff on quadratic
margins, at(c.c1sess=(0(1)5) educgrp=(1 2 3)) vsquish // predictions per session
marginsplot, name(predicted_educ, replace) // plot educ predictions
predict prededuc // save fixed-effect predicted outcomes
corr nm3rt prededuc // get total r to make r2

```

STATA output:

```

Mixed-effects ML regression      Number of obs      =      606
Group variable: id              Number of groups   =      101
                                Obs per group: min =       6
                                avg =      6.0
                                max =       6
                                Wald chi2(13)  =     106.94
                                Prob > chi2    =      0.0000
Log likelihood = -4147.6829

```

| | nm3rt | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|--|---------------------------|-----------|-----------|-------|-------|----------------------|-----------|
| | c1sess | -106.4987 | 40.28349 | -2.64 | 0.008 | -185.4529 | -27.54452 |
| | c.c1sess#c.c1sess | 12.47966 | 6.848972 | 1.82 | 0.068 | -.9440805 | 25.9034 |
| | age80 | 20.28963 | 8.560341 | 2.37 | 0.018 | 3.511673 | 37.06759 |
| | c.age80#c.c1sess | -4.575964 | 3.267261 | -1.40 | 0.161 | -10.97968 | 1.827749 |
| | c.age80#c.c1sess#c.c1sess | .6176862 | .5534216 | 1.12 | 0.264 | -.4670003 | 1.702373 |
| | reas22 | -36.62127 | 10.76417 | -3.40 | 0.001 | -57.71865 | -15.52389 |
| | c.reas22#c.c1sess | 2.978327 | 1.280262 | 2.33 | 0.020 | .4690609 | 5.487594 |
| | educgrp | | | | | | |
| | 1 | -51.37682 | 151.0698 | -0.34 | 0.734 | -347.4683 | 244.7146 |
| | 2 | 37.64254 | 120.8739 | 0.31 | 0.755 | -199.2659 | 274.5509 |
| | educgrp#c.c1sess | | | | | | |
| | 1 | -70.24589 | 59.07811 | -1.19 | 0.234 | -186.0368 | 45.54507 |
| | 2 | -4.357662 | 48.13262 | -0.09 | 0.928 | -98.69587 | 89.98055 |
| | educgrp#c.c1sess#c.c1sess | | | | | | |
| | 1 | 11.06526 | 10.03239 | 1.10 | 0.270 | -8.597857 | 30.72837 |
| | 2 | -1.464111 | 8.188464 | -0.18 | 0.858 | -17.51321 | 14.58498 |
| | _cons | 1961.886 | 101.7896 | 19.27 | 0.000 | 1762.382 | 2161.39 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|-----------|-----------|----------------------|--|
| id: Unstructured | | | | |
| var(c1sess) | 24143.68 | 5618.86 | 15300.54 | 38097.82 → linear var down by 2.95% |
| var(c1sess2) | 582.0323 | 164.4561 | 334.5305 | 1012.648 → quad var down by 4.01% |
| var(_cons) | 228602.6 | 34699.74 | 169776.2 | 307812.1 → intercept var down by 0.04% |
| cov(c1sess,c1sess2) | -3636.004 | 939.9484 | -5478.269 | -1793.739 |
| cov(c1sess,_cons) | -33285.68 | 10917.47 | -54683.52 | -11887.83 |
| cov(c1sess2,_cons) | 4127.595 | 1789.937 | 619.382 | 7635.808 |
| var(Residual) | 20298.12 | 1649.105 | 17310.14 | 23801.87 → residual var not reduced |

LR test vs. linear regression: chi2(6) = 826.31 Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.

```
. estat ic, n(101),
```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 101 | . | -4147.683 | 21 | 8337.366 | 8392.283 |

Note: N=101 used in calculating BIC

```
. estimates store Educ,
. lrtest Educ Reas2,
```

Is the education model (6a) better than the revised reasoning model (5b)?
 No, $-2\Delta LL = 4.7$ on $df=6$, $p = .583$

```
Likelihood-ratio test                                LR chi2(6) =      4.70
(Assumption: Reas2 nested in Educ)                   Prob > chi2 =    0.5831
```

```
. * Estimating group means at first and last sessions
. margins ib(last).educgrp, at(c1sess=(0))
Predictive margins                                Number of obs =      606
Expression   : Linear prediction, fixed portion, predict()
at           : c1sess =      0
```

```
-----+-----
      |               Delta-method
      |      Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
educgrp |
  1 |      1884.284   110.9362    16.99   0.000     1666.853     2101.715
  2 |      1973.304    66.56898    29.64   0.000     1842.831     2103.776
  3 |      1935.661   101.2482    19.12   0.000     1737.218     2134.104
-----+-----
```

```
. margins ib(last).educgrp, at(c1sess=(5))
Predictive margins                                Number of obs =      606
Expression   : Linear prediction, fixed portion, predict()
at           : c1sess =      5
```

```
-----+-----
      |               Delta-method
      |      Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
educgrp |
  1 |      1599.713    94.54367    16.92   0.000     1414.41     1785.015
  2 |      1704.939    56.6088     30.12   0.000     1593.988     1815.89
  3 |      1725.687    86.06347    20.05   0.000     1557.006     1894.369
-----+-----
```

```
. * Contrasts between groups on intercept, linear, and quadratic slopes
. test 1.educgrp=3.educgrp                                // Low vs. High: Intercept
( 1)  [nm3rt]1.educgrp - [nm3rt]3b.educgrp = 0
      chi2( 1) =    0.12
      Prob > chi2 =   0.7338

. test 2.educgrp=3.educgrp                                // Med vs. High: Intercept
( 1)  [nm3rt]2.educgrp - [nm3rt]3b.educgrp = 0
      chi2( 1) =    0.10
      Prob > chi2 =   0.7555

. test 1.educgrp=2.educgrp                                // Low vs. Med: Intercept
( 1)  [nm3rt]1.educgrp - [nm3rt]2.educgrp = 0
      chi2( 1) =    0.46
      Prob > chi2 =   0.4960

. test 1.educgrp#c.c1sess=3.educgrp#c.c1sess              // Low vs. High: Linear
( 1)  [nm3rt]1.educgrp#c.c1sess - [nm3rt]3b.educgrp#co.c1sess = 0
      chi2( 1) =    1.41
      Prob > chi2 =   0.2344

. test 2.educgrp#c.c1sess=3.educgrp#c.c1sess              // Med vs. High: Linear
( 1)  [nm3rt]2.educgrp#c.c1sess - [nm3rt]3b.educgrp#co.c1sess = 0
      chi2( 1) =    0.01
      Prob > chi2 =   0.9279

. test 1.educgrp#c.c1sess=2.educgrp#c.c1sess              // Low vs. Med: Linear
```

```

( 1) [nm3rt]1.educgrp#c.c1sess - [nm3rt]2.educgrp#c.c1sess = 0
      chi2( 1) =      1.69
      Prob > chi2 =    0.1939

.      test 1.educgrp#c.c1sess#c.c1sess=3.educgrp#c.c1sess#c.c1sess // Low vs. High: Quad
( 1) [nm3rt]1.educgrp#c.c1sess#c.c1sess - [nm3rt]3b.educgrp#co.c1sess#co.c1sess = 0
      chi2( 1) =      1.22
      Prob > chi2 =    0.2700

.      test 2.educgrp#c.c1sess#c.c1sess=3.educgrp#c.c1sess#c.c1sess // Med vs. High: Quad
( 1) [nm3rt]2.educgrp#c.c1sess#c.c1sess - [nm3rt]3b.educgrp#co.c1sess#co.c1sess = 0
      chi2( 1) =      0.03
      Prob > chi2 =    0.8581

.      test 1.educgrp#c.c1sess#c.c1sess=2.educgrp#c.c1sess#c.c1sess // Low vs. Med: Quad
( 1) [nm3rt]1.educgrp#c.c1sess#c.c1sess - [nm3rt]2.educgrp#c.c1sess#c.c1sess = 0
      chi2( 1) =      2.12
      Prob > chi2 =    0.1454

.      contrast educgrp, // omnibus group diff on intercept

Contrasts of marginal linear predictions
Margins      : asbalanced
-----+-----
      |      df      chi2      P>chi2
-----+-----
nm3rt  |
educgrp |      2      0.48      0.7869
-----+-----

.      contrast educgrp#c.c1sess, // omnibus group diff on linear
Contrasts of marginal linear predictions
Margins      : asbalanced
-----+-----
      |      df      chi2      P>chi2
-----+-----
nm3rt  |
educgrp#c.c1sess |      2      1.92      0.3827
-----+-----

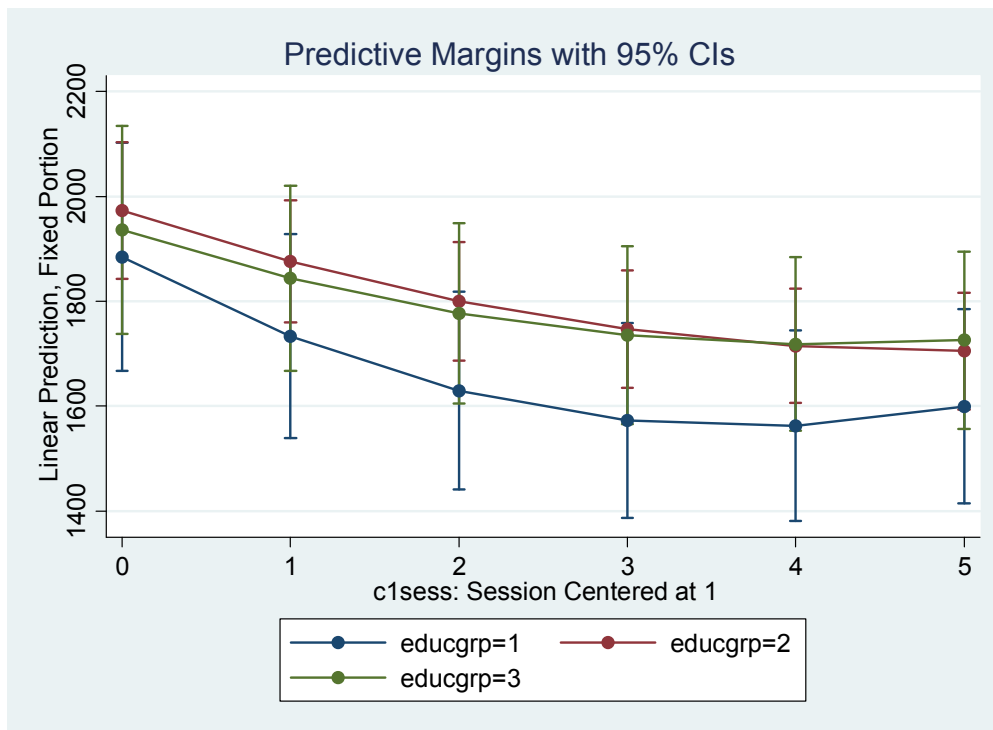
.      contrast educgrp#c.c1sess#c.c1sess, // omnibus group diff on quadratic
Contrasts of marginal linear predictions
Margins      : asbalanced
-----+-----
      |      df      chi2      P>chi2
-----+-----
nm3rt  |
educgrp#c.c1sess#c.c1sess |      2      2.18      0.3358
-----+-----

.      margins, at(c.c1sess=(0(1)5) educgrp=(1 2 3)) vsquish // predictions per session
Predictive margins      Number of obs      =      606
Expression : Linear prediction, fixed portion, predict()
1._at      : c1sess      =      0
              educgrp      =      1
2._at      : c1sess      =      0
              educgrp      =      2
3._at      : c1sess      =      0
              educgrp      =      3
(output continues for all other sessions)

```

| | Delta-method | | | | [95% Conf. Interval] | |
|-----|--------------|-----------|-------|-------|----------------------|----------|
| | Margin | Std. Err. | z | P> z | | |
| _at | | | | | | |
| 1 | 1884.284 | 110.9362 | 16.99 | 0.000 | 1666.853 | 2101.715 |
| 2 | 1973.304 | 66.56898 | 29.64 | 0.000 | 1842.831 | 2103.776 |
| 3 | 1935.661 | 101.2482 | 19.12 | 0.000 | 1737.218 | 2134.104 |
| 4 | 1733.602 | 98.99478 | 17.51 | 0.000 | 1539.576 | 1927.628 |
| 5 | 1875.98 | 59.24278 | 31.67 | 0.000 | 1759.867 | 1992.094 |
| 6 | 1844.159 | 90.05891 | 20.48 | 0.000 | 1667.647 | 2020.672 |
| 7 | 1629.804 | 96.28618 | 16.93 | 0.000 | 1441.086 | 1818.521 |
| 8 | 1800.482 | 57.64972 | 31.23 | 0.000 | 1687.491 | 1913.474 |
| 9 | 1777.411 | 87.64532 | 20.28 | 0.000 | 1605.63 | 1949.193 |
| 10 | 1572.889 | 94.93902 | 16.57 | 0.000 | 1386.812 | 1758.967 |
| 11 | 1746.809 | 56.86718 | 30.72 | 0.000 | 1635.352 | 1858.267 |
| 12 | 1735.417 | 86.46259 | 20.07 | 0.000 | 1565.953 | 1904.88 |
| 13 | 1562.859 | 92.79589 | 16.84 | 0.000 | 1380.983 | 1744.736 |
| 14 | 1714.961 | 55.5437 | 30.88 | 0.000 | 1606.098 | 1823.825 |
| 15 | 1718.175 | 84.43878 | 20.35 | 0.000 | 1552.678 | 1883.672 |
| 16 | 1599.713 | 94.54367 | 16.92 | 0.000 | 1414.41 | 1785.015 |
| 17 | 1704.939 | 56.6088 | 30.12 | 0.000 | 1593.988 | 1815.89 |
| 18 | 1725.687 | 86.06347 | 20.05 | 0.000 | 1557.006 | 1894.369 |

```
. marginsplot, name(predicted_educ, replace) // plot educ predictions
Variables that uniquely identify margins: c1sess educgrp
```



```
. predict prededuc // save fixed-effect predicted outcomes
(option xb assumed)
```

```
. corr nm3rt prededuc // get total r to make r2
(obs=606)
```

| | nm3rt prededuc | |
|----------|----------------|--------|
| nm3rt | 1.0000 | |
| prededuc | 0.4151 | 1.0000 |

$R = .4151$, so R^2 for time+age+reas+educ = .172

The fixed effects of time, age, and reasoning before accounted for ~16.0% of the variance in RT, so there is a net increase of 1.2% due to education (which is not significant).

Simple Processing Speed – Example Conditional Models of Change Results

The extent to which individual differences in response time (RT) over six sessions for a simple processing speed test (number match three) could be predicted from baseline age, abstract reasoning, and education level was examined in a series of multilevel models (i.e., general linear mixed models) in which the six practice sessions were nested within each participant. Given the interest in comparing models differing in fixed effects, maximum likelihood (ML) was used in estimating and reporting all model parameters. The significance of new fixed effects were evaluated with individual Wald tests (i.e., of estimate / SE) as well as with likelihood ratio tests (i.e., $-2\Delta LL$), with degrees of freedom equal to the number of new fixed effects. Session (i.e., the index of time) was centered at the first occasion, age was centered at 80 years, abstract reasoning was centered at 22 (near the mean of the scale), and graduate-level education was the reference group for education level (with separate contrasts for high school or less and for bachelor's level education).

The best-fitting unconditional growth model specified quadratic decline across the six sessions (i.e., a decelerating negative function) with significant individual differences in the intercept, linear, and quadratic effects. Accordingly, effect size was evaluated via pseudo- R^2 values for the proportion reduction in each random effect variance, as well as with total R^2 , the squared correlation between the actual outcome values and the outcomes predicted by the model fixed effects. In the unconditional growth model, the fixed effects for linear and quadratic change across sessions accounted for approximately 4% of the total variation in RT.

Next, age was added as a predictor of the intercept, linear slope, and quadratic slope. The age model fit significantly better than the unconditional model as indicated by a significant likelihood ratio test, $-2\Delta LL(3) = 11.6$, $p = .009$; the AIC was lower, although the BIC was not. However, only the fixed effect of age on the intercept was significant, indicating that for every additional year of age above 80, RT at the first session was predicted to be significantly higher by 29.05 ($p = .001$). In terms of pseudo- R^2 , age accounted for 11.29% of the random intercept variance, 4.50% of the random linear slope variance, and 2.64% of the random quadratic slope variance. As expected given that baseline age is a time-invariant predictor, the residual variance was not reduced. The total cumulative R^2 from session and age was $R^2 = .11$, approximately a 7% increase due to age (which was significant, as indicated by the likelihood ratio test). Although the interactions of age with the linear and quadratic slopes were not significant, they were retained in the model to fully control for age effects before examining the effects of other predictors.

Abstract reasoning was then added as a predictor of the intercept, linear slope, and quadratic slope. The abstract reasoning model fit significantly better than the age model, $-2\Delta LL(3) = 12.5$, $p = .006$; the AIC was lower, although the BIC was not. However, only the fixed effect of reasoning on the intercept was significant. The nonsignificant effect of reasoning on the quadratic slope was then removed, revealing a significant effect of reasoning on both the intercept and linear slope, such that for every unit higher reasoning above 22, RT at the first session was expected to be lower by 32.82 and the linear rate of improvement in RT (as evaluated at the first session given the quadratic slope) was expected to be less negative by 2.94 (i.e., faster initial RT with less improvement in persons with greater reasoning). These two effects still resulted in a significant improvement in model fit over the age model, $-2\Delta LL(2) = 10.1$, $p = .006$, with a lower AIC and BIC. Reasoning accounted for 5.68% of the random intercept variance but had no measurable reduction of the random linear and quadratic slope variances. The total cumulative R^2 from session, age, and reasoning was $R^2 = .16$, approximately a 5% increase due to reasoning (which was significant, as indicated by the likelihood ratio test).

Finally, education level (high school or less, bachelor's level, or graduate level) was then added as a predictor of the intercept, linear slope, and quadratic slope. The education model did not fit significantly better than the reasoning model, $-2\Delta LL(6) = 4.7$, $p = .583$, with a higher AIC and BIC. None of the omnibus main effects of group on the intercept, linear, or quadratic slopes were significant, $\chi^2(2) < 1.92$, p 's $> .05$, and none of the pairwise group comparisons were significant as well. Education accounted for 0.04% of the random intercept variance, 2.95% of the random linear slope variance, and 4.01% of the random quadratic slope variance. The total cumulative R^2 from session, age, reasoning, and education was $R^2 = .17$, approximately a 1% increase due to education (which was not significant, as indicated by the likelihood ratio test).

(From here one might remove nonsignificant model effects and/or add other effects as needed to fully answer all research questions...)

Example 4: Examining BP and WP Effects of Negative Mood Predicting Next-Morning Glucose (complete data, syntax, and output available for SAS, SPSS, and STATA electronically)

These data were simulated loosely based on real data reported in the citation below. The daily diary study followed persons with Type II diabetes for 21 consecutive days to examine within-person relationships between mood, stress, and morning glucose (an index of how well-controlled the diabetes is). Here we will examine between-person and within-person relationships between daily negative mood and glucose the next morning (which was log-transformed given skewness) and how these relationships are moderated by sex.

Skaff, M., Mullan J., Fisher, L., Almeida, D., Hoffman, L., Masharani, U., & Mohr, D. (2009). [Effects of mood on daily fasting glucose in Type 2 Diabetes](#). *Health Psychology*, 28(3), 265-272.

SAS Data Setup:

