Lecture 5: Multilevel Models (MLMs) for Longitudinal Data

- Topics:
  - From clustered to longitudinal: terminology and considerations of "time"
  - 3 options for modeling longitudinal dependency
  - Random intercepts vs random time slopes
  - Time-varying predictors (and 1 example)
  - Nonlinear change (and 3 examples)
  - When to eject to a "truly" multivariate MLM (via SEM or M-SEM)

From Clustered to Longitudinal

- The good news: Much of what you've learned about univariate (=one outcome) **two-level** MLMs for clustered data readily translate into longitudinal data
  - **Clustered**: L1 = people, L2 = clusters (e.g., schools, camps, teams)
  - **Longitudinal ("panel")**: L1 = occasions, L2 = people → **what we will work with for now!**

- **Three-level** designs within each type are also possible
  - Clustered: L1 = people, L2 = schools, L3 = districts
  - "Intensive" longitudinal: L1 = occasions, L2 = days, L3 = people

- **Combination designs** are also possible, but time may not be L1
  - "Clustered longitudinal": L1 = occasions, L2 = people, L3 = clusters
  - One-shot surveys: L1 = respondents, L2 = occasions, L3 = clusters
  - Cross-sectional surveys: L1 = respondents, L2 = interviewers, L3 = occasions

Sources of Longitudinal Relations

- **Level 2 = Between-Person*** (**BP**) Variation:
  - "**INTER**-individual differences" from "**time-invariant**" **measures**
  - All longitudinal studies that begin as cross-sectional studies have this

- **Level 1 = Within-Person*** (**WP**) Variation:
  - "**INTRA**-individual differences" from "**time-varying**" **measures**
  - Only longitudinal studies can provide this extra type of 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 (even "time" in accelerated designs in which people start at different points in "time", such as age or grade in school)
  - BP = more/less *than other people*; WP = more/less *than usual*

* I use "person" here, but the "between" units can be anything that is measured repeatedly (e.g., schools, countries, companies…)

From Clustered to Longitudinal

- Predictors and their terminology shifts accordingly:
  - **L2** predictors are now person "**time-invariant**" characteristics (e.g., at baseline)
  - **L1** predictors are now occasion-specific "**time-varying**" measures (including "time")
  - My subscripts will shift from pc (person and cluster) to ti (time and individual)

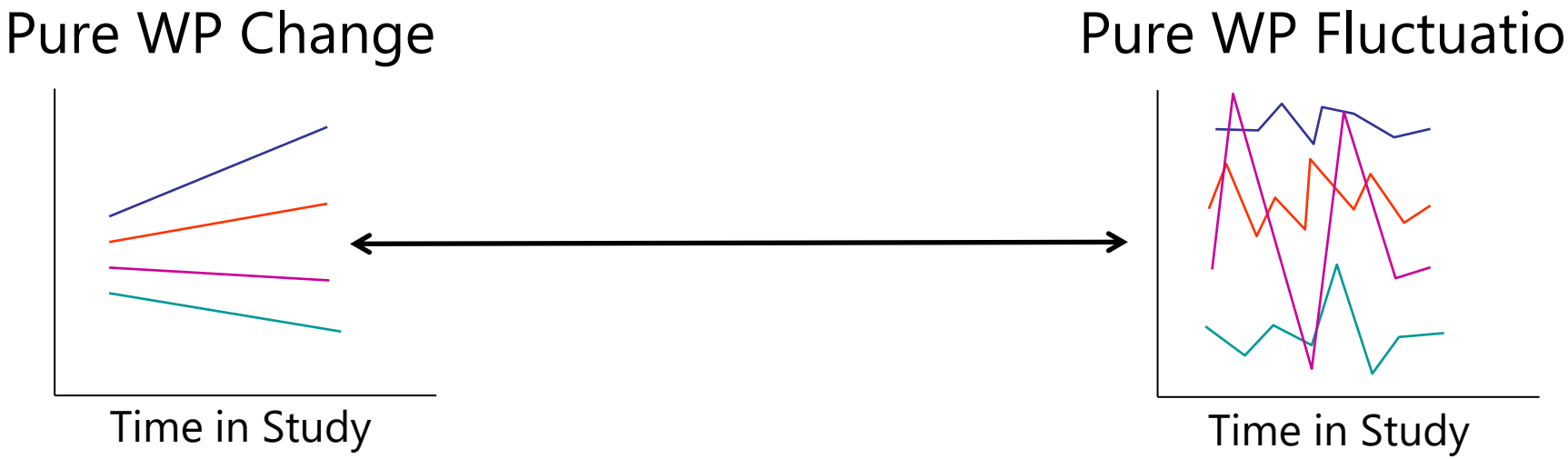
- As in clustered data, we can use random effects to capture sources of dependency (now regarding the correlation of occasions from the same person)
  - **Fixed effects** are still **constants** estimated for L1 or L2 predictors
  - **Random effects** are still **variables**—intercepts and slopes of L1 time-varying predictors can vary across persons (with estimated variance across L2 persons now)

- However, "**time**" is a **salient design element** that must be considered in several ways that do not apply to clustered data
  - You should NOT plan on just fitting a random linear slope for time and calling it good!
  - This is why I have spent 2 semesters on longitudinal models alone (and easily could have spent 3)!
    - First course: RM ANOVA/ ACS models, random effects, nonlinear change, L2 predictors (Hoffman 2015 ch. 1–7)
    - Second course: L1 predictors, accelerated designs (ch. 8–10), MLM → SEM, multivariate change and mediation

A Longitudinal Data Continuum

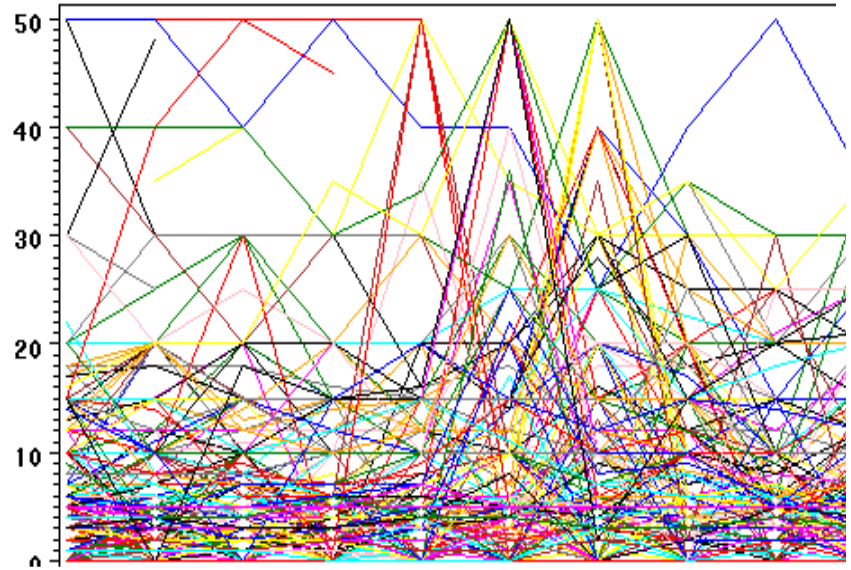
- **Within-Person (WP) Change**: Expect systematic effect(s) of time in study
  - e.g., "(Latent) Growth Curve Models" → **Time is deliberately sampled to capture change**
  - Lends itself to questions about patterns of change and moderation thereof