```
* SAS code to read data into work library and center predictors;
DATA work.example4; SET filepath.example6;
* Level-2 effect of Negative Mood (mean=0, SD=1);
  PMnm0 = PMnegmood - 0; LABEL PMnm0= "PMnm0: Person Mean Negative Mood (0=0)";
* Level-1 effect to use with PERSON-MEAN-CENTERING;
  WPnm = negmood - PMnegmood; LABEL wpnm= "WPnm: Within-Person Negative Mood (0=PM)";
* Level-1 effect to use with GRAND-MEAN-CENTERING;
  TVnm0 = negmood - 0; LABEL TVnm0= "TVnm0: Time-Varying Negative Mood (0=0)";
RUN;
```

SPSS Data Setup:

```
* SPSS code to import data and center predictors.
GET FILE = "example/Example6.sav".
DATASET NAME example6 WINDOW=FRONT.
COMPUTE PMnm0 = PMnegmood - 0.
COMPUTE WPnm = negmood - PMnegmood.
COMPUTE TVnm0 = negmood - 0.
VARIABLE LABELS
  PMnm0 "PMnm0: Person Mean Negative Mood (0=0)"
  WPnm "WPnm: Within-Person Negative Mood (0=PM)"
  TVnm0 "TVnm0: Time-Varying Negative Mood (0=0)".
EXECUTE.
```

STATA Data Setup:

```
* STATA code to center predictors
* level-2 effect of negative mood
gen PMnm0 = PMnegmood - 0
label variable PMnm0 "PMnm0: Person Mean Negative Mood (0=0)"
* level-1 effect to use with PERSON-MEAN-CENTERING
gen WPnm = negmood - PMnegmood
label variable WPnm "WPnm: Within-Person Negative Mood (0=PM)"
* level-1 effect to use with GRAND-MEAN-CENTERING
gen TVnm0 = negmood - 0
label variable TVnm0 "TVnm0: Time-Varying Negative Mood (0=0)"
```


Model 1a. Empty Model for LN Morning Glucose (Daily Outcome)

```
TITLE "SAS Model 1a: Empty Model for Daily Glucose Outcome";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lglucAM = / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / VCORR SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC; RUN;
```

```
TITLE "SPSS Model 1a: Empty Model for Daily Glucose Outcome".
MIXED lglucAM BY ID day
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED =
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID).
```

$$\begin{aligned}\text{Level 1: } \text{Glucose}_{ti} &= \beta_{0i} + e_{ti} \\ \text{Level 2: } \beta_{0i} &= \gamma_{00} + U_{0i}\end{aligned}$$

```
* STATA Model 1a: Empty Model for Daily Glucose Outcome
xtmixed lglucAM , || id: , variance ml covariance(un) residuals(independent,t(day)),
  estimates store empty // save LL for LRT
```

STATA output:

```
Mixed-effects ML regression                Number of obs      =      4140
Group variable: id                        Number of groups   =       207
                                           Obs per group: min =        20
                                           avg =       20.0
                                           max =        20
                                           Wald chi2(0)      =         .
                                           Prob > chi2       =         .
```

| lglucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------------------------|----------|-----------|--------|-------|----------------------|----------|
| -----+----- | | | | | | |
| _cons | 4.942683 | .0181322 | 272.59 | 0.000 | 4.907145 | 4.978221 |
| -----+----- | | | | | | |
| Random-effects Parameters | Estimate | Std. Err. | | | [95% Conf. Interval] | |
| -----+----- | | | | | | |
| id: Identity | | | | | | |
| var(_cons) | .0665423 | .0066897 | | | .0546417 | .0810348 |
| -----+----- | | | | | | |
| var(Residual) | .0302851 | .0006829 | | | .0289757 | .0316537 |
| -----+----- | | | | | | |

Calculate the ICC for the glucose outcome:

$$ICC = \frac{.06654}{.06654 + .03029} = .69$$

This LR test tells us that the random intercept variance is significantly greater than 0, and thus so is the ICC.

```
LR test vs. linear regression: chibar2(01) = 4024.09 Prob >= chibar2 = 0.0000
Covariance Parameter Estimates
```

Model 1b. Empty Model for Negative Mood (Daily Predictor)

```
TITLE "SAS Model 1b: Empty Model for Daily Negative Mood Predictor";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL negmood = / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / VCORR SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC; RUN;
```

```
TITLE "SPSS Model 1b: Empty Model for Daily Negative Mood Predictor".
MIXED negmood BY ID day
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED =
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID).
```

$$\begin{aligned}\text{Level 1: } \text{Mood}_{ti} &= \beta_{0i} + e_{ti} \\ \text{Level 2: } \beta_{0i} &= \gamma_{00} + U_{0i}\end{aligned}$$

```
* STATA Model 1b: Empty Model for Daily Negative Mood Predictors
xtmixed negmood , || id: , ///
  variance ml covariance(un) residuals(independent,t(day))
```

STATA output:

```

Wald chi2(0)      =      .
Prob > chi2       =      .

-----+-----
      negmood |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _cons |   .1597403   .0418067    3.82   0.000    .0778007    .24168

-----+-----
Random-effects Parameters   |   Estimate   Std. Err.     [95% Conf. Interval]
-----+-----
id: Identity
      var(_cons) |   .3355036   .0355674     .2725584    .4129855
-----+-----
      var(Residual) |   .525824   .0118575     .5030898    .5495855
-----+-----
LR test vs. linear regression: chibar2(01) =   1500.40 Prob >= chibar2 = 0.0000

```

Calculate the ICC for the mood predictor:

$$ICC = \frac{.3355}{.3355 + .5258} = .39$$

This LR test tells us that the random intercept variance is significantly greater than 0, and thus so is the ICC.

Model 2a. Fixed Effects of Negative Mood using Person-Mean-Centering (PMC)

Level 1: $Glucose_{ti} = \beta_{0i} + \beta_{1i} (Mood_{ti} - \overline{Mood_i}) + e_{ti}$

Level 2: Intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01} (\overline{Mood_i} - 0) + U_{0i}$

Within-Person Mood: $\beta_{1i} = \gamma_{10}$

```

TITLE "SAS Model 2a: Fixed Effects of Negative Mood using PMC";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lglucAM = WPnm PMnm0 / SOLUTION DDFM=Satterthwaite OUTPM=MoodPred;
  RANDOM INTERCEPT / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC;
  ESTIMATE "Within-Person Mood Effect"      WPnm 1;
  ESTIMATE "Between-Person Mood Effect"     PMnm0 1;
  ESTIMATE "Contextual Mood Effect"         PMnm0 1 WPnm -1;
RUN; PROC CORR NOSIMPLE DATA=MoodPred; VAR lglucAM pred; RUN;

```

```

TITLE "SPSS Model 2a: Fixed Effects of Negative Mood using PMC".
MIXED lglucAM BY ID day WITH WPnm PMnm0
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = WPnm PMnm0
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID)
  /SAVE = FIXPRED (predmood)
  /TEST = "Within-Person Mood Effect"      WPnm 1
  /TEST = "Between-Person Mood Effect"     PMnm0 1
  /TEST = "Contextual Mood Effect"         PMnm0 1 WPnm -1.
CORRELATIONS lglucAM predmood.

```

```

* STATA Model 2a: Fixed Effects of Negative Mood using PMC
xtmixed lglucAM c.WPnm c.PMnm0, || id: , ///
  variance ml covariance(un) residuals(independent,t(day)),
  estat ic, n(207),
  predict predmood, // save fixed-effect predicted outcomes
  estimates store FixWP, // save LL for LRT
  lrtest FixWP empty, // LRT against empty model
  lincom 1*c.WPnm // within-person mood effect
  lincom 1*c.PMnm0 // between-person mood effect
  lincom 1*c.PMnm0 - 1*c.WPnm // contextual mood effect
corr lglucAM predmood

```

STATA output:

```

Log likelihood =      978.269                Wald chi2(2)      =      15.20
                                                Prob > chi2        =      0.0005

```

| lgLucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------|----------|-----------|--------|-------|----------------------|----------|
| WPnm | .0109743 | .0038207 | 2.87 | 0.004 | .0034859 | .0184626 |
| PMnm0 | .0803976 | .030461 | 2.64 | 0.008 | .0206952 | .1401 |
| _cons | 4.930171 | .0184512 | 267.20 | 0.000 | 4.894008 | 4.966335 |

```

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

```

```

id: Identity
      var(_cons) |      .0643486   .0064737   .0528329   .0783743 → intercept var down 3.29%
      var(Residual) |      .0302214   .0006815   .0289147   .031587 → residual var down 0.23%

```

```

LR test vs. linear regression: chibar2(01) =   3941.45 Prob >= chibar2 = 0.0000

```

```

.      estat ic, n(207),

```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|-----------|-----------|
| . | 207 | . | 978.269 | 5 | -1946.538 | -1929.874 |

```

Note: N=207 used in calculating BIC

```

```

.      lrtest FixWP empty,           // LRT against empty model

```

```

Likelihood-ratio test                LR chi2(2) =      15.08
(Assumption: empty nested in FixWP)  Prob > chi2 =      0.0005

```

Is this a better model than the empty model (1a)—is the total R^2 significant?
Yes, $ML -2\Delta LL(2) = 15$,
 $p = .004$

```

.      lincom 1*c.WPnm                // within-person mood effect

```

| lgLucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------|----------|-----------|------|-------|----------------------|----------|
| (1) | .0109743 | .0038207 | 2.87 | 0.004 | .0034859 | .0184626 |

```

.      lincom 1*c.PMnm0                // between-person mood effect

```

| lgLucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------|----------|-----------|------|-------|----------------------|-------|
| (1) | .0803976 | .030461 | 2.64 | 0.008 | .0206952 | .1401 |

```

.      lincom 1*c.PMnm0 - 1*c.WPnm    // contextual mood effect

```

| lgLucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------|----------|-----------|------|-------|----------------------|---------|
| (1) | .0694233 | .0306963 | 2.26 | 0.024 | .0092597 | .129587 |

```

.      corr lgLucAM predmood

```

```

(obs=4140)

```

| | lgLucAM | predmood |
|----------|---------|----------|
| lgLucAM | 1.0000 | |
| predmood | 0.1527 | 1.0000 |

Total R^2 from mood = .023

What does the level-1 effect (WPnm) represent in this model?

The level-1 effect is the within-person effect of negative mood. For every unit relative increase in your own negative mood that day, that next day's glucose goes up by .01097 (WP relation among daily levels).

What does the level-2 effect (PMnm0) represent in this model?

The level-2 effect is the between-person effect of negative mood. For every unit higher person mean negative mood, mean glucose is higher by .08040 (BP relation among mean levels).

What does the “contextual mood effect” represent?

It is the difference in the between-person and within-person effects: the between-person mood effect is significantly greater than the within-person mood effect by .0694 (so convergence was not obtained). So after controlling for current negative mood, there is an incremental effect of .0694 per unit higher person mean negative mood.

Model 2b. Random Effect of WP Negative Mood under PMC

| |
|--|
| Level 1: $\text{Glucose}_{ti} = \beta_{0i} + \beta_{1i} (\text{Mood}_{ti} - \overline{\text{Mood}_i}) + e_{ti}$ Level 2: Intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01} (\overline{\text{Mood}_i} - 0) + U_{0i}$ Within-Person Mood: $\beta_{1i} = \gamma_{10} + U_{1i}$ |
|--|

```

TITLE "SAS Model 2b: Random Effect of WP Negative Mood using PMC";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lglucAM = WPnm PMnm0 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT WPnm / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC;
  ESTIMATE "Within-Person Mood Effect"      WPnm    1;
  ESTIMATE "Between-Person Mood Effect"     PMnm0   1;
  ESTIMATE "Contextual Mood Effect"         PMnm0   1 WPnm -1; RUN;

```

```

TITLE "SPSS Model 2b: Random Effect of WP Negative Mood using PMC".

```

```

MIXED lglucAM BY ID day WITH WPnm PMnm0
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = WPnm PMnm0
  /RANDOM = INTERCEPT WPnm | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID)
  /TEST = "Within-Person Mood Effect"      WPnm    1
  /TEST = "Between-Person Mood Effect"     PMnm0   1
  /TEST = "Contextual Mood Effect"         PMnm0   1 WPnm -1.

```

```

* STATA Model 2b: Random Effect of WP Negative Mood using PMC
xtmixed lglucAM c.WPnm c.PMnm0, || id: WPnm, ///
  variance ml covariance(un) residuals(independent,t(day)),
  estat ic, n(207),
  estimates store RandWP,
  lrtest RandWP FixWP,
  lincom 1*c.WPnm           // within-person mood effect
  lincom 1*c.PMnm0          // between-person mood effect
  lincom 1*c.PMnm0 - 1*c.WPnm // contextual mood effect

```

STATA output:

| | | | | | | |
|----------------------------|----------|----------------------|--------|----------------------|----------------------|----------|
| Log likelihood = 979.72265 | | Wald chi2(2) = 14.03 | | Prob > chi2 = 0.0009 | | |
| ----- | | | | | | |
| lglucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
| -----+----- | | | | | | |
| WPnm | .0110375 | .0041371 | 2.67 | 0.008 | .0029288 | .0191462 |
| PMnm0 | .0802152 | .030471 | 2.63 | 0.008 | .0204931 | .1399372 |
| _cons | 4.930206 | .0184585 | 267.10 | 0.000 | 4.894028 | 4.966384 |
| ----- | | | | | | |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|-----------|-----------|----------------------|-------------------------------------|
| id: Unstructured | | | | |
| var(WPnm) | .0005056 | .0003348 | .0001381 | .0018515 random WPnm slope variance |
| var(_cons) | .0644045 | .0064789 | .0528795 | .0784413 random intercept variance |
| cov(WPnm,_cons) | -.0002049 | .001067 | -.0022962 | .0018865 int-WPnm slope covariance |
| var(Residual) | .029953 | .000692 | .0286269 | .0313406 |

LR test vs. linear regression: $\chi^2(3) = 3944.36$ Prob > $\chi^2 = 0.0000$

Note: LR test is conservative and provided only for reference.

. estat ic, n(207),

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|-----------|-----------|
| . | 207 | . | 979.7227 | 7 | -1945.445 | -1922.116 |

Note: N=207 used in calculating BIC

. estimates store RandWP, // save LL for LRT
. lrtest RandWP FixWP, // LRT against fixed effect

Likelihood-ratio test LR $\chi^2(2) = 2.91$
(Assumption: FixWP nested in RandWP) Prob > $\chi^2 = 0.2337$

Is this a better model than the fixed effects model (2a)? How do we know?

No, $ML - 2\Delta LL(2) = 2.91$, $p = .235$

Each person does not need his or her own effect of worse negative mood than usual.

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

Model 2c. Adding Moderation Effects by Sex (0=M, 1=F) for Each Mood Effect under PMC

$$\begin{aligned} \text{Level 1: } \text{Glucose}_{it} &= \beta_{0i} + \beta_{1i} (\text{Mood}_{it} - \overline{\text{Mood}_i}) + e_{it} \\ \text{Level 2: } \text{Intercept: } \beta_{0i} &= \gamma_{00} + \gamma_{01} (\overline{\text{Mood}_i} - 0) + \gamma_{02} (\text{Woman}_i) + \gamma_{03} (\overline{\text{Mood}_i} - 0)(\text{Woman}_i) + U_{0i} \\ \text{Within-Person Mood: } \beta_{1i} &= \gamma_{10} + \gamma_{12} (\text{Woman}_i) \end{aligned}$$

```
TITLE "SAS Model 2c: Fixed Effects of Sex (0=M, 1=F) by PMC Negative Mood";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lglucAM = WPnm PMnm0 sexmf WPnm*sexmf PMnm0*sexmf
    / SOLUTION DDFM=Satterthwaite OUTPM=SexPred;
  RANDOM INTERCEPT / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC;

ESTIMATE "Intercept: Men (Mood=0)" intercept 1 sexmf 0;
ESTIMATE "Intercept: Women (Mood=0)" intercept 1 sexmf 1;
ESTIMATE "Intercept: Women Diff (Mood=0)" sexmf 1;
ESTIMATE "Within-Person Mood Effect: Men" WPnm 1 WPnm*sexmf 0;
ESTIMATE "Within-Person Mood Effect: Women" WPnm 1 WPnm*sexmf 1;
ESTIMATE "Within-Person Mood Effect: Women Diff" WPnm*sexmf 1;
ESTIMATE "Between-Person Mood Effect: Men" PMnm0 1 PMnm0*sexmf 0;
ESTIMATE "Between-Person Mood Effect: Women" PMnm0 1 PMnm0*sexmf 1;
ESTIMATE "Between-Person Mood Effect: Women Diff" PMnm0*sexmf 1;
ESTIMATE "Contextual Mood Effect: Men" PMnm0 1 PMnm0*sexMF 0 WPnm -1 WPnm*sexMF 0;
ESTIMATE "Contextual Mood Effect: Women" PMnm0 1 PMnm0*sexMF 1 WPnm -1 WPnm*sexMF -1;
ESTIMATE "Contextual Mood Effect: Women Diff" PMnm0*sexMF 1 WPnm*sexMF -1;
RUN; PROC CORR NOSIMPLE DATA=SexPred; VAR lglucAM pred; RUN;
```

TITLE "SPSS Model 2c: Fixed Effects of Sex (0=M, 1=F) by PMC Negative Mood".

```
MIXED lglucAM BY ID day WITH WPnm PMnm0 sexmf
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = WPnm PMnm0 sexmf WPnm*sexmf PMnm0*sexmf
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
```

```

/REPEATED = day | SUBJECT(ID) COVTYPE(ID)
/SAVE = FIXPRED (predsex)
/TEST = "Intercept: Men (Mood=0)"          intercept 1 sexmf 0
/TEST = "Intercept: Women (Mood=0)"        intercept 1 sexmf 1
/TEST = "Intercept: Women Diff (Mood=0)"    sexmf 1
/TEST = "Within-Person Mood Effect: Men"    WPnm 1 WPnm*sexmf 0
/TEST = "Within-Person Mood Effect: Women"  WPnm 1 WPnm*sexmf 1
/TEST = "Within-Person Mood Effect: Women Diff" WPnm*sexmf 1
/TEST = "Between-Person Mood Effect: Men"    PMnm0 1 PMnm0*sexmf 0
/TEST = "Between-Person Mood Effect: Women"  PMnm0 1 PMnm0*sexmf 1
/TEST = "Between-Person Mood Effect: Women Diff" PMnm0*sexmf 1
/TEST = "Contextual Mood Effect: Men"        PMnm0 1 PMnm0*sexMF 0 WPnm -1 WPnm*sexMF 0
/TEST = "Contextual Mood Effect: Women"      PMnm0 1 PMnm0*sexMF 1 WPnm -1 WPnm*sexMF -1
/TEST = "Contextual Mood Effect: Women Diff" PMnm0*sexMF 1 WPnm*sexMF -1.
CORRELATIONS lglucAM predsex.

* STATA Model 2c: SPSS Model 2c: Fixed Effects of Sex (0=M, 1=F) by PMC Negative Mood
xtmixed lglucAM c.WPnm c.PMnm0 c.sexmf c.WPnm#c.sexmf c.PMnm0#c.sexmf, ///
|| id: , variance ml covariance(un) residuals(independent,t(day)),
estat ic, n(207),
estimates store Sexeffects, // save LL for LRT
lrtest Sexeffects FixWP, // LRT against main effects model
predict predsex, // save fixed-effect predicted outcomes
lincom 1*_cons + 0*c.sexmf // intercept: men (mood=0)
lincom 1*_cons + 1*c.sexmf // intercept: women (mood=0)
lincom 1*c.sexmf // intercept: women diff (mood=0)
lincom 1*c.WPnm + 0*c.WPnm#c.sexmf // within-person mood effect: men
lincom 1*c.WPnm + 1*c.WPnm#c.sexmf // within-person mood effect: women
lincom 1*c.WPnm#c.sexmf // within-person mood effect: women diff
lincom 1*c.PMnm0 + 0*c.PMnm0#c.sexmf // between-person mood effect: men
lincom 1*c.PMnm0 + 1*c.PMnm0#c.sexmf // between-person mood effect: women
lincom 1*c.PMnm0#c.sexmf // between-person mood effect: women diff
lincom 1*c.PMnm0 + 0*PMnm0#c.sexmf - 1*c.WPnm + 0*c.WPnm#c.sexmf // contextual mood: men
lincom 1*c.PMnm0 + 1*pmnm0#c.sexmf - 1*c.WPnm - 1*c.WPnm#c.sexmf // contextual mood: women
lincom 1*c.PMnm0#c.sexmf - 1*WPnm#c.sexmf // contextual mood: women diff
margins, at(c.WPnm=(-1 0 1) c.PMnm0=(-1 1) c.sexmf=(0 1)) vsquish // create predicted values
marginsplot, noci name(predicted_mood, replace) xdimension(WPnm) // plot predicted, no CI
corr lglucAM predsex

```

STATA output:

```

Log likelihood = 994.02512      Wald chi2(5)      =      47.55
                               Prob > chi2      =      0.0000
-----+-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      WPnm |      .0311885   .0059366     5.25   0.000     .0195529   .0428241
      PMnm0 |      .1996279   .0484871     4.12   0.000     .104595   .2946608
      sexmf |     -.0361935   .0362613    -1.00   0.318    -.1072643   .0348772
c.WPnm#c.sexmf |     -.0344341   .0077425    -4.45   0.000    -.0496092  -.019259
c.PMnm0#c.sexmf |     -.184933   .0613487    -3.01   0.003    -.3051743  -.0646918
      _cons |      4.953854   .0273373    181.21  0.000     4.900274   5.007434
-----+-----

Random-effects Parameters |   Estimate  Std. Err.      [95% Conf. Interval]
-----+-----
id: Identity              |
      var(_cons) |      .0607399   .0061183     .0498578   .0739972 → intercept var down by 5.61%
-----+-----
      var(Residual) |      .0300694   .0006781     .0287694   .0314282 → residual var down by 0.50%
-----+-----

LR test vs. linear regression: chibar2(01) = 3804.78 Prob >= chibar2 = 0.0000
.      estat ic, n(207),
-----+-----
      Model |   Obs    ll(null)    ll(model)    df        AIC        BIC
-----+-----
      . |   207          .    994.0251      8    -1972.05    -1945.388

```

```
. estimates store Sexeffects, // save LL for LRT
. lrtest Sexeffects FixWP, // LRT against main effects model
```

```
Likelihood-ratio test LR chi2(3) = 31.51
(Assumption: FixWP nested in Sexeffects) Prob > chi2 = 0.0000
```

Is this a better model than the fixed main effects model (2b)?
Yes, $ML - 2\Delta LL(3) = 31$,
 $p < .001$

What does the intercept now represent in this model?

The intercept of 4.9539 is the expected glucose for a man with a $PMnm = 0$ and $WPnm = 0$.

What does the level-1 effect (WPnm) represent in this model?

The level-1 effect is the simple within-person effect of negative mood specifically for a man. For every unit relative increase in your own negative mood that day, that next day's glucose goes up by 0.03119 (significant).

What does the level-2 effect (PMnm0) represent in this model?

The level-2 effect is the simple between-person effect of negative mood specifically for a man. For every unit increase in your person mean negative mood, mean glucose is higher by 0.1996 (significant).

What does the main effect of sex represent in this model?

The simple effect of sex is the difference between men and women for someone with a person mean negative mood of 0 on day when they are at their mean. In those persons, women are -0.03619 lower in mean glucose (n.s.).

What does the WPnm*Sex interaction represent in this model?

The $WP*Sex$ interaction tells us that the WP mood effect is 0.03443 smaller in women (significant interaction).

What does the PMnm0*Sex interaction represent in this model?

The $BP*Sex$ interaction tells us BP mood effect is 0.1849 smaller in women (significant interaction).

Which effects are not directly given by the model?

The effects for women and all of the contextual effects, as shown below.

```
. lincom 1*c.WPnm + 1*c.WPnm#c.sexmf // within-person mood effect: women
-----
lglucAM | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
(1) | -.0032456 .0049702 -0.65 0.514 -.0129871 .0064959
-----

. lincom 1*c.PMnm0 + 1*c.PMnm0#c.sexmf // between-person mood effect: women
-----
lglucAM | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
(1) | .0146949 .0375854 0.39 0.696 -.0589712 .088361
-----

. lincom 1*c.PMnm0 + 0*c.PMnm0#c.sexmf - 1*c.WPnm + 0*c.WPnm#c.sexmf // contextual mood: men
-----
lglucAM | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
(1) | .1684394 .0488639 3.45 0.001 .0726679 .2642109
-----

. lincom 1*c.PMnm0 + 1*c.PMnm0#c.sexmf - 1*c.WPnm - 1*c.WPnm#c.sexmf // contextual mood: women
-----
lglucAM | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
(1) | .0179405 .0378969 0.47 0.636 -.0563361 .0922171
-----

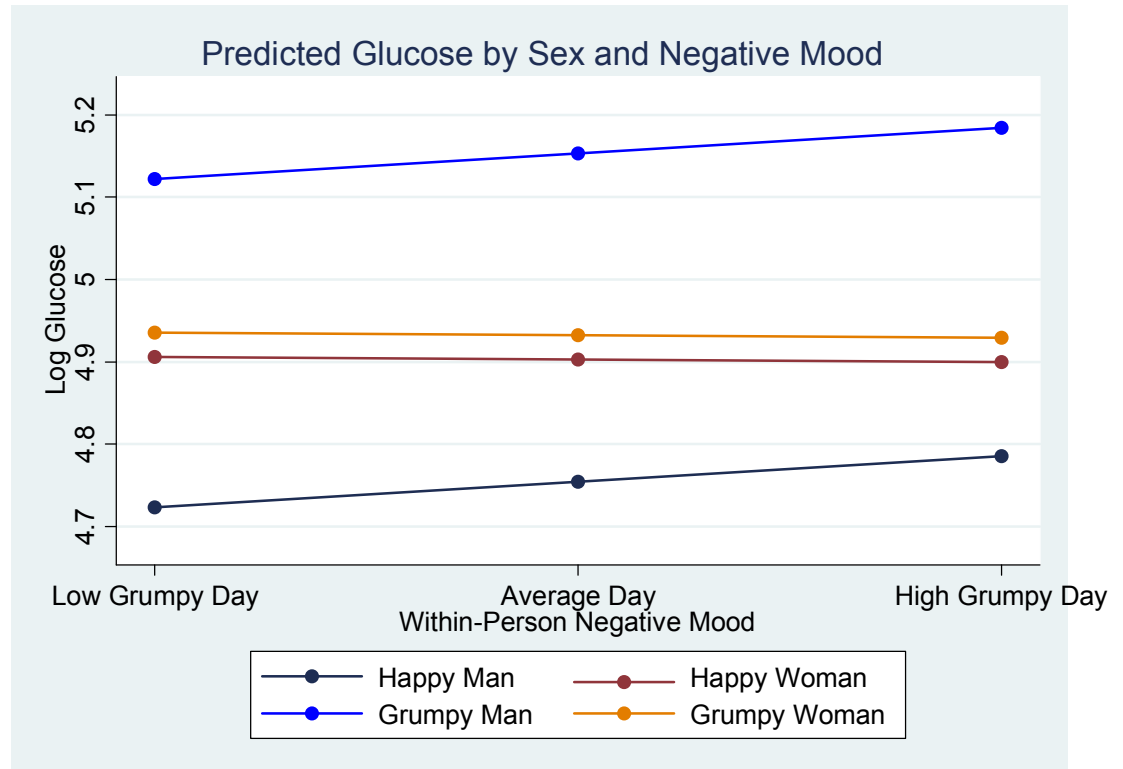
. lincom 1*c.PMnm0#c.sexmf - 1*c.WPnm#c.sexmf // contextual mood: women diff
-----
lglucAM | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
(1) | -.1504989 .0618374 -2.43 0.015 -.2716979 -.0293
-----

. margins, at(c.WPnm=(-1 0 1) c.PMnm0=(-1 1) c.sexmf=(0 1)) vsquish // create predicted values
```


Expression : Linear prediction, fixed portion, predict()

| | | | | |
|--------|---|-------|---|----|
| 1._at | : | WPnm | = | -1 |
| | | PMnmO | = | -1 |
| | | sexmf | = | 0 |
| 2._at | : | WPnm | = | -1 |
| | | PMnmO | = | -1 |
| | | sexmf | = | 1 |
| 3._at | : | WPnm | = | -1 |
| | | PMnmO | = | 1 |
| | | sexmf | = | 0 |
| 4._at | : | WPnm | = | -1 |
| | | PMnmO | = | 1 |
| | | sexmf | = | 1 |
| 5._at | : | WPnm | = | 0 |
| | | PMnmO | = | -1 |
| | | sexmf | = | 0 |
| 6._at | : | WPnm | = | 0 |
| | | PMnmO | = | -1 |
| | | sexmf | = | 1 |
| 7._at | : | WPnm | = | 0 |
| | | PMnmO | = | 1 |
| | | sexmf | = | 0 |
| 8._at | : | WPnm | = | 0 |
| | | PMnmO | = | 1 |
| | | sexmf | = | 1 |
| 9._at | : | WPnm | = | 1 |
| | | PMnmO | = | -1 |
| | | sexmf | = | 0 |
| 10._at | : | WPnm | = | 1 |
| | | PMnmO | = | -1 |
| | | sexmf | = | 1 |
| 11._at | : | WPnm | = | 1 |
| | | PMnmO | = | 1 |
| | | sexmf | = | 0 |
| 12._at | : | WPnm | = | 1 |
| | | PMnmO | = | 1 |
| | | sexmf | = | 1 |

Figure 1



| | Delta-method | | | | | |
|-----|--------------|-----------|--------|-------|----------------------|----------|
| | Margin | Std. Err. | z | P> z | [95% Conf. Interval] | |
| _at | | | | | | |
| 1 | 4.723038 | .0600243 | 78.69 | 0.000 | 4.605392 | 4.840683 |
| 2 | 4.906211 | .0505436 | 97.07 | 0.000 | 4.807148 | 5.005275 |
| 3 | 5.122294 | .0516162 | 99.24 | 0.000 | 5.021128 | 5.22346 |
| 4 | 4.935601 | .0381615 | 129.33 | 0.000 | 4.860806 | 5.010396 |
| 5 | 4.754226 | .059742 | 79.58 | 0.000 | 4.637134 | 4.871319 |
| 6 | 4.902966 | .0502812 | 97.51 | 0.000 | 4.804416 | 5.001515 |
| 7 | 5.153482 | .0512596 | 100.54 | 0.000 | 5.053015 | 5.253949 |
| 8 | 4.932356 | .0378448 | 130.33 | 0.000 | 4.858181 | 5.00653 |
| 9 | 4.785415 | .0600482 | 79.69 | 0.000 | 4.667723 | 4.903107 |
| 10 | 4.89972 | .0505089 | 97.01 | 0.000 | 4.800725 | 4.998716 |
| 11 | 5.184671 | .0515882 | 100.50 | 0.000 | 5.08356 | 5.285782 |
| 12 | 4.92911 | .0381781 | 129.11 | 0.000 | 4.854282 | 5.003938 |

```
. corr lglucAM predsex
(obs=4140)
```

| | lgglucAM | predsex |
|----------|----------|---------|
| lgglucAM | 1.0000 | |
| predsex | 0.2493 | 1.0000 |

Total R^2 from mood+sex = .062, for a net increase of .039 from sex effects

Model 3. Predicting Glucose from Time-Varying Negative Mood only (GMC):

$$\text{Level 1: Glucose}_{ti} = \beta_{0i} + \beta_{1i} (\text{Mood}_{ti} - 0) + e_{ti}$$

$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$

$$\text{Time-Varying Mood: } \beta_{1i} = \gamma_{10}$$

```

TITLE "SAS Model 3: Fixed Effect of TV Negative Mood only using GMC";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lglucAM = TVnm0 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC; RUN;

TITLE "SPSS Model 3: Fixed Effect of TV Negative Mood only using GMC".
MIXED lglucAM BY ID day WITH TVnm0
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = TVnm0
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID).

* STATA Model 3: Fixed Effect of TV Negative Mood only using GMC
xtmixed lglucAM c.TVnm0, || id: , variance ml covariance(un) residuals(independent,t(day)),
  estat ic, n(207)

```

STATA output:

```

Log likelihood = 975.74178      Wald chi2(1)      =      10.04
                               Prob > chi2      =      0.0015

-----+-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      TVnm0 |   .0120181   .0037921     3.17   0.002     .0045856   .0194505
      _cons |   4.940763   .0180634    273.52   0.000     4.90536   4.976167
-----+-----

Random-effects Parameters |   Estimate  Std. Err.   [95% Conf. Interval]
-----+-----
id: Identity              |
      var(_cons) |   .065954   .0066337     .0541534   .0803259 → intercept var down by 0.89%
-----+-----
      var(Residual) |   .0302219   .0006815     .0289152   .0315876 → residual var down by 0.23%
-----+-----

LR test vs. linear regression: chibar2(01) = 3982.05 Prob >= chibar2 = 0.0000

.      estat ic, n(207)

-----+-----
      Model |   Obs    ll(null)   ll(model)    df       AIC       BIC
-----+-----
      . |   207          .    975.7418     4   -1943.484   -1930.153
-----+-----

Note: N=207 used in calculating BIC

```

What does the effect of TVnm0 represent in this model?

It is the smushed (conflated, convergence) effect of mood.

Model 3a. Fixed Effects of Negative Mood using Grand-Mean-Centering (GMC)

$$\begin{aligned} \text{Level 1: } \text{Glucose}_{ti} &= \beta_{0i} + \beta_{1i} (\text{Mood}_{ti} - 0) + e_{ti} \\ \text{Level 2: } \text{Intercept: } \beta_{0i} &= \gamma_{00} + \gamma_{01} (\overline{\text{Mood}_i} - 0) + U_{0i} \\ \text{Time-Varying Mood: } \beta_{1i} &= \gamma_{10} \end{aligned}$$

```

TITLE "SAS Model 3a: Fixed Effects of Negative Mood using GMC";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lgLucAM = TVnm0 PMnm0 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC;
  ESTIMATE "Within-Person Mood Effect"      TVnm0 1;
  ESTIMATE "Between-Person Mood Effect"     TVnm0 1 PMnm0 1;
  ESTIMATE "Contextual Mood Effect"         PMnm0 1; RUN;

TITLE "SPSS Model 3a: Fixed Effects of Negative Mood using GMC".
MIXED lgLucAM BY ID day WITH TVnm0 PMnm0
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = TVnm0 PMnm0
  /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID)
  /TEST = "Within-Person Mood Effect"      TVnm0 1
  /TEST = "Between-Person Mood Effect"     TVnm0 1 PMnm0 1
  /TEST = "Contextual Mood Effect"         PMnm0 1.

* STATA Model 3a: Fixed Effects of Negative Mood using GMC
xtmixed lgLucAM c.TVnm0 c.PMnm0, || id: , ///
  variance ml covariance(un) residuals(independent,t(day)),
  estat ic, n(207),
  estimates store FixTV,          // save LL for LRT
  lincom 1*c.TVnm0                // within-person mood effect
  lincom 1*c.TVnm0 + 1*c.PMnm0    // between-person mood effect
  lincom 1*c.PMnm0                // contextual mood effect

```

STATA output:

```

Log likelihood =      978.269      Wald chi2(2)      =      15.20
                                Prob > chi2      =      0.0005

```

| lgLucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------|----------|-----------|--------|-------|----------------------|----------|
| TVnm0 | .0109743 | .0038207 | 2.87 | 0.004 | .0034859 | .0184626 |
| PMnm0 | .0694233 | .0306963 | 2.26 | 0.024 | .0092597 | .129587 |
| _cons | 4.930171 | .0184512 | 267.20 | 0.000 | 4.894008 | 4.966335 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|----------|-----------|----------------------|----------|
| id: Identity | | | | |
| var(_cons) | .0643486 | .0064737 | .0528329 | .0783743 |
| var(Residual) | .0302214 | .0006815 | .0289147 | .031587 |

LR test vs. linear regression: chibar2(01) = 3941.45 Prob >= chibar2 = 0.0000

```

.      estat ic, n(207),

```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|-----------|-----------|
| . | 207 | . | 978.269 | 5 | -1946.538 | -1929.874 |

Note: N=207 used in calculating BIC

Note that the fit is the same as model 2a (and thus the R² values are, too)

```

.      lincom 1*c.TVnm0                                // within-person mood effect
-----+-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   .0109743   .0038207    2.87   0.004    .0034859    .0184626
-----+-----
.      lincom 1*c.TVnm0 + 1*c.PMnm0                    // between-person mood effect
-----+-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   .0803976   .030461    2.64   0.008    .0206952    .1401
-----+-----
.      lincom 1*c.PMnm0                                // contextual mood effect
-----+-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   .0694233   .0306963    2.26   0.024    .0092597    .129587
-----+-----

```

What does the level-2 effect (PMnm0) represent in this model?

It is the difference in the between-person and within-person effects (the contextual effect): the between-person mood effect is significantly greater than the within-person mood effect by .0694 (so convergence was not obtained). In other words, after controlling for current negative mood, there is an incremental effect of .0694 per unit higher person mean negative mood.

Model 3b. Random Effect of TV Negative Mood under GMC

Level 1: $\text{Glucose}_{ti} = \beta_{0i} + \beta_{1i} (\text{Mood}_{ti} - 0) + e_{ti}$
 Level 2: Intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01} (\overline{\text{Mood}_i} - 0) + U_{0i}$
 Time-Varying Mood: $\beta_{1i} = \gamma_{10} + U_{1i}$

```

TITLE "SAS Model 3b: Random Effect of TV Negative Mood using GMC";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lglucAM = TVnm0 PMnm0 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT TVnm0 / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC;
  ESTIMATE "Within-Person Mood Effect"      TVnm0 1;
  ESTIMATE "Between-Person Mood Effect"     TVnm0 1 PMnm0 1;
  ESTIMATE "Contextual Mood Effect"         PMnm0 1; RUN;

TITLE "SPSS Model 3b: Random Effect of TV Negative Mood using GMC".
MIXED lglucAM BY ID day WITH TVnm0 PMnm0
  /METHOD = ML
  /PRINT  = SOLUTION TESTCOV
  /FIXED  = TVnm0 PMnm0
  /RANDOM  = INTERCEPT TVnm0 | SUBJECT(ID) COVTYPE(UN)
  /REPEATED = day | SUBJECT(ID) COVTYPE(ID)
  /TEST   = "Within-Person Mood Effect"      TVnm0 1
  /TEST   = "Between-Person Mood Effect"     TVnm0 1 PMnm0 1
  /TEST   = "Contextual Mood Effect"         PMnm0 1.

* STATA Model 3b: Random Effect of WP Negative Mood using GMC
xtmixed lglucAM c.TVnm0 c.PMnm0, || id: TVnm0, ///
  variance ml covariance(un) residuals(independent,t(day)),
  estat ic, n(207),
  estimates store RandTV,          // save LL for LRT
  lrtest RandTV FixTV,            // LRT against fixed effect
  lincom 1*c.TVnm0                // within-person mood effect
  lincom 1*c.TVnm0 + 1*c.PMnm0    // between-person mood effect
  lincom 1*c.PMnm0                // contextual mood effect

```

STATA output:

```

                                Wald chi2(2)      =      13.72
Log likelihood =      980.1989      Prob > chi2      =      0.0010

```

| lgLucAM | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------|----------|-----------|--------|-------|----------------------|----------|
| TVnm0 | .0110189 | .0041807 | 2.64 | 0.008 | .0028248 | .019213 |
| PMnm0 | .0701465 | .0306592 | 2.29 | 0.022 | .0100555 | .1302374 |
| _cons | 4.930203 | .0184342 | 267.45 | 0.000 | 4.894073 | 4.966333 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|-----------|-----------|----------------------|------------------------------------|
| id: Unstructured | | | | |
| var(TVnm0) | .0005787 | .0003394 | .0001833 | .0018268 TV mood slope variance |
| var(_cons) | .0639986 | .0064641 | .0525044 | .0780092 intercept variance |
| cov(TVnm0,_cons) | -.0003279 | .0010502 | -.0023863 | .0017305 int-mood slope covariance |
| var(Residual) | .0299214 | .0006904 | .0285984 | .0313056 |

```

LR test vs. linear regression:      chi2(3) = 3945.31      Prob > chi2 = 0.0000

```

Note: LR test is conservative and provided only for reference.

```

.      estat ic, n(207),

```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|-----------|-----------|
| . | 207 | . | 980.1989 | 7 | -1946.398 | -1923.069 |

Note: N=207 used in calculating BIC

```

.      estimates store RandTV,      // save LL for LRT
.      lrtest RandTV FixTV,      // LRT against fixed effect
Likelihood-ratio test      LR chi2(2) =      3.86
(Assumption: FixTV nested in RandTV)      Prob > chi2 =      0.1452

```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

Is this a better model than the fixed effects model (3a)? How do we know?

No, ML $-2\Delta LL(2) = 3.86$,
 $p = .145$

Each person does not need his or her own effect of worse negative mood than usual.

Model 3c. Adding Moderation Effects by Sex (0=M, 1=F) for Each Mood Effect under GMC

Level 1: $\text{Glucose}_{ti} = \beta_{0i} + \beta_{1i}(\text{Mood}_{ti} - 0) + e_{ti}$
 Level 2: Intercept: $\beta_{0i} = \gamma_{00} + \gamma_{01}(\overline{\text{Mood}}_i - 0) + \gamma_{02}(\text{Woman}_i) + \gamma_{03}(\overline{\text{Mood}}_i - 0)(\text{Woman}_i) + U_{0i}$
 Time-Varying Mood: $\beta_{1i} = \gamma_{10} + \gamma_{12}(\text{Woman}_i)$

```

TITLE "SAS Model 3c: Fixed Effects of Sex (0=M, 1=F) by GMC Negative Mood";
PROC MIXED DATA=work.example4 COVTEST NOCLPRINT NOITPRINT NAMELEN=100 IC METHOD=ML;
  CLASS ID day;
  MODEL lgLucAM = TVnm0 PMnm0 sexmf TVnm0*sexmf PMnm0*sexmf
    / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / SUBJECT=ID TYPE=UN;
  REPEATED day / SUBJECT=ID TYPE=VC;
ESTIMATE "Intercept: Men (Mood=0)"      intercept 1 sexmf 0;
ESTIMATE "Intercept: Women (Mood=0)"    intercept 1 sexmf 1;
ESTIMATE "Intercept: Women Diff (Mood=0)" sexmf 1;
ESTIMATE "Within-Person Effect: Men"    TVnm0 1 TVnm0*sexmf 0;
ESTIMATE "Within-Person Effect: Women"  TVnm0 1 TVnm0*sexmf 1;
ESTIMATE "Within-Person Effect: Women Diff" TVnm0*sexmf 1;
ESTIMATE "Between-Person Effect: Men"    TVnm0 1 TVnm0*sexmf 0 PMnm0 1 PMnm0*sexmf 0;
ESTIMATE "Between-Person Effect: Women"  TVnm0 1 TVnm0*sexmf 1 PMnm0 1 PMnm0*sexmf 1;
ESTIMATE "Between-Person Effect: Women Diff" TVnm0*sexmf 1 PMnm0*sexmf 1;
ESTIMATE "Contextual Effect: Men"        PMnm0 1 PMnm0*sexmf 0;
ESTIMATE "Contextual Effect: Women"      PMnm0 1 PMnm0*sexmf 1;
ESTIMATE "Contextual Effect: Women Diff" PMnm0*sexmf 1; RUN;

```

```

TITLE "SPSS Model 3c: Fixed Effects of Sex (0=M, 1=F) by GMC Negative Mood".
MIXED lgLucAM BY ID day WITH TVnm0 PMnm0 sexmf
    /METHOD = ML
    /PRINT = SOLUTION TESTCOV
    /FIXED = TVnm0 PMnm0 sexmf TVnm0*sexmf PMnm0*sexmf
    /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
    /REPEATED = day | SUBJECT(ID) COVTYPE(ID)
/TEST = "Intercept: Men (Mood=0)"           intercept 1 sexmf 0
/TEST = "Intercept: Women (Mood=0)"         intercept 1 sexmf 1
/TEST = "Intercept: Women Diff (Mood=0)"    sexmf 1
/TEST = "Within-Person Mood Effect: Men"    TVnm0 1 TVnm0*sexmf 0
/TEST = "Within-Person Mood Effect: Women"  TVnm0 1 TVnm0*sexmf 1
/TEST = "Within-Person Mood Effect: Women Diff" TVnm0*sexmf 1
/TEST = "Between-Person Mood Effect: Men"    TVnm0 1 TVnm0*sexmf 0 PMnm0 1 PMnm0*sexmf 0
/TEST = "Between-Person Mood Effect: Women"  TVnm0 1 TVnm0*sexmf 1 PMnm0 1 PMnm0*sexmf 1
/TEST = "Between-Person Mood Effect: Women Diff" TVnm0*sexmf 1 PMnm0*sexmf 1
/TEST = "Contextual Mood Effect: Men"        PMnm0 1 PMnm0*sexMF 0
/TEST = "Contextual Mood Effect: Women"      PMnm0 1 PMnm0*sexMF 1
/TEST = "Contextual Mood Effect: Women Diff" PMnm0*sexMF 1.

* STATA Model 3c: SPSS Model 2c: Fixed Effects of Sex (0=M, 1=F) by GMC Negative Mood
xtmixed lgLucAM c.TVnm0 c.PMnm0 c.sexmf c.TVnm0#c.sexmf c.PMnm0#c.sexmf, ///
    || id: , variance ml covariance(un) residuals(independent,t(day)),
    estat ic, n(207),
lincom 1*_cons + 0*c.sexmf // intercept: men (mood=0)
lincom 1*_cons + 1*c.sexmf // intercept: women (mood=0)
lincom 1*c.sexmf // intercept: women diff (mood=0)
lincom 1*c.TVnm0 + 0*c.TVnm0#c.sexmf // within-person mood effect: men
lincom 1*c.TVnm0 + 1*c.TVnm0#c.sexmf // within-person mood effect: women
lincom 1*c.TVnm0#c.sexmf // within-person mood effect: women diff
lincom 1*c.TVnm0 + 0*c.TVnm0#c.sexmf + 1*c.PMnm0 + 0*c.PMnm0#c.sexmf // between-person: men
lincom 1*c.TVnm0 + 1*c.TVnm0#c.sexmf + 1*c.PMnm0 + 1*c.PMnm0#c.sexmf // between-person: women
lincom 1*c.TVnm0#c.sexmf + 1*c.PMnm0#c.sexmf // between-person: women diff
lincom 1*c.PMnm0 + 0*c.PMnm0#c.sexmf // contextual mood effect: men
lincom 1*c.PMnm0 + 1*c.PMnm0#c.sexmf // contextual mood effect: women
lincom 1*c.PMnm0#c.sexmf // contextual mood effect: women diff

```

STATA output (and non-directly provided estimates for simple effects):

```

Log likelihood = 994.02512      Wald chi2(5) = 47.55
                                Prob > chi2 = 0.0000
-----+-----
      lgLucAM |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      TVnm0 |   .0311885   .0059366     5.25  0.000   .0195529   .0428241
      PMnm0 |   .1684394   .0488639     3.45  0.001   .0726679   .2642109
      sexmf |  -.0361935   .0362613    -1.00  0.318  - .1072643   .0348772
c.TVnm0#c.sexmf | -.0344341   .0077425    -4.45  0.000  - .0496092  -.019259
c.PMnm0#c.sexmf | -.1504989   .0618374    -2.43  0.015  - .2716979  -.0293
      _cons |   4.953854   .0273373   181.21  0.000   4.900274   5.007434
-----+-----

Random-effects Parameters |   Estimate   Std. Err.      [95% Conf. Interval]
-----+-----
id: Identity
      var(_cons) |   .0607399   .0061183   .0498578   .0739972
-----+-----
      var(Residual) |   .0300694   .0006781   .0287694   .0314282
-----+-----
LR test vs. linear regression: chibar2(01) = 3804.78 Prob >= chibar2 = 0.0000

.      estat ic, n(207),
-----+-----
      Model |   Obs   ll(null)   ll(model)    df       AIC       BIC
-----+-----
      . |   207         .    994.0251     8    -1972.05    -1945.388
-----+-----

```

```

.      lincom 1*c.TVnm0 + 1*c.TVnm0#c.sexmf          // within-person mood effect: women
-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   -.0032456   .0049702   -0.65   0.514    -.0129871    .0064959
-----

.      lincom 1*c.TVnm0 + 0*c.TVnm0#c.sexmf + 1*c.PMnm0 + 0*c.PMnm0#c.sexmf    // between-person mood effect: men
-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   .1996279   .0484871    4.12   0.000    .104595    .2946608
-----

.      lincom 1*c.TVnm0 + 1*c.TVnm0#c.sexmf + 1*c.PMnm0 + 1*c.PMnm0#c.sexmf    // between-person mood effect: women
-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   .0146949   .0375854    0.39   0.696    -.0589712    .088361
-----

.      lincom 1*c.TVnm0#c.sexmf + 1*c.PMnm0#c.sexmf          // between-person mood effect: women diff
-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   -.184933   .0613487   -3.01   0.003    -.3051743    -.0646918
-----

.      lincom 1*c.PMnm0 + 1*c.PMnm0#c.sexmf          // contextual mood effect: women
-----
      lglucAM |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      (1) |   .0179405   .0378969    0.47   0.636    -.0563361    .0922171
-----

```


Sample Results Section (note the order of the models is different than what is in the handout):

The effects of negative mood and sex on next day's morning glucose level were examined in 207 persons with type-2 diabetes over a 20-day period. Glucose was natural log transformed (after adding 1 to each score) to improve normality. Intraclass correlations as calculated from an empty means, random intercept only model were .69 for glucose and .39 for negative mood, such that 69% and 39% of the variance in each variable was between persons, respectively. Preliminary analyses suggested that a random intercept only model for the variances of glucose over time had acceptable fit, and thus all conditional (predictor) models were examined using that structure as a baseline.

The time-varying (level-1) predictor for negative mood (left uncentered, given that 0 represented average level of the measure) was first entered into the model. A significant positive effect was obtained, such that higher daily levels of negative mood were related to higher daily levels of glucose. However, the inclusion of a single parameter for the effect of negative mood presumes that its between-person and within-person effects would be equivalent. This convergence hypothesis was tested explicitly by including person mean negative mood (also left uncentered, given that 0 represented average level of the original measure) as a level-2 predictor. The effect of person mean negative mood was significant, indicating that after controlling for absolute level of daily negative mood, persons with higher mean negative mood had higher mean glucose. Given that the significance of the level-2 effect also indicates that the between-person and within-person effects of negative mood were not equivalent, the model was re-specified to facilitate interpretation of these separate effects using group-mean-centering (i.e., person-mean-centering in longitudinal data). Specifically, a new level-1 predictor variable was created by subtracting each person's mean from daily negative mood, while the level-2 effect continued to be represented by the person mean. In this specification using person-mean-centering, the level-2 mean of negative mood represents the between-person effect directly and the level-1 within-person deviation of negative mood represents the within-person effect directly. Both the between- and within-person effects of negative mood were significantly positive. A random level-1 effect of negative mood was tested within both models, and was not found to be significant in either, $-2\Delta LL (\sim 2) < 5.14$, $p > .05$, indicating no significant individual differences in the within-person effect of negative mood.

Three effects of sex were then entered into the person-mean-centered model, including a main effect of sex and interactions with the between- and within-person effects of negative mood. The main effect of sex was non-significant, indicating no sex differences in mean glucose among persons with average levels of mean negative mood on average days (i.e., when average persons were at their mean). Given that both interactions were significant, however, results for both men and women will be presented as derived from ESTIMATE statements for the effects estimated specifically for each group within the overall model. Parameters for this final model are given in Table 1.

As shown, the intercept of 4.95 represents the expected morning LN glucose for a man with an average level of mean negative mood on an average day (i.e., both mean and person-mean-centered negative mood at 0). Men showed significant between- and within-person effects of negative mood, such that for every unit higher in mean negative mood, mean glucose was expected to be 0.20 higher (i.e., the between-person effect), and for every unit higher in negative mood on a given day relative to his own mean, glucose that next morning was expected to be 0.03 higher as well (i.e., the within-person effect). Thus, in men, being higher overall in negative mood and higher than usual in negative mood were each related to higher levels of glucose, and these effects were significantly different in magnitude (contextual effect = 0.17, SE = 0.05, $p < .001$). Said differently the contextual effect also indicates a significant contribution of person mean negative mood after controlling for daily negative mood.

As shown in Figure 1, however, these patterns were not found in women, as indicated by the significant interactions with sex. Specifically, the between-person and within-person effects of negative mood in women were 0.015 (SE = 0.038) and -0.003 (SE = 0.005), respectively. Neither effect was significant nor did they differ significantly in magnitude (contextual effect = 0.018, SE = .038). Both effects of negative mood were significantly smaller than in men (interaction terms of sex with between-person and within-person negative mood of -0.185 and -0.034, respectively). Finally, the contextual effect of negative mood, or the difference between the between-person and within-person effects of negative mood, was significantly larger for men (0.151, SE = 0.062, $p = .016$).

(Table 1 would have all parameter estimates from final model, see chapter 8 for examples)

(Figure 1 would show the within-person effect of negative mood for men and women with low or high mean negative mood – see plot for an example)

Example 5: Two-Level Clustered Data Example: Students within Schools (only syntax and output available for SAS, SPSS, and STATA electronically)

These are real data taken from the results of a math test given at the end of 10th grade in a Midwestern Rectangular State. These data include 13,802 students from 94 schools, with 31–515 students in each school ($M = 275$). We will examine how student free and reduced lunch status (0=pay for lunch, 1= receive free or reduced lunch) predicts math test scores.

SAS Code for Data Manipulation:

```
* Importing data into work library;
DATA work.grade10; SET example.grade10;
    * Selecting cases that are complete for analysis variables;
    WHERE NMISS(studentID, schoolID, frlunch, math)=0;
    LABEL studentID= "studentID: Student ID number"
           schoolID= "schoolID: School ID number"
           frlunch= "frlunch: 0=No, 1=Free/Reduced Lunch"
           math= "math: Math Test Score Outcome"; RUN;

* Getting school means to use as predictors;
PROC SORT DATA=work.grade10; BY schoolID studentID; RUN;
PROC MEANS NOPRINT N DATA= work.grade10;
    BY schoolID;
    VAR frlunch math;
    OUTPUT OUT=SchoolMeans
           MEAN(frlunch math)= SMfrlunch SMmath; RUN;

* Labeling new school mean variables;
DATA work.SchoolMeans; SET work.SchoolMeans;
    SchoolN = _FREQ_; * Saving N per school;
    DROP _TYPE_ _FREQ_; * Dropping unneeded SAS-created variables;
    LABEL SMfrlunch= "SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch"
           SMmath= "SMmath: School Mean Math Outcome"
           SchoolN= "SchoolN: # Students Contributing Data"; RUN;

* Merging school means back with individual data;
DATA work.grade10; MERGE work.grade10 work.SchoolMeans; BY schoolID;
    * Selecting only schools with data from at least 30 students;
    IF SchoolN < 31 THEN DELETE; RUN;

TITLE "Getting means to center predictors with";
PROC MEANS MEAN STDDEV MIN MAX DATA=work.grade10;
    VAR math frlunch SMmath SMfrlunch SchoolN; RUN; TITLE;

* Centering school mean predictors;
DATA work.grade10; SET work.grade10;
    SMfrlunch30 = SMfrlunch - .30; LABEL SMfrlunch30= "SMfrlunch30: 0=.30"; RUN;
```

SPSS Code for Data Manipulation:

```
* SPSS code to import data and create/center predictors.
DATASET NAME grade10 WINDOW=FRONT.
VARIABLE LABELS
    studentID "studentID: Student ID number"
    schoolID "schoolID: School ID number"
    frlunch "frlunch: 0=No, 1=Free/Reduced Lunch"
    math "math: Math Test Score".

* Selecting complete cases for analysis.
SELECT IF (NMISS(studentID, schoolID, frlunch, math)=0).
EXECUTE.

* Getting school means to use as level-2 predictors - SPSS 14+ can merge them back automatically.
SORT CASES BY schoolID studentID.
```

```

AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /PRESORTED
  /BREAK = schoolID
  /SMfrlunch = MEAN(frlunch)
  /SMmath = MEAN(math)
  /SchoolN = N.

* Labeling new school mean variables.
VARIABLE LABELS
  SMfrlunch    "SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch"
  SMmath       "SMmath: School Mean Math Outcome"
  SchoolN      "SchoolN: # Students Contributing Data".

* Selecting schools with data from at least 30 students.
SELECT IF (SchoolN GT 30).

* Descriptive statistics.
DESCRIPTIVES VARIABLES=math frlunch SMmath SMfrlunch SchoolN
  /STATISTICS=MEAN STDDEV MIN MAX.

* Centering school mean predictor.
COMPUTE SMfrlunch30 = SMfrlunch - .30.
VARIABLE LABELS SMfrlunch30 "SMfrlunch30: 0=.30".
EXECUTE.

```

STATA Code for Data Manipulation:

```

* label existing variables
label variable studentID "studentID: Student ID number"
label variable schoolID  "schoolID: School ID number"
label variable frlunch   "frlunch: Student Free/Reduced Lunch 0=No 1=Yes"
label variable math      "math: Student Free/Reduced Lunch 0=No 1=Yes"

* get school means of variables and label them
egen SMfrlunch = mean(frlunch), by (schoolID)
egen SMmath    = mean(math), by (schoolID)
label variable SMfrlunch "SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch"
label variable SMmath    "SMmath: School Mean Math Outcome"

* get number of students per school
egen SchoolN = count(studentID), by (schoolID)
label variable SchoolN "SchoolN: # Students Contributing Data"

* then drop schools with <= 30 students
drop if SchoolN < 31

* get means to center with
summarize math frlunch SMmath SMfrlunch SchoolN

* centering school mean predictor
gen SMfrlunch30 = SMfrlunch - .30
label variable SMfrlunch30 "SMfrlunch30: Percentage Students with Free Lunch (0=30%)"

```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|-----------|-------|----------|-----------|----------|----------|
| math | 13082 | 48.11856 | 17.25905 | 0 | 83 |
| frlunch | 13082 | .3075218 | .461485 | 0 | 1 |
| SMmath | 13082 | 48.11856 | 6.81813 | 29.45098 | 61.61364 |
| SMfrlunch | 13082 | .3075218 | .2220852 | 0 | .8032787 |
| SchoolN | 13082 | 274.9502 | 155.3319 | 31 | 515 |

Model 1: Two-Level Empty Means, Random Intercept for Math Test Outcome

$$\begin{aligned} \text{Level 1: } \text{Math}_{ij} &= \beta_{0j} + e_{ij} \\ \text{Level 2: } \beta_{0j} &= \gamma_{00} + U_{0j} \end{aligned}$$

```
TITLE1 "SAS Model 1: 2-Level Empty Means, Random Intercept for Math Outcome";
PROC MIXED DATA=work.grade10 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS schoolID studentID;
  MODEL math = / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID; RUN;
```

```
TITLE "SPSS Model 1: 2-Level Empty Means, Random Intercept for Math Outcome".
MIXED math BY schoolID studentID
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED =
  /RANDOM = INTERCEPT | SUBJECT(schoolID) COVTYPE(UN).
```

```
* STATA Model 1a: 2-Level Empty Means, Random Intercept for Math Outcome
xtmixed math , || schoolID: , ///
  variance ml covariance(un) residuals(independent),
  estat ic, n(94)
```

STATA output:

```
Mixed-effects ML regression          Number of obs      =      13082
Group variable: schoolID             Number of groups   =         94
                                     Obs per group: min =         31
                                     avg      =      139.2
                                     max      =         515
                                     Wald chi2(0)      =         .
Log likelihood = -54895.45            Prob > chi2         =         .
```

| math | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-------|----------|-----------|-------|-------|----------------------|
| _cons | 47.75613 | .7191927 | 66.40 | 0.000 | 46.34654 49.16572 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|----------|-----------|----------------------|
| schoolID: Identity | | | |
| var(_cons) | 44.93635 | 7.039956 | 33.05554 61.08735 |
| var(Residual) | 253.1756 | 3.141541 | 247.0926 259.4084 |

```
LR test vs. linear regression: chibar2(01) = 1857.08 Prob >= chibar2 = 0.0000
.      estat ic, n(94)
```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 94 | . | -54895.45 | 3 | 109796.9 | 109804.5 |

Note: N=94 used in calculating BIC
Covariance Parameter Estimates

Calculate the ICC for the correlation of students in the same school for math:

$$ICC = \frac{44.94}{44.94 + 253.18} = .15$$

This LR test tells us that the random intercept variance is significantly greater than 0,

Design effect using mean #students per school: $= 1 + ((n - 1) * ICC) \rightarrow 1 + [(275-1)*.15] = 42.1$

Effective sample size: $N_{\text{effective}} = (\# \text{Total Obs}) / \text{Design Effect} \rightarrow 13,082 / 42.1 = 311!!!$

95% random effect confidence interval for the intercept across schools: Fixed effect $\pm 1.96 * \text{SQRT}(\text{variance})$

$48 \pm 1.96 * \text{SQRT}(45) = 35 \text{ to } 61 \rightarrow 95\% \text{ of schools are predicted to have school mean math from 35 to 61}$

Model 2: Adding a Fixed Effect of Student Free/Reduced Lunch (Level 1)

| |
|--|
| Level 1: $\text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{FRLunch}_{ij}) + e_{ij}$ Level 2: Intercept: $\beta_{0j} = \gamma_{00} + U_{0j}$ Free/Reduced Lunch: $\beta_{1j} = \gamma_{10}$ |
|--|

```

TITLE1 "SAS Model 2: Adding Fixed Effect of Student Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS schoolID studentID;
  MODEL math = frlunch / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID; RUN;

TITLE "SPSS 2: Adding Fixed Effect of Student Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = frlunch
  /RANDOM = INTERCEPT | SUBJECT(schoolID) COVTYPE(UN).

* STATA Model 2: Adding Fixed Effect of Student Free/Reduced Lunch
xtmixed math c.frlunch, || schoolID: , ///
  variance ml covariance(un) residuals(independent),
  estat ic, n(94)

```

STATA output:

```

Log likelihood = -54508.069      Wald chi2(1)      =      808.17
                                Prob > chi2        =      0.0000
-----
      math |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
    frlunch |   -9.43162   .3317684   -28.43   0.000   -10.08187   -8.781366
      _cons |   50.61611   .5766321    87.78   0.000    49.48594   51.74629
-----+-----

Random-effects Parameters      |   Estimate   Std. Err.      [95% Conf. Interval]
-----+-----
schoolID: Identity              |
      var(_cons) |    26.89008   4.439001     19.45701    37.16277 → int var down by 40.16%
-----+-----
      var(Residual) |    239.3289   2.969964    233.5781   245.2213 → res var down by 5.47%
-----+-----

LR test vs. linear regression: chibar2(01) =   891.06 Prob >= chibar2 = 0.0000

.      estat ic, n(94)
-----+-----
      Model |   Obs    ll(null)   ll(model)    df         AIC         BIC
-----+-----
. |       94         .   -54508.07     4    109024.1    109034.3
-----+-----

Note: N=94 used in calculating BIC

```

What does the effect of student free/reduced lunch represent in this model?

Children who get free/reduced lunch score 9.43 points lower than children who don't.

What are we assuming about the effect of student free/reduced lunch?

We are assuming no contextual effect (that the between-school and within-school effects of FRLunch are equal).

Model 3: Adding a Fixed Effect of School Proportion Free/Reduced Lunch (Level 2)

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{FRLunch}_{ij}) + e_{ij}$$

$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolFRLunch}_j} - .30) + U_{0j}$$

$$\text{Free/Reduced Lunch: } \beta_{1j} = \gamma_{10}$$

```
TITLE1 "SAS Model 3: Adding Fixed Effect of School Proportion Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS schoolID studentID;
  MODEL math = frlunch SMfrlunch30 / SOLUTION DDFM=Satterthwaite OUTPM=work.LunchSave;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
  ESTIMATE "FR Lunch Between-School Effect" frlunch 1 SMfrlunch30 1;
RUN;
PROC CORR NOSIMPLE DATA=work.LunchSave; VAR math pred; RUN;
```

```
TITLE "SPSS Model 3: Adding Fixed Effect of School Proportion Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = frlunch SMfrlunch30
  /RANDOM = INTERCEPT | SUBJECT(schoolID) COVTYPE(UN)
  /SAVE = FIXPRED(lunchpred)
  /TEST = "FR Lunch Between-School Effect" frlunch 1 SMfrlunch30 1.
CORRELATIONS /VARIABLES = math lunchpred.
```

```
* STATA Model 3: Adding Fixed Effect of School Proportion Free/Reduced Lunch
xtmixed math c.frlunch c. SMfrlunch30, || schoolID: , ///
  variance ml covariance(un) residuals(independent),
  estat ic, n(94),
  predict lunchpred, // save fixed-effect predicted outcomes
  estimates store FixFRLunch, // save LL for LRT
  lincom 1*frlunch + 1*SMfrlunch30 // FR lunch between-school effect
corr math lunchpred // calculate total R2
```

STATA output:

```
Log likelihood = -54482.416      Wald chi2(2)      =      926.41
                                Prob > chi2        =      0.0000
```

| | math | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-------------|------|-----------|-----------|--------|-------|----------------------|
| frlunch | | -9.172883 | .3344153 | -27.43 | 0.000 | -9.828325 -8.517441 |
| SMfrlunch30 | | -16.85017 | 2.000813 | -8.42 | 0.000 | -20.77169 -12.92865 |
| _cons | | 50.60542 | .4341687 | 116.56 | 0.000 | 49.75447 51.45638 |

```
-----+-----
```

| Random-effects Parameters | | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|---------------|----------|-----------|--|
| schoolID: Identity | | | | |
| | var(_cons) | 13.48454 | 2.542895 | 9.317898 19.51437 → int var down by 49.85% |
| | var(Residual) | 239.3978 | 2.971595 | 233.6439 245.2935 → res var up by 0.03% |

```
-----+-----
```

LR test vs. linear regression: chibar2(01) = 354.12 Prob >= chibar2 = 0.0000

```
. estat ic, n(94),
```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 94 | . | -54482.42 | 5 | 108974.8 | 108987.5 |

```
-----+-----
```

Note: N=94 used in calculating BIC

```
. predict lunchpred, // save fixed-effect predicted outcomes
```

```
(option xb assumed)
.      estimates store FixFRLunch,          // save LL for LRT

.      lincom 1*c.frlunch + 1*c.SMfrlunch30    // FR lunch between-school effect
-----+-----
      math |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      (1) |    -26.02305   1.972668   -13.19   0.000   -29.88941    -22.1567

. corr math lunchpred          // calculate total R2
(obs=13082)
-----+-----
      math |      math lunchpred
-----+-----
      math |      1.0000
lunchpred |      0.4038      1.0000
```

R = .4038, so total $R^2 \sim .163$

Total reduction from **both** lunch effects:
Intercept variance → 69.99% (of 15%)
Residual variance → 5.44% (of 85%)

What does the effect of school proportion free/reduced lunch represent in this model?

This is the contextual effect for FRLunch: holding child lunch status constant, for every 10% more children in your school who get free/reduced lunch, school mean math is lower by 1.69 points. Before controlling for individual lunch status, the reduction is 2.60 points per 10% (between-school effect, given in separate estimate).

What does the effect of student free/reduced lunch NOW represent in this model?

This is the pure within-school effect: holding school lunch status constant, children who receive free/reduced lunch score 9.17 points lower than children who don't.

Model 4: Adding a Random Effect of Student Free/Reduced Lunch (over Schools)

Level 1: $\text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{FRLunch}_{ij}) + e_{ij}$
 Level 2: Intercept: $\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{SchoolFRLunch}_j - .30) + U_{0j}$
 Free/Reduced Lunch: $\beta_{1j} = \gamma_{10} + U_{1j}$

```
TITLE1 "SAS Model 4: Adding Random Effect of Student Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS schoolID studentID;
  MODEL math = frlunch SMfrlunch30 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT frlunch / G TYPE=UN SUBJECT=schoolID; RUN;
```

```
TITLE "SPSS Model 4: Adding Random Effect of Student Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV G
  /FIXED = frlunch SMfrlunch30
  /RANDOM = INTERCEPT frlunch | SUBJECT(schoolID) COVTYPE(UN).
```

```
* STATA Model 4: Adding Random Effect of Student Free/Reduced Lunch
xtmixed math c.frlunch c. SMfrlunch30, || schoolID: frlunch, ///
  variance ml covariance(un) residuals(independent),
  estat recovariance, level(schoolID),
  estat ic, n(94),
  estimates store RandFRLunch          // save LL for LRT
  lrtest RandFRLunch FixFRLunch       // LRT against fixed effect model
```

STATA output:

```
Log likelihood = -54438.694          Wald chi2(2)          =          400.83
                                     Prob > chi2           =           0.0000

-----+-----
      math |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
```


| | | | | | | |
|-------------|----------|----------|--------|-------|-----------|-----------|
| frlunch | -8.45 | .5611734 | -15.06 | 0.000 | -9.54988 | -7.350121 |
| SMfrlunch30 | -17.0879 | 1.917936 | -8.91 | 0.000 | -20.84698 | -13.32881 |
| _cons | 50.25931 | .5145964 | 97.67 | 0.000 | 49.25072 | 51.2679 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] | |
|---------------------------|-----------|-----------|----------------------|---------------------------------------|
| schoolID: Unstructured | | | | |
| var(frlunch) | 12.68699 | 3.311004 | 7.607035 | 21.15934 random slope var for frlunch |
| var(_cons) | 19.93184 | 3.745681 | 13.79068 | 28.80772 random intercept var |
| cov(frlunch,_cons) | -11.89358 | 3.164502 | -18.09589 | -5.691274 int-lunch covariance |
| var(Residual) | 236.8373 | 2.946808 | 231.1316 | 242.684 |

LR test vs. linear regression: $\chi^2(3) = 441.57$ Prob > $\chi^2 = 0.0000$

Note: LR test is conservative and provided only for reference.

. estat ic, n(94),

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------|-----|----------|-----------|----|----------|----------|
| . | 94 | . | -54438.69 | 7 | 108891.4 | 108909.2 |

Note: N=94 used in calculating BIC

. estimates store RandFRLunch // save LL for LRT

. lrtest RandFRLunch FixFRLunch // LRT against fixed effect model

Likelihood-ratio test LR $\chi^2(2) = 87.45$

(Assumption: FixFRLunch nested in RandFRLunch) Prob > $\chi^2 = 0.0000$

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

Is model 4 better than model 3? Yes
 $-2\Delta LL(2) = 88.2, p < .0001$

So what does this mean about the effect of student free/reduced lunch?

The difference in math between kids who get free/reduced lunch and kids who don't varies significantly over schools.

95% random effects CI for the random FRLunch slope: $\rightarrow -8.45 \pm 1.96 * \text{SQRT}(12.69) = -15.43 \text{ to } -1.47$

On average, the gap related to lunch status is 8.45 points, but across 95% of the schools, that gap is predicted to be anywhere from 1.47 to 15.43 points.

Model 5: Adding a Cross-Level Interaction of Student by School Free/Reduced Lunch

$$\begin{aligned} \text{Level 1: } \text{Math}_{ij} &= \beta_{0j} + \beta_{1j}(\text{FRLunch}_{ij}) + e_{ij} \\ \text{Level 2: } \text{Intercept: } \beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SchoolFRLunch}}_j - .30) + U_{0j} \\ \text{Free/Reduced Lunch: } \beta_{1j} &= \gamma_{10} + \gamma_{11}(\overline{\text{SchoolFRLunch}}_j - .30) + U_{1j} \end{aligned}$$

```
TITLE1 "SAS Model 5: Adding Cross-Level Interaction of Student by School Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS schoolID studentID;
  MODEL math = frlunch SMfrlunch30 frlunch*SMfrlunch30 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT frlunch / TYPE=UN SUBJECT=schoolID; RUN;
```

```
TITLE "SPSS Model 5: Adding Cross-Level Interaction of Student by School Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
  /METHOD = ML
  /PRINT = SOLUTION TESTCOV
  /FIXED = frlunch SMfrlunch30 frlunch*SMfrlunch30
  /RANDOM = INTERCEPT frlunch | SUBJECT(schoolID) COVTYPE(UN).
```

```
* STATA Model 5: Adding Cross-Level Interaction of Student by School Free/Reduced Lunch
xtmixed math c.frlunch c.smfrlunch30 c.frlunch#c.smfrlunch30, ///
  || schoolID: frlunch, variance ml covariance(un) residuals(independent),
  estat ic, n(94)
```

STATA output:

```

                                Wald chi2(3)      =    413.76
Log likelihood = -54437.502      Prob > chi2      =    0.0000
-----+-----
      math |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      frlunch | -8.688252   .5673922   -15.31   0.000   -9.80032   -7.576183
    SMfrlunch30 | -19.45931   2.473474    -7.87   0.000  -24.30723  -14.61139
c.frlunch#c.SMfrlunch30 | 4.140733   2.633721     1.57   0.116   -1.021265   9.302731
      _cons | 50.22283   .5140769    97.70   0.000   49.21526   51.2304
-----+-----

Random-effects Parameters      |   Estimate   Std. Err.      [95% Conf. Interval]
-----+-----
schoolID: Unstructured         |
      var(frlunch) |   11.79733   3.165294     6.972708   19.96026 → slope var down by 7.01%
      var(_cons) |   19.82708   3.701312    13.75171   28.58648 → int var down by 0.53%
    cov(frlunch,_cons) |  -11.34396   3.087016    -17.3944   -5.293523
-----+-----
      var(Residual) |   236.8234   2.946467    231.1183   242.6694 → res var down by .01%
-----+-----

LR test vs. linear regression:   chi2(3) =   442.87   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
.      estat ic, n(94)
-----+-----
      Model |   Obs     ll(null)   ll(model)     df         AIC         BIC
-----+-----
      . |    94           .   -54437.5         8     108891     108911.4
-----+-----

```

What does the effect of student free/reduced lunch NOW represent in this model?

This is the difference between kids who get free/reduced lunch and those who don't in schools where 30% of the kids get free/reduced lunch: those kids who get free/reduced lunch are lower by 8.69.

What does the effect of school proportion free/reduced lunch NOW represent in this model?

This is the contextual (incremental between-school) effect for a kid who does not receive free/reduced lunch: for those kids, for every 10% more kids in their school that receive free/reduced lunch, their school mean math is lower by 1.94.

What does the cross-level interaction of student by school free/reduced lunch represent?

The effect of being a kid who receives free/reduced lunch is reduced nonsignificantly by 0.4 for every 10% more children in their school who get free/reduced lunch. But this effect is currently smushed—it assumes without testing that school FRLunch moderates the within-school and between-school effects of FRLunch to the same extent.

Model 6: Adding a Level-2 Interaction of Quadratic School Free/Reduced Lunch

$$\begin{aligned}
 \text{Level 1: } \text{Math}_{ij} &= \beta_{0j} + \beta_{1j}(\text{FRLunch}_{ij}) + e_{ij} \\
 \text{Level 2: } \text{Intercept: } \beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SchoolFRLunch}}_j - .30) + \gamma_{02}(\overline{\text{SchoolFRLunch}}_j - .30)^2 + U_{0j} \\
 \text{Free/Reduced Lunch: } \beta_{1j} &= \gamma_{10} + \gamma_{11}(\overline{\text{SchoolFRLunch}}_j - .30) + U_{1j}
 \end{aligned}$$

```

TITLE1 "SAS Model 6: Adding Level-2 Interaction of Quadratic School Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 IC METHOD=ML;
  CLASS schoolID studentID;
  MODEL math = frlunch SMfrlunch30 frlunch*SMfrlunch30 SMfrlunch30*SMfrlunch30
    / SOLUTION DDFM=Satterthwaite OUTPM=work.TotalSave;
  RANDOM INTERCEPT frlunch / TYPE=UN SUBJECT=schoolID;
  ESTIMATE "FR Lunch Between-School Main Effect" frlunch 1 SMfrlunch30 1;
  ESTIMATE "FR Lunch Between-School Interaction" frlunch*SMfrlunch30 1 SMfrlunch30*SMfrlunch30 1;
RUN; PROC CORR NOSIMPLE DATA=work.TotalSave; VAR math pred; RUN;

```

```

TITLE "SPSS Model 6: Adding Level-2 Interaction of Quadratic School Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
    /METHOD = ML
    /PRINT = SOLUTION TESTCOV
    /FIXED = frlunch SMfrlunch30 frlunch*SMfrlunch30 SMfrlunch30*SMfrlunch30
    /RANDOM = INTERCEPT frlunch | SUBJECT(schoolID) COVTYPE(UN)
    /SAVE = FIXPRED(totalpred)
    /TEST = "FR Lunch Between-School Main Effect" frlunch 1 SMfrlunch30 1
    /TEST = "FR Lunch Between-School Interaction" frlunch*SMfrlunch30 1 SMfrlunch30*SMfrlunch30 1.
CORRELATIONS /VARIABLES = math totalpred.

* STATA Model 6: Adding Level-2 Interaction of Quadratic School Free/Reduced Lunch
xtmixed math c.frlunch c.SMfrlunch30 c.frlunch#c.SMfrlunch30 c.SMfrlunch30#c.SMfrlunch30, ///
    || schoolID: frlunch, variance ml covariance(un) residuals(independent),
    estat ic, n(94),
    predict totalpred, // save fixed-effect predicted outcomes
lincom 1*c.frlunch + 1*c.SMfrlunch30 // FR lunch between-school main effect
lincom 1*c.frlunch#c.SMfrlunch30 + 1*c.SMfrlunch30#c.SMfrlunch30 // FR lunch BS interaction
margins, at(c.frlunch=(0 1) c.SMfrlunch30=(-.2 0 .2 .4)) vsquish // create predicted values
marginsplot, noci name(predicted_lunch, replace) xdimension(frlunch) // plot predicted, no CI
corr math totalpred // calculate total R2

```

STATA output:

```

Log likelihood = -54436.242      Wald chi2(4)      =      418.05
                                Prob > chi2        =      0.0000

```

| math | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|-----------------------------|-----------|-----------|--------|-------|----------------------|-----------|
| frlunch | -8.835737 | .5769075 | -15.32 | 0.000 | -9.966455 | -7.705019 |
| SMfrlunch30 | -17.98486 | 2.595472 | -6.93 | 0.000 | -23.07189 | -12.89783 |
| c.frlunch#c.SMfrlunch30 | 5.428146 | 2.764887 | 1.96 | 0.050 | .0090667 | 10.84723 |
| c.SMfrlunch30#c.SMfrlunch30 | -14.2013 | 8.815645 | -1.61 | 0.107 | -31.47965 | 3.077044 |
| _cons | 50.85941 | .6398308 | 79.49 | 0.000 | 49.60537 | 52.11346 |

```

-----
Random-effects Parameters      |      Estimate      Std. Err.      [95% Conf. Interval]
-----+-----
schoolID: Unstructured         |
    var(frlunch)               |      11.81828      3.178501      6.976308      20.02088 → slope var up by -.18%
    var(_cons)                 |      18.95016      3.572456      13.09621      27.4208 → int var down by 4.42%
    cov(frlunch,_cons)         |     -10.92287      3.032665     -16.86678     -4.978956
-----+-----
    var(Residual)              |      236.8186      2.946416      231.1136      242.6645 → res var same
-----
LR test vs. linear regression:      chi2(3) =      426.87      Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
.      estat ic, n(94),
-----
Model |      Obs      ll(null)      ll(model)      df      AIC      BIC
-----+-----
.      |      94          .      -54436.24      9      108890.5      108913.4
-----

```

Note: N=94 used in calculating BIC

What does the cross-level interaction of student by school free/reduced lunch NOW represent?

The effect of being a kid who receives free/reduced lunch (now after allowing for differential moderation across levels of the effects of free/reduced lunch at both levels by school mean free/reduced lunch) is reduced significantly by 0.54 for every 10% more children in their school who get free/reduced lunch.

What does the level-2 interaction of quadratic school free/reduced lunch represent?

After controlling for kid free/reduced lunch status, the contextual (incremental between-school) effect of school mean free/reduced lunch as evaluated at 30% FRLunch becomes nonsignificantly more negative by 2*1.13 for every 10% more kids in their school with free/reduced lunch.

```
. lincom 1*c.frlunch + 1*c.SMfrlunch30 // FR lunch between-school main effect
```

| | math | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-----|------|----------|-----------|--------|-------|----------------------|
| (1) | | -26.8206 | 2.603258 | -10.30 | 0.000 | -31.92289 -21.71831 |

```
. lincom 1*c.frlunch#c.SMfrlunch30 + 1*c.SMfrlunch30#c.SMfrlunch30 // FR lunch between-school interaction
```

| | math | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-----|------|-----------|-----------|-------|-------|----------------------|
| (1) | | -8.773157 | 8.41717 | -1.04 | 0.297 | -25.27051 7.724192 |

If we don't control for kid free/reduced lunch, the between-school effect of -2.68 per 10% of school mean free/reduced lunch as evaluated at 30% Frlunch becomes nonsignificantly more negative by 2×0.88 for every 10% more kids in their school with free/reduced lunch.

So school mean free/reduced lunch moderates the within-school Frlunch effect, but not the contextual (incremental between-school) or between-school effects.

```
. margins, at(c.frlunch=(0 1) c.SMfrlunch30=(-.2 0 .2 .4)) vsquish // create predicted values
Adjusted predictions Number of obs = 13082
```

Expression : Linear prediction, fixed portion, predict()

| | | | Delta-method | | | | |
|-------|-------------|-------|--------------|-----------|---|------|----------------------|
| | | | Margin | Std. Err. | z | P> z | [95% Conf. Interval] |
| 1._at | : frlunch | = 0 | | | | | |
| | SMfrlunch30 | = -.2 | | | | | |
| 2._at | : frlunch | = 0 | | | | | |
| | SMfrlunch30 | = 0 | | | | | |
| 3._at | : frlunch | = 0 | | | | | |
| | SMfrlunch30 | = .2 | | | | | |
| 4._at | : frlunch | = 0 | | | | | |
| | SMfrlunch30 | = .4 | | | | | |
| 5._at | : frlunch | = 1 | | | | | |
| | SMfrlunch30 | = -.2 | | | | | |
| 6._at | : frlunch | = 1 | | | | | |
| | SMfrlunch30 | = 0 | | | | | |
| 7._at | : frlunch | = 1 | | | | | |
| | SMfrlunch30 | = .2 | | | | | |
| 8._at | : frlunch | = 1 | | | | | |
| | SMfrlunch30 | = .4 | | | | | |

```
. marginsplot, noci name(predicted_lunch, replace) xdimension(frlunch) //
plot predicted, no CI
Variables that uniquely identify margins: frlunch SMfrlunch30
```

```
. corr math totalpred // calculate total R2
(obs=13082)
```

| | math | totalpred |
|-----------|--------|-----------|
| math | 1.0000 | |
| totalpred | 0.4051 | 1.0000 |

$R = .4051$, so total $R^2 = .164$

Additional reduction from **both** interactions:
 Intercept variance → 4.93%
 Lunch slope variance → 6.85%
 Residual variance → 0.01%

Sample Results Section (note that “smushed” models are not reported)...

The extent to which student free/reduced lunch status could predict student math outcomes was examined in a series of multilevel models in which the 13,802 students were modeled as nested within their 94 schools. Maximum likelihood (ML) was used in estimating and reporting all model parameters. The significance of fixed effects was evaluated with individual Wald tests (i.e., of estimate / SE), whereas random effects were evaluated via likelihood ratio tests (i.e., $-2\Delta LL$ with degrees of freedom equal to the number of new random effects variances and covariances). Effect size was evaluated via pseudo- R^2 values for the proportion reduction in each variance component, as well as with total R^2 , the squared correlation between the actual math outcomes and the math outcomes predicted by the fixed effects.

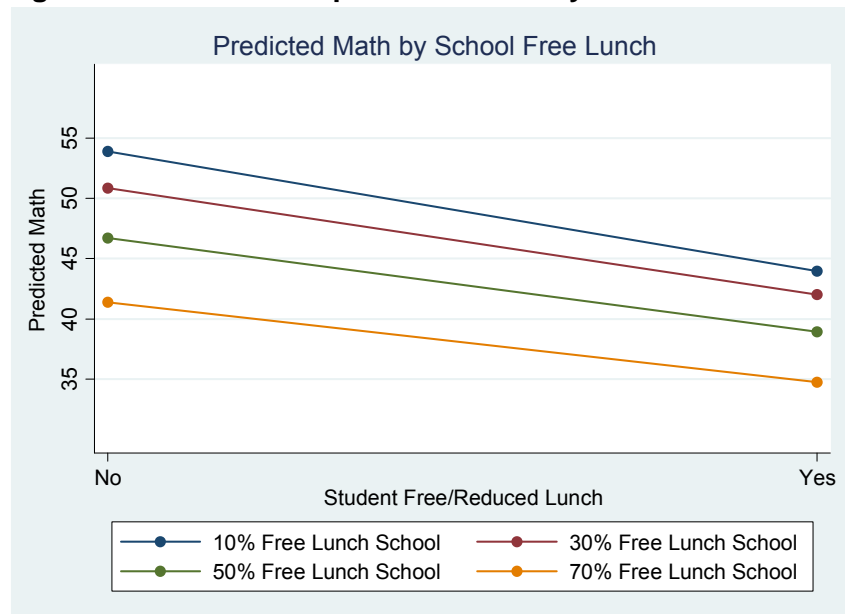
As derived from an empty means, random intercept model, student math scores had an intraclass correlation of .15, indicating that 15% of the variance in math scores was between schools. A 95% random effects confidence interval, calculated as fixed intercept $\pm 1.96 \times \text{SQRT}(\text{random intercept variance})$, revealed that 95% of the sample schools were predicted to have intercepts for school mean math scores between 35 to 61. Children who did not receive free/reduced lunch were treated as the reference group. Given the large variability across schools in the proportion of students who received free/reduced lunch (0–80% of students), a contextual effect at level 2 was represented by the school proportion of students who receive free/reduced lunch centered near the sample mean of 30%.

The effects of free/reduced lunch status at each level were then added to the model. The within-school effect was significant and accounted for 5.44% of the residual variance, and indicated that students who receive free/reduced lunch are expected to have lower math scores than other students in their school by 9.18. The between-school effect was also significant and accounted for 70% of the remaining random intercept variance, and indicated that for every additional 10% of students who receive free/reduced lunch, that school's mean math score is expected to be lower by 2.60. After controlling for student free/reduced lunch, the contextual free/reduced lunch effect of -1.69 per additional 10% of students was still significant. A random slope for the effect of free/reduced lunch also resulted in a significant improvement in model fit, $-2\Delta\text{LL}(2) = 88.2$, $p < .001$, indicating that the size of the disadvantage related to free/reduced lunch differed significantly across schools. A 95% random effects confidence interval for the student free/reduced lunch effect, calculated as fixed slope $\pm 1.96 \times \text{SQRT}(\text{random slope variance})$, revealed that 95% of the schools were predicted to have lunch-related gaps between students ranging from -15.45 to -1.46 .

The extent to which school differences in the lunch-related disadvantage in math could be predicted from school lunch composition was then examined by adding a cross-level intra-variable interaction between the student and school lunch predictors, as well as the quadratic effect of school lunch composition to control for a contextual interaction effect. The within-school lunch effect was significantly moderated by school lunch composition (which reduced its random slope variance by 6.85%), although the moderation of the between-school and contextual effects was not significant, reducing the random intercept variance by another 4.93%, for a total $R^2 = .164$.

The significant intra-variable cross-level interaction, as shown by the nonparallel slopes of the lines in Figure 1, indicated that the lunch-related disadvantage in math scores of 8.84, as found for students receiving free/reduced lunch in schools in which 30% of students received free/reduced lunch, became significantly less negative by 0.54 for every additional 10% of students who received free/reduced lunch. Alternatively, the contextual school effect of -1.80 per 10% free/reduced lunch students (in baseline students in schools with 30% free/reduced lunch students) was reduced by 0.54 in free/reduced lunch students. The level-2 quadratic effect, seen by the widening distance between the lines in Figure 1, indicated that the same contextual school effect became nonsignificantly more negative by 1.42 for every additional 10% free/reduced lunch students (i.e., controlling for student lunch status), or that the between-school effect of -2.68 per 10% students became nonsignificantly more negative by 0.88 per 10% students (i.e., not controlling for student lunch status).

Figure 1: Plot of model-predicted math by free/reduced lunch status



Example 6: Clustered Longitudinal Data for Time within Person within Twin Pair

The data for this example come from the Octogenarian Twin Study of Aging, a longitudinal study (with 5 occasions spanning 8 years) of same-sex twin pairs initially age 79-100. We will be examining change over time in a measure of crystallized intelligence (information test), as well as prediction of that change from a measure of physical functioning (grip strength measured in pounds). These data are already stacked such that one row contains the data for one occasion for one person. The ID variables PairID and TwinID index which twin pair and which person, respectively, and Case is a unique identifier for each person. Time is unbalanced across persons, so the REPEATED statement will not be used (because we have to assume a VC R matrix anyway).

Model 1a: Empty Means, 2-Level Model for Information Test Outcome

$$\begin{aligned}\text{Level 1: } \text{Info}_{ti} &= \beta_{0i} + e_{ti} \\ \text{Level 2: } \beta_{0i} &= \gamma_{00} + U_{0i}\end{aligned}$$

This model has 2 variance components: residual at level 1 and random intercept at level 2. It assumes that all people are independent (does not account for twin pair membership).

```
TITLE "SAS Model 1a: Empty Means, 2-Level Model for Information Test Outcome";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID*TwinID; RUN;
```

```
TITLE "SPSS Model 1a: Empty Means, 2-Level Model for Information Test Outcome".
MIXED info BY PairID TwinID
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED =
  /RANDOM = INTERCEPT | SUBJECT(PairID*TwinID) COVTYPE(UN).
```

```
* STATA Model 1a: Empty Means, 2-Level Model for Information Test Outcome
xtmixed info , || Case: , variance reml covariance(unstructured)
  estat ic, n(594)
  estimates store TwoLevel
```

STATA output:

```
Mixed-effects REML regression      Number of obs      =      1734
Group variable: Case              Number of groups    =       594
                                   Obs per group: min =         1
                                   avg      =        2.9
                                   max      =         5
                                   Wald chi2(0)      =         .
                                   Prob > chi2       =         .

Log restricted-likelihood = -6073.7202
```

| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
|-------------|----------|-----------|-------|-------|----------------------|
| -----+----- | | | | | |
| _cons | 25.46294 | .4909978 | 51.86 | 0.000 | 24.5006 26.42527 |

| Random-effects Parameters | Estimate | Std. Err. | [95% Conf. Interval] |
|---------------------------|----------|-----------|----------------------|
| -----+----- | | | |
| Case: Identity | | | |
| var(_cons) | 130.5222 | 8.38369 | 115.0827 148.0331 |
| -----+----- | | | |
| var(Residual) | 26.66816 | 1.120232 | 24.5605 28.95669 |

```
LR test vs. linear regression: chibar2(01) = 1411.30 Prob >= chibar2 = 0.0000
. estat ic, n(594)
```

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|-------------|-----|----------|-----------|----|----------|---------|
| -----+----- | | | | | | |
| . | 594 | . | -6073.72 | 3 | 12153.44 | 12166.6 |

Note: N=594 used in calculating BIC

Calculate the ICC for the proportion of between-person variation in Info:

$$ICC = \frac{130.52}{130.52 + 26.67} = .83$$

This LR test tells us that the random intercept variance is significantly greater than 0,

This model now has 3 variance components: residual at level-1, random intercept at level 2, and random intercept at level 3. It now allows a correlation between people from the same twin pair.


```
. estimates store ThreeLevel
. lrtest ThreeLevel TwoLevel
```

Is the 3-level model a better fit than the 2-level model?
 Yes, $-2\Delta LL(1) = 101.5$, $p < .001$

```
Likelihood-ratio test                LR chi2(1)  =    101.50
(Assumption: TwoLevel nested in ThreeLevel)  Prob > chi2 =    0.0000
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative. Note: LR tests based on REML are valid only when the fixed-effects specification is identical for both models.

Proportion variance at each level:

```
Level 1 (time) = 26.75 / 157.83 = .17
Level 2 (person) = 47.34 / 157.83 = .30
Level 3 (pair) = 83.73 / 157.83 = .53
```

ICC for time within person & pair =

$83.73 + 47.34 / (83.73 + 47.34 + 26.75) = .83$

ICC for person within pair = $83.72 / (83.72 + 47.33) = .64$

This ICC = .64 is significantly greater than 0 via $-2\Delta LL$ for 3- vs. 2-level.

Now let's do the same thing for our two time-varying predictors: age and grip strength.

Age Model: Empty Means, 3-Level Model for Age Predictor

```
TITLE "SAS Age Model: Empty Means, 3-Level Model for Age Predictor";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID;           * Level 3;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID*TwinID;    * Level 2; RUN;
```

```
TITLE "SPSS Age Model: Empty Means, 3-Level Model for Age Predictor".
```

```
MIXED info BY PairID TwinID
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED =
  /RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)
  /RANDOM = INTERCEPT | SUBJECT(PairID*TwinID) COVTYPE(UN).
```

```
* STATA Age Model: Empty Means, 3-Level Model for Age Predictor
```

```
xtmixed info , || PairID: , covariance(unstructured) ///
  || Case: , variance reml covariance(unstructured)
```

STATA output:

```
-----+-----
      age |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _cons |   85.96476   .1585134   542.32   0.000    85.65408    86.27544
-----+-----

Random-effects Parameters |   Estimate    Std. Err.     [95% Conf. Interval]
-----+-----
PairID: Identity  var(_cons) |   6.553374    .6752503     5.354986    8.019948  level-3 between-pair = 47%
-----+-----
Case: Identity   var(_cons) |   2.84e-23    5.82e-23     5.09e-25    1.58e-21  level-2 within-pair = 0%
-----+-----
                  var(Residual) |   7.466046    .2842018     6.929293    8.044377  level-1 within-person = 53%
-----+-----

LR test vs. linear regression:      chi2(2) =    459.38    Prob > chi2 = 0.0000
```

Because there is no age variance at level 2, age will be a predictor at levels 1 and 3 only.

Grip Strength Model: Empty Means, 3-Level Model for Grip Strength Predictor

```
TITLE "SAS Grip Model: Empty Means, 3-Level Model for Grip Strength Predictor";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL gripp = / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID;          * Level 3;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID*TwinID;    * Level 2; RUN;
```

```
TITLE "SPSS Grip Model: Empty Means, 3-Level Model for Grip Strength Predictor".
```

```
MIXED gripp BY PairID TwinID
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED =
  /RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)
  /RANDOM = INTERCEPT | SUBJECT(PairID*TwinID) COVTYPE(UN).
```

```
* STATA Grip Model: Empty Means, 3-Level Model for Grip Strength Predictor
xtmixed gripp , || PairID: , covariance(unstructured) ///
  || Case: , variance reml covariance(unstructured)
```

STATA output:

```
-----
      gripp |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _cons |   8.06599   .1268694    63.58   0.000     7.817331     8.31465
-----+-----

Random-effects Parameters   |   Estimate   Std. Err.     [95% Conf. Interval]
-----+-----
PairID: Identity  var(_cons) |   3.085847   .4673646     2.293276     4.152336 level-3 between-pair = 36%
-----+-----
Case: Identity    var(_cons) |   2.552534   .3436612     1.960513     3.323329 level-2 within-pair = 29%
-----+-----
                        var(Residual) |   3.049563   .1271551     2.810255     3.309249 level-1 within-person = 35%
-----+-----

LR test vs. linear regression:      chi2(2) =    795.50   Prob > chi2 = 0.0000
```

Because there is grip strength variance at all levels, grip strength will be a predictor at all levels.

We now need to create our predictor variables, including a mean of grip strength at the pair and person levels. We then code time as “time-in-study” and use baseline age as between-pair age. This gives us a convenient demarcation of age at baseline as the cross-sectional effect of age, and time-in-study as the longitudinal effect of age.

SAS Data Manipulation:

```
* Importing data into work library and creating person mean gripp for level-2;
DATA work.octodata; SET octo.octodata;
  PMgripp = MEAN(OF gripp1-gripp5);
  LABEL PMgripp= "PMgripp: Person Mean Gripp"; RUN;

* Getting twin pair means for grip strength to use at level-3;
PROC SORT DATA=work.octodata; BY PairID TwinID Wave; run;
PROC MEANS NOPRINT DATA=work.octodata; BY PairID; VAR PMgripp;
  OUTPUT OUT=PairMeans MEAN(PMgripp) = FMgripp; RUN;

* Merging PairMeans with datafile and centering predictors;
DATA work.octodata; MERGE work.octodata work.PairMeans; BY PairID;
  LABEL FMgripp= "FMgripp: Family Mean Gripp";

*** Age Variables ***;
  * Centering age at time 1 at 85 to use at level-3;
    BFage85 = agew1 - 85; LABEL BFage85= "BFage85: Age at Time1 (0=85)";
  * Within-person centering age at level-1 (like PERSON MEAN CENTERING);
    time = age - agew1; LABEL time= "time: Time Since Entry (0= Age Wave 1)";
```

```

*** Grip Strength Variables ***;
* Centering family mean gripp at 9 to use at level-3;
  BFgripp9 = FMgripp - 9;
* Centering person mean gripp at 9 to use at level-2;
  BPgripp9 = PMgripp - 9;          * GRAND MEAN CENTERING;
  WFgripp  = PMgripp - FMgripp;    * PERSON MEAN CENTERING;
* Centering time-varying gripp to use at level-1;
  TVgripp9 = gripp - 9;          * GRAND MEAN CENTERING;
  WPgripp  = gripp - PMgripp;    * PERSON MEAN CENTERING;
LABEL BFgripp9= "BFgripp9: Between-Family Mean Grip Strength in Pounds (0=9)"
      BPgripp9= "BPgripp9: Between-Person Mean Grip Strength in Pounds (0=9)"
      WFgripp=  "WFgripp: Within-Family Deviation from Mean Grip Strength in Pounds"
      TVgripp9= "TVgripp9: Time-Varying Grip Strength in Pounds (0=9)"
      WPgripp=  "WPgripp: Within-Person Deviation from Mean Grip Strength in Pounds";

* Selecting only cases with complete data;
  IF NMISS(agem1, age, FMgripp, PMgripp, gripp, info)>0 THEN DELETE; RUN;

```

SPSS Data Manipulation:

```

SORT CASES BY PairID TwinID Wave.
* Getting person gripp means to use as level-2 predictor.
COMPUTE PMgripp = MEAN(gripp1 TO gripp5).
EXECUTE.
* Getting pair gripp means to use as level-3 predictor.
AGGREGATE /OUTFILE=* MODE=ADDVARIABLES /PRESORTED /BREAK = PairID /FMgripp = MEAN(PMgripp).
VARIABLE LABELS FMgripp "FMgripp: Family Mean Gripp" PMgripp "PMgripp: Person Mean Gripp".

*** Age Variables ***.
* Centering age at time 1 at 85 to use at level-3.
  COMPUTE BFace85 = agew1 - 85.
* Within-person centering age at level-1 (like PERSON MEAN CENTERING).
  COMPUTE time = age - agew1.
  VARIABLE LABELS BFace85 "BFace85: Age at Time1 (0=85)"
                 time    "time: Time Since Entry (0= Age Wave 1)".

*** Grip Strength Variables ***.
* Centering family mean gripp at 9 to use at level-3.
  COMPUTE BFgripp9 = FMgripp - 9.
* Centering person mean gripp at 9 to use at level-2.
  COMPUTE BPgripp9 = PMgripp - 9.
  COMPUTE WFgripp  = PMgripp - FMgripp.
* Centering time-varying gripp to use at level-1.
  COMPUTE TVgripp9 = gripp - 9.
  COMPUTE WPgripp  = gripp - PMgripp.
VARIABLE LABELS
  BFgripp9 "BFgripp9: Between-Family Mean Grip Strength in Pounds (0=9)"
  BPgripp9 "BPgripp9: Between-Person Mean Grip Strength in Pounds (0=9)"
  WFgripp  "WFgripp: Within-Family Deviation from Mean Grip Strength in Pounds"
  TVgripp9 "TVgripp9: Time-Varying Grip Strength in Pounds (0=9)"
  WPgripp  "WPgripp: Within-Person Deviation from Mean Grip Strength in Pounds".

* Selecting only complete cases.
  SELECT IF (NMISS(agem1, age, FMgripp, PMgripp, gripp, info)=0).
  EXECUTE.

```

STATA Data Manipulation:

```

* Creating person mean gripp for level-2
egen PMgripp = rmean(GRIPP1-GRIPP5)
label variable PMgripp "PMgripp: Person Mean Gripp"
* Creating family mean gripp for level-3
egen FMgripp = mean(PMgripp), by(PairID)
label variable FMgripp "FMgripp: Family Mean Gripp"

```

```

* Age variables
* centering age at time 1 at 85 to use at level-3
gen BFage85 = agew1 - 85
label variable BFage85 "BFage85: Age at Time1 (0=85)"
* within person centering age at level-1 (like PERSON MEAN CENTERING)
gen time = age - agew1
label variable time "time: Time since entry (0= Age Wave 1)"

* Grip Strength Variables
* centering family mean gripp at 9 use at level-3
gen BFgripp9 = FMgripp - 9
* centering person mean gripp at 9 to use at level-2
gen BPgripp9 = PMgripp - 9 // GRAND MEAN CENTERING
gen WFgripp = PMgripp - FMgripp // PERSON MEAN CENTERING
* centering time-varying gripp to use at level-1
gen TVgripp9 = gripp - 9 // GRAND MEAN CENTERING
gen WPgripp = gripp - PMgripp // PERSON MEAN CENTERING

label variable BFgripp9 "BFgripp9: Between-Family Mean Grip Strength in Pounds (0=9)"
label variable BPgripp9 "BPgripp9: Between-Person mean gripp strength in pounds (0=9)"
label variable WFgripp "WFgripp: Within-Family deviation from mean grip strength in Pounds"
label variable TVgripp9 "TVgripp9: Time-Varying Grip Strength in Pounds (0=9)"
label variable WPgripp "WPgripp: Within-Person Deviation from Mean Grip Strength in Pounds"

* Selecting only cases with complete data
egen nummiss = rowmiss(agew1 age FMgripp PMgripp gripp info)
drop if nummiss>0

```

Model 2a: Fixed Quadratic, Random Intercepts at Levels 2 and 3

Level 1: $\text{Info}_{ij} = \beta_{0ij} + \beta_{1ij}(\text{Age}_{ij} - \text{PairAge1}_j) + \beta_{2ij}(\text{Age}_{ij} - \text{PairAge1}_j)^2 + e_{ij}$

Level 2:

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

Linear Time: $\beta_{1ij} = \delta_{10j}$

Quadratic Time: $\beta_{2ij} = \delta_{20j}$

Level 3:

Intercept: $\delta_{00j} = \gamma_{000} + \gamma_{001}(\text{PairAge1}_j - 85) + V_{00j}$

Linear Time: $\delta_{10j} = \gamma_{100}$

Quadratic Time: $\delta_{20j} = \gamma_{200}$

```

TITLE "SAS Model 2a: Fixed Quadratic, Random Intercept for Pair and Twin";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = BFage85 time time*time / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID; * Level 3;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID*TwinID; * Level 2; RUN;

```

```

TITLE "SPSS Model 2a: Fixed Quadratic, Random Intercept for Pair and Twin".
MIXED info BY PairID TwinID WITH BFage85 time

```

```

/METHOD = REML
/PRINT = SOLUTION TESTCOV
/FIXED = BFage85 time time*time
/RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)
/RANDOM = INTERCEPT | SUBJECT(PairID*TwinID) COVTYPE(UN).

```

```

* STATA Model 2a: Fixed Quadratic, Random Intercepts at Levels 2 and 3
xtmixed info c.BFage85 c.time c.time#c.time , || PairID: , covariance(unstructured) ///
  || Case: , variance reml covariance(unstructured)
estat ic, n(337)
estimates store FixQuad

```

STATA output:

```

Wald chi2(3)      = 195.45
Log restricted-likelihood = -5939.0225      Prob > chi2      = 0.0000
-----+-----
      info |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      BFace85 | -.8073689   .1942406    -4.16   0.000   -1.188074   -.4266643
        time | -.2350914   .1456677    -1.61   0.107   -.5205948    .050412
c.time#c.time | -.0555854   .0187153    -2.97   0.003   -.0922667   -.018904
        _cons | 25.10103   .6834791    36.73   0.000    23.76144   26.44062
-----+-----

Random-effects Parameters   |   Estimate   Std. Err.   [95% Conf. Interval]
-----+-----
PairID: Identity  var(_cons) |   79.53662   9.694711    62.6346   100.9997
-----+-----
Case: Identity   var(_cons) |   52.4136   5.67978    42.38419   64.81628
-----+-----
                var(Residual) |   22.77218   .9601037    20.96606   24.73389
-----+-----

LR test vs. linear regression:      chi2(2) = 1636.90   Prob > chi2 = 0.0000
-----+-----
      Model |   Obs    ll(null)   ll(model)    df         AIC         BIC
-----+-----
          . |   337         .   -5939.023     7    11892.05    11918.79

```

This model has 3 variance components: residual at level-1, random intercept at level-2, and random intercept at level-3. It now also has 3 new fixed effects: BFace85, time, and time².

We do not compare REML deviances because these models differ in fixed effects. Instead, we use their p-values. This is our new unconditional growth model baseline, as obtained from testing sequential models not shown here.

Model 2b: Fixed Quadratic, Random Linear Slope at Level 2

Level 1: $\text{Info}_{ij} = \beta_{0ij} + \beta_{1ij}(\text{Age}_{ij} - \text{PairAge1}_j) + \beta_{2ij}(\text{Age}_{ij} - \text{PairAge1}_j)^2 + e_{ij}$

Level 2:

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

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

Quadratic Time: $\beta_{2ij} = \delta_{20j}$

Level 3:

Intercept: $\delta_{00j} = \gamma_{000} + \gamma_{001}(\text{PairAge1}_j - 85) + V_{00j}$

Linear Time: $\delta_{10j} = \gamma_{100}$

Quadratic Time: $\delta_{20j} = \gamma_{200}$

```

TITLE "SAS Model 2b: Add Random Linear Slope for Twin";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = BFace85 time time*time / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID; * Level 3;
  RANDOM INTERCEPT time / TYPE=UN SUBJECT=PairID*TwinID; * Level 2; RUN;

```

```

TITLE "SPSS Model 2b: Add Random Linear Slope for Twin".
MIXED info BY PairID TwinID WITH BFace85 time
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED = BFace85 time time*time
  /RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)
  /RANDOM = INTERCEPT time | SUBJECT(PairID*TwinID) COVTYPE(UN).

```

```

* STATA Model 2b: Add Random Linear Slope for Twin
xtmixed info c.BFace85 c.time c.time#c.time , || PairID: , covariance(unstructured) ///
  || Case: time , variance reml covariance(unstructured)
estat ic, n(337)
estimates store RandLin2
lrtest RandLin2 FixQuad

```

STATA output:

```

Wald chi2(3)      =    188.20
Log restricted-likelihood = -5872.9993      Prob > chi2      =    0.0000
-----+-----
      info |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      BFage85 |   -.7307761   .1909202    -3.83   0.000   -1.104973   -.3565793
        time |   -.1454705   .132939    -1.09   0.274   -.4060262   .1150853
c.time#c.time |   -.1021417   .0165422   -6.17   0.000   -.1345639   -.0697195
        _cons |   25.27722   .6626819    38.14   0.000    23.97839    26.57605
-----+-----

Random-effects Parameters   |   Estimate   Std. Err.     [95% Conf. Interval]
-----+-----
PairID: Identity  var(_cons) |   80.10376   9.410732     63.62858    100.8448 → level-3 intercept var
-----+-----
Case: Unstructured
      var(time) |   1.178443   .1805631     .8727425    1.591224 → level-2 linear var
      var(_cons) |   44.31214   5.257737     35.11767    55.91389 → level-2 intercept var
      cov(time,_cons) |   1.622178   .7900245     .0737584    3.170598 → level-2 int-linear cov
-----+-----
      var(Residual) |   15.12274   .8324702     13.57607    16.84563 → level-1 residual var
-----+-----
LR test vs. linear regression:      chi2(4) = 1768.94   Prob > chi2 = 0.0000
-----+-----
      Model |   Obs    ll(null)    ll(model)    df         AIC         BIC
-----+-----
      . |   337          .   -5872.999     9        11764    11798.38
-----+-----

.      estimates store RandLin2
.      lrtest RandLin2 FixQuad

This model has 2 new variance components at level 2:
random linear slope and intercept-slope covariance.
Do we need the random linear slope for twin?
Yes,  $-2\Delta LL(2) = 132.0, p < .001$ 

Likelihood-ratio test      LR chi2(2) = 132.05
(Assumption: FixQuad nested in RandLin2)      Prob > chi2 = 0.0000

```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative. Note: LR tests based on REML are valid only when the fixed-effects specification is identical for both models.

Model 2c: Fixed Quadratic, Random Linear Slope at Levels 2 and 3

Level 1: $\text{Info}_{tij} = \beta_{0ij} + \beta_{1ij}(\text{Age}_{tij} - \text{PairAge1}_j) + \beta_{2ij}(\text{Age}_{tij} - \text{PairAge1}_j)^2 + e_{tij}$

Level 2:

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

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

Quadratic Time: $\beta_{2ij} = \delta_{20j}$

Level 3:

Intercept: $\delta_{00j} = \gamma_{000} + \gamma_{001}(\text{PairAge1}_j - 85) + V_{00j}$

Linear Time: $\delta_{10j} = \gamma_{100} + V_{10j}$ ←

Quadratic Time: $\delta_{20j} = \gamma_{200}$

```

TITLE "SAS Model 2c: Add Random Linear Slope for Pair";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = BFace85 time time*time / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT time / TYPE=UN SUBJECT=PairID;          * Level 3;
  RANDOM INTERCEPT time / TYPE=UN SUBJECT=PairID*TwinID;    * Level 2; RUN;

TITLE "SPSS Model 2c: Add Random Linear Slope for Pair".
MIXED info BY PairID TwinID WITH BFace85 time
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED = BFace85 time time*time
  /RANDOM = INTERCEPT time | SUBJECT(PairID) COVTYPE(UN)
  /RANDOM = INTERCEPT time | SUBJECT(PairID*TwinID) COVTYPE(UN).

* STATA Model 2c: Add Random Linear Slope for Pair
xtmixed info c.BFace85 c.time c.time#c.time , || PairID: time, covariance(unstructured) ///
  || Case: time , variance reml covariance(unstructured)
estat ic, n(337)
estimates store RandLin23
lrtest RandLin23 RandLin2

```

STATA output:

```

Log restricted-likelihood = -5872.6076      Wald chi2(3)      =      182.94
                                           Prob > chi2      =      0.0000
-----
      info |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      BFace85 |  -.7438709   .190867   -3.90   0.000   -1.117963   -.3697784
        time |  -.1429383   .133292   -1.07   0.284   -.4041859   .1183093
c.time#c.time |  -.1016908   .0165408   -6.15   0.000   -.1341103   -.0692713
        _cons |   25.25502   .6639108   38.04   0.000    23.95378    26.55626
-----

Random-effects Parameters |   Estimate    Std. Err.      [95% Conf. Interval]
-----+-----
PairID: Unstructured
      var(time) |   .0640187   .1696797     .000355    11.5449 → level-3 linear var
      var(_cons) |   80.86105   9.503706    64.22388   101.8081 → level-3 intercept var
      cov(time,_cons) |  -.7329904   .9257944   -2.547514    1.081533 → level-3 int-linear cov
-----
Case: Unstructured
      var(time) |   1.116498   .2415957     .7305816    1.706266 → level-2 linear var
      var(_cons) |   44.00753   5.22105    34.87711    55.52819 → level-2 intercept var
      cov(time,_cons) |   1.957119   .8826687     .2271198    3.687117 → level-2 int-linear cov
-----
      var(Residual) |   15.11455   .8311075    13.57031    16.8345 → level-1 residual var
-----
LR test vs. linear regression:      chi2(6) = 1769.73   Prob > chi2 = 0.0000
-----
      Model |   Obs    ll(null)   ll(model)    df       AIC       BIC
-----+-----
      . |   337      .   -5872.608    11    11767.22    11809.24
.      estimates store RandLin23
.      lrtest RandLin23 RandLin2

```

This model has 2 new variance components at level 3: random linear slope and intercept-slope covariance. **Do we need the random linear slope for pair, too?** No, $-2\Delta LL(2) = 0.8$, $p = .67$

```

Likelihood-ratio test
(Assumption: RandLin2 nested in RandLin23)      LR chi2(2) =      0.78
                                                  Prob > chi2 =      0.6759

```

ICC of person within pair:

For Intercepts = $80.86 / (80.86 + 44.01) = .65$

For Slopes = $0.06 / (0.06 + 1.12) = .05 (\approx 0)$

Because the ICC for the slope at the pair level is not significantly different from 0, we will remove it.

TWO EQUIVALENT MODELS: PERSON-MEAN-CENTERING VS. GRAND-MEAN-CENTERING

Model 3a: Separate Effects of Grip Strength at Each Level via Person-Mean-Centering

```

TITLE "SAS Model 3a: Grip Strength at each level via PERSON MEAN CENTERING";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = BFace85 time time*time WPgripp WFgripp BFgripp9
    / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID; * Level 3;
  RANDOM INTERCEPT time / TYPE=UN SUBJECT=PairID*TwinID; * Level 2;
  ESTIMATE "Level-2 Contextual Effect" WFgripp 1 WPgripp -1;
  ESTIMATE "Level-3 Contextual Effect" BFgripp9 1 WFgripp -1; RUN;

TITLE "SPSS Model 3a: Grip Strength at each level via PERSON MEAN CENTERING".
MIXED info BY PairID TwinID WITH BFace85 time WPgripp WFgripp BFgripp9
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED = BFace85 time time*time WPgripp WFgripp BFgripp9
  /RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)
  /RANDOM = INTERCEPT time | SUBJECT(PairID*TwinID) COVTYPE(UN)
  /TEST = "Level-2 Contextual Effect" WFgripp 1 WPgripp -1
  /TEST = "Level-3 Contextual Effect" BFgripp9 1 WFgripp -1.

* STATA Model 3a: Grip Strength at each level via PERSON MEAN CENTERING
xtmixed info c.BFace85 c.time c.time#c.time c.WPgripp c.WFgripp c.BFgripp9 , ///
  || PairID: , covariance(unstructured) || Case: time, variance reml covariance(unstructured)
  estat ic, n(337)
  lincom 1*c.WFgripp - 1*c.WPgripp // Level-2 Contextual Effect
  lincom 1*c.BFgripp9 - 1*c.WFgripp // Level-3 Contextual Effect

```

Model 3b: Testing 3-Level Convergence of Grip Strength Effects via Grand-Mean-Centering

```

TITLE "SAS Model 3b: Grip Strength Convergence across levels via GRAND MEAN CENTERING";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = BFace85 time time*time TVgripp9 BPgripp9 BFgripp9
    / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID; * Level 3;
  RANDOM INTERCEPT time / TYPE=UN SUBJECT=PairID*TwinID; * Level 2;
  ESTIMATE "Level-2 Within-Family Effect" TVgripp9 1 BPgripp9 1;
  ESTIMATE "Level-3 Between-Pair Effect" TVgripp9 1 BPgripp9 1 BFgripp9 1; RUN;

TITLE "SPSS Model 3b: Grip Strength Convergence across levels via GRAND MEAN CENTERING".
MIXED info BY PairID TwinID WITH BFace85 time TVgripp9 BPgripp9 BFgripp9
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV
  /FIXED = BFace85 time time*time TVgripp9 BPgripp9 BFgripp9
  /RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)
  /RANDOM = INTERCEPT time | SUBJECT(PairID*TwinID) COVTYPE(UN)
  /TEST = "Level-2 Within-Family Effect" TVgripp9 1 BPgripp9 1
  /TEST = "Level-3 Between-Pair Effect" TVgripp9 1 BPgripp9 1 BFgripp9 1.

* STATA Model 3b: Grip Strength Convergence across levels via GRAND MEAN CENTERING
xtmixed info c.BFace85 c.time c.time#c.time c.TVgripp9 c.BPgripp9 c.BFgripp9 , ///
  || PairID: , covariance(unstructured) || Case: time, variance reml covariance(unstructured)
  estat ic, n(337)
  lincom 1*c.TVgripp9 + 1*c.BPgripp9 // Level-2 Within-Family Effect
  lincom 1*c.TVgripp9 + 1*c.BPgripp9 + 1*c.BFgripp9 // Level-3 Between-Pair Effect

```

STATA output:

```

Wald chi2(6) = 270.77
Log restricted-likelihood = -5838.9589 Prob > chi2 = 0.0000
-----

```

| Random-effects Parameters | | Estimate | Std. Err. | [95% Conf. Interval] | | |
|--------------------------------|-----------------|-----------|-----------|----------------------|----------|----------|
| | | | | | | |
| PairID: Identity | var(_cons) | 71.39084 | 8.596088 | 56.38327 | 90.39297 | |
| | | | | | | |
| Case: Unstructured | | | | | | |
| | var(time) | .9945399 | .1647409 | .7188274 | 1.376004 | |
| | var(_cons) | 41.90059 | 5.043539 | 33.09495 | 53.04917 | |
| | cov(time,_cons) | 1.224168 | .7247081 | -.1962339 | 2.64457 | |
| | | | | | | |
| | var(Residual) | 15.31228 | .8413071 | 13.74903 | 17.05327 | |
| | | | | | | |
| LR test vs. linear regression: | | chi2(4) = | 1700.54 | Prob > chi2 = 0.0000 | | |
| | | | | | | |
| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
| | | | | | | |
| . | 337 | . | -5838.959 | 12 | 11701.92 | 11747.76 |

Because the models we will examine for grip strength are equivalent, the variance components and fit statistics are the same for both.

Model 3a: Separate Effects of Grip Strength at Each Level via Person-Mean-Centering

Level 1: $\text{Info}_{ij} = \beta_{0ij} + \beta_{1ij}(\text{Age}_{ij} - \text{PairAge1}_j) + \beta_{2ij}(\text{Age}_{ij} - \text{PairAge1}_j)^2 + \beta_{3ij}(\text{Grip}_{ij} - \overline{\text{Grip}}_{ij}) + e_{ij}$

Level 2:

Intercept: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(\overline{\text{Grip}}_{ij} - \overline{\text{Grip}}_j) + U_{0ij}$ **Within-person grip (WPgripp)**

Linear Time: $\beta_{1ij} = \delta_{10j} + U_{1ij}$ **Within-family grip (WFgripp)**

Quadratic Time: $\beta_{2ij} = \delta_{20j}$

Within-Person Grip: $\beta_{3ij} = \delta_{30j}$

Level 3:

Intercept: $\delta_{00j} = \gamma_{000} + \gamma_{001}(\text{PairAge1}_j - 85) + \gamma_{002}(\overline{\text{Grip}}_j - 9) + V_{00j}$ **Between-family grip (BFgripp9)**

Linear Time: $\delta_{10j} = \gamma_{100}$

Quadratic Time: $\delta_{20j} = \gamma_{200}$

Within-Person Grip: $\delta_{30j} = \gamma_{300}$

Within-Family Grip: $\delta_{01j} = \gamma_{010}$

| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|--|-----------|-----------|-------|-------|----------------------|-----------|
| BFAge85 | -.3463256 | .1921102 | -1.80 | 0.071 | -.7228547 | .0302034 |
| time | .0884511 | .1386187 | 0.64 | 0.523 | -.1832365 | .3601387 |
| c.time#c.time | -.1010033 | .0165328 | -6.11 | 0.000 | -.133407 | -.0685995 |
| WPgripp | .5031221 | .0979629 | 5.14 | 0.000 | .3111184 | .6951259 |
| WFgripp | .9143609 | .2251347 | 4.06 | 0.000 | .4731049 | 1.355617 |
| BFgripp9 | 1.511445 | .2463971 | 6.13 | 0.000 | 1.028515 | 1.994374 |
| _cons | 27.04317 | .7528849 | 35.92 | 0.000 | 25.56754 | 28.5188 |
| . lincom 1*c.WFgripp - 1*c.WPgripp // Level-2 Contextual Effect | | | | | | |
| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
| (1) | .4112388 | .241575 | 1.70 | 0.089 | -.0622395 | .8847171 |
| . lincom 1*c.BFgripp9 - 1*c.WFgripp // Level-3 Contextual Effect | | | | | | |
| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
| (1) | .5970839 | .3275152 | 1.82 | 0.068 | -.044834 | 1.239002 |

Model 3b: Testing 3-Level Convergence of Grip Strength Effects via Grand-Mean-Centering

Level 1: $\text{Info}_{tij} = \beta_{0ij} + \beta_{1ij}(\text{Age}_{tij} - \text{PairAge1}_j) + \beta_{2ij}(\text{Age}_{tij} - \text{PairAge1}_j)^2 + \beta_{3ij}(\text{Grip}_{tij} - 9) + e_{tij}$

Level 2:

Intercept: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(\overline{\text{Grip}}_j - 9) + U_{0ij}$ Within-person grip (TVgripp9)

Linear Time: $\beta_{1ij} = \delta_{10j} + U_{1ij}$ Contextual between-person grip (BPgripp9)

Quadratic Time: $\beta_{2ij} = \delta_{20j}$

Within-Person Grip: $\beta_{3ij} = \delta_{30j}$

Level 3:

Intercept: $\delta_{00j} = \gamma_{000} + \gamma_{001}(\text{PairAge1}_j - 85) + \gamma_{002}(\overline{\text{Grip}}_j - 9) + V_{00j}$ Contextual between-family grip (BFgripp9)

Linear Time: $\delta_{10j} = \gamma_{100}$

Quadratic Time: $\delta_{20j} = \gamma_{200}$

Within-Person Grip: $\delta_{30j} = \gamma_{300}$

Within-Family Grip: $\delta_{01j} = \gamma_{010}$

| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|---------------|-----------|-----------|-------|-------|----------------------|-----------|
| BFace85 | -.3463256 | .1921102 | -1.80 | 0.071 | -.7228547 | .0302034 |
| time | .0884511 | .1386187 | 0.64 | 0.523 | -.1832365 | .3601387 |
| c.time#c.time | -.1010033 | .0165328 | -6.11 | 0.000 | -.133407 | -.0685995 |
| TVgripp9 | .5031221 | .0979629 | 5.14 | 0.000 | .3111184 | .6951259 |
| BPgripp9 | .4112388 | .241575 | 1.70 | 0.089 | -.0622395 | .884717 |
| BFgripp9 | .5970839 | .3275152 | 1.82 | 0.068 | -.044834 | 1.239002 |
| _cons | 27.04317 | .7528849 | 35.92 | 0.000 | 25.56754 | 28.5188 |

level-1, total within-person
level-1 = level-2 effect?
level-2 = level-3 effect?

. lincom 1*c.TVgripp9 + 1*c.BPgripp9 // Level-2 Within-Family Level 2 Effect

| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|------|----------|-----------|------|-------|----------------------|----------|
| (1) | .9143609 | .2251347 | 4.06 | 0.000 | .4731049 | 1.355617 |

. lincom 1*c.TVgripp9 + 1*c.BPgripp9 + 1*c.BFgripp9 // Level-3 Between-Pair, Level 3 Effect

| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
|------|----------|-----------|------|-------|----------------------|----------|
| (1) | 1.511445 | .2463971 | 6.13 | 0.000 | 1.028515 | 1.994374 |

It appears that although there is a significant positive effect of grip strength at each level, those effects may not be significantly different in magnitude. Accordingly, let's simplify the model by removing the contextual effect at level 3, such that the level-2 and level-3 effects of grip strength are assumed to be the same.

Model 3c: Separate Effects of Grip Strength at Level 1 and Level-2&3 via Person-Mean-Centering

```
TITLE "SAS Model 3c: Grip Strength at Level 1 and Level 2&3 via PERSON MEAN CENTERING";
PROC MIXED DATA=work.octodata NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
  CLASS PairID TwinID;
  MODEL info = BFace85 time time*time WPgripp BPgripp9 / SOLUTION DDFM=Satterthwaite;
  RANDOM INTERCEPT / TYPE=UN SUBJECT=PairID; * Level 3;
  RANDOM INTERCEPT time / TYPE=UN SUBJECT=PairID*TwinID; * Level 2;
  ESTIMATE "Level-2&3 Contextual Effect" BFgripp9 1 WPgripp -1; RUN;
```

TITLE "SPSS Model 3c: Grip Strength at Level 1 and Level 2&3 via PERSON MEAN CENTERING".

MIXED info BY PairID TwinID WITH BFace85 time WPgripp BPgripp9

/METHOD = REML

/PRINT = SOLUTION TESTCOV

/FIXED = BFace85 time time*time WPgripp BPgripp9

/RANDOM = INTERCEPT | SUBJECT(PairID) COVTYPE(UN)

/RANDOM = INTERCEPT time | SUBJECT(PairID*TwinID) COVTYPE(UN)

/TEST = "Level-2&3 Contextual Effect" BPgripp9 1 WPgripp -1.

*** STATA Model 3c: Grip Strength at Level 1 and Level 2&3 via PERSON MEAN CENTERING**

xtmixed info c.BFace85 c.time c.time#c.time c.WPgripp c.BPgripp9 , ///

|| PairID: , covariance(unstructured) || Case: time, variance reml covariance(unstructured)
estat ic, n(337)

lincom 1*c.BPgripp9 - 1*c.WPgripp // Level-2&3 Contextual Effect

Level 1: $\text{Info}_{ij} = \beta_{0ij} + \beta_{1ij}(\text{Age}_{ij} - \text{PairAge1}_j) + \beta_{2ij}(\text{Age}_{ij} - \text{PairAge1}_j)^2 + \beta_{3ij}(\text{Grip}_{ij} - 9) + e_{ij}$

Level 2:

Intercept: $\beta_{0ij} = \delta_{00j} + \delta_{01j}(\overline{\text{Grip}}_{ij} - 9) + U_{0ij}$ **Within-person grip (WPgripp)**

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

Quadratic Time: $\beta_{2ij} = \delta_{20j}$ **Between-person grip (BPgripp9)**

Within-Person Grip: $\beta_{3ij} = \delta_{30j}$

Level 3:

Intercept: $\delta_{00j} = \gamma_{000} + \gamma_{001}(\text{PairAge1}_j - 85) + V_{00j}$

Linear Time: $\delta_{10j} = \gamma_{100}$

Quadratic Time: $\delta_{20j} = \gamma_{200}$

Within-Person Grip: $\delta_{30j} = \gamma_{300}$

Within-Family Grip: $\delta_{01j} = \gamma_{010}$

STATA output:

| | | | | | | |
|--|-----------------|-----------|-----------|--------------|----------------------|----------------------|
| Log restricted-likelihood = -5840.4202 | | | | Wald chi2(5) | = | 267.14 |
| | | | | Prob > chi2 | = | 0.0000 |
| ----- | | | | | | |
| | info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
| ----- | | | | | | |
| | BFace85 | -.4275427 | .1874185 | -2.28 | 0.023 | -.7948762 -.0602091 |
| | time | .0904993 | .1386117 | 0.65 | 0.514 | -.1811746 .3621732 |
| c.time#c.time | | -.1011004 | .0165312 | -6.12 | 0.000 | -.133501 -.0686997 |
| | WPgripp | .5071004 | .097934 | 5.18 | 0.000 | .3151532 .6990476 |
| | BPgripp9 | 1.184309 | .1695658 | 6.98 | 0.000 | .8519664 1.516652 |
| | _cons | 26.47672 | .687956 | 38.49 | 0.000 | 25.12835 27.82508 |
| ----- | | | | | | |
| ----- | | | | | | |
| Random-effects Parameters | | | Estimate | Std. Err. | [95% Conf. Interval] | |
| ----- | | | | | | |
| PairID: Identity | var(_cons) | | 71.96326 | 8.654375 | 56.85183 | 91.09136 |
| ----- | | | | | | |
| Case: Unstructured | | | | | | |
| | var(time) | | .9952698 | .1647394 | .7195284 | 1.376682 |
| | var(_cons) | | 41.97828 | 5.046668 | 33.16595 | 53.13209 |
| | cov(time,_cons) | | 1.234527 | .7220033 | -.180573 | 2.649628 |
| ----- | | | | | | |
| | var(Residual) | | 15.3081 | .8408553 | 13.74567 | 17.04813 |
| ----- | | | | | | |
| LR test vs. linear regression: | | | chi2(4) = | 1705.54 | Prob > chi2 = 0.0000 | |

| Model | Obs | ll(null) | ll(model) | df | AIC | BIC |
|--|----------|-----------|-----------|-------|----------------------|----------|
| . | 337 | . | -5840.42 | 11 | 11702.84 | 11744.86 |
| . lincom 1*c.BPgripp9 - 1*c.WPgripp // Level-2&3 Contextual Effect | | | | | | |
| info | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
| (1) | .6772089 | .1925793 | 3.52 | 0.000 | .2997605 | 1.054657 |

One could then test interactions, keeping in mind the need to differentiate effects across all three levels as needed...

Sample Results Section (note this combines across models somewhat)

The extent of individual change in crystallized intelligence (as measured by the information test) and the relationship between intelligence, age, and grip strength was examined in a sample of 337 same-sex twins measured every two years for up to five occasions. Multilevel models were estimated using restricted maximum likelihood. The significance of fixed effects was evaluated with individual Wald tests (i.e., of estimate / SE), whereas random effects were evaluated via likelihood ratio tests (i.e., $-2\Delta LL$ with degrees of freedom equal to the number of new random effects variances and covariances).

A two-level empty means, random intercept model of time nested within person was initially specified and indicated that 83% of the information test outcome variance was between persons. The addition of a random intercept for twin pair resulted in a significant improvement in model fit, $-2\Delta LL(1) = 101.5$, $p < .001$, and revealed that 64% of that between-person variance was due to twin pair (i.e., shared variance between twins from the same pair). Thus, a three-level model was necessary, given that 17% of the variance was at level 1 (within persons over time), 30% was at level 2 (within pairs), and 53% was at level 3 (between pairs). A three-level empty means, random intercept model to decompose the variance in time-varying age revealed that 47% was between pairs (given that the twins initially varied in age from 80 to 100), whereas the remaining 53% was within persons over time—there was no level-2 age variance. Thus, the level-3 cross-sectional and level-1 longitudinal effects of age were modeled separately using baseline age (centered at 85) and time in study, respectively. Preliminary analyses revealed that a linear effect of age at baseline and a quadratic effect of time in study resulted in the best-fitting model to describe mean change. Although a random linear time slope for twin significantly improved model fit, $-2\Delta LL(2) = 132.0$, $p < .001$, the subsequent addition of a random linear time slope for twin pair did not significantly improve model fit, $-2\Delta LL(2) = 0.8$, $p = .67$, indicating that the 5% of the random linear time slope variance that was due to twin pair was not distinguishable from 0. As a result, the random linear time slope was retained at the twin level only (i.e., level 2 but not level 3).

The prediction of the information test outcome from time-varying grip strength was then examined. A three-level empty means, random intercept model to decompose the variance in grip strength revealed that 36% was between pairs, 29% was within pairs, and 35% was within persons over time. Predictors for grip strength were included via person-mean-centering, in which the within-person effect was represented by the deviation of each occasion's grip strength around each person's mean, the within-pair effect was represented by the deviation of each twin's mean grip strength around each pair's mean, and the between-pair effect was represented by the family mean grip strength (centered at 9 pounds). There was a significant main effect of grip strength at each level. Within persons, for every additional pound of grip strength more than one's own mean, information test at that occasion was expected to be higher by 0.50. Within pairs, for every additional pound of person mean grip strength more than one's family mean, information test for that twin was expected to be higher by 0.91. Between pairs, for every additional pound of family mean grip strength more than other families, information test for the twin pair was expected to be higher by 1.51.

Contextual effects for the differences in effect size across levels were requested using separate statements (i.e., as would be provided directly using grand-mean-centering but including the person and pair means). The pair-level contextual effect was not significant, indicating that the within-pair and between-pair effects were equivalent. Consequently, the model was re-specified to include within-person grip strength, as described previously, along with between-person grip strength to represent the combination of the twin and pair levels, calculated as each person's mean grip strength centered at 9. The between-person effect of grip strength was significant, such that for every additional pound of mean grip strength more than other people, information test for that twin was expected to be higher by 1.18. This effect was significantly larger than the within-person effect of grip strength of 0.51 (i.e., a significant person contextual effect), and thus both the within-person and between-person effects of grip strength were retained.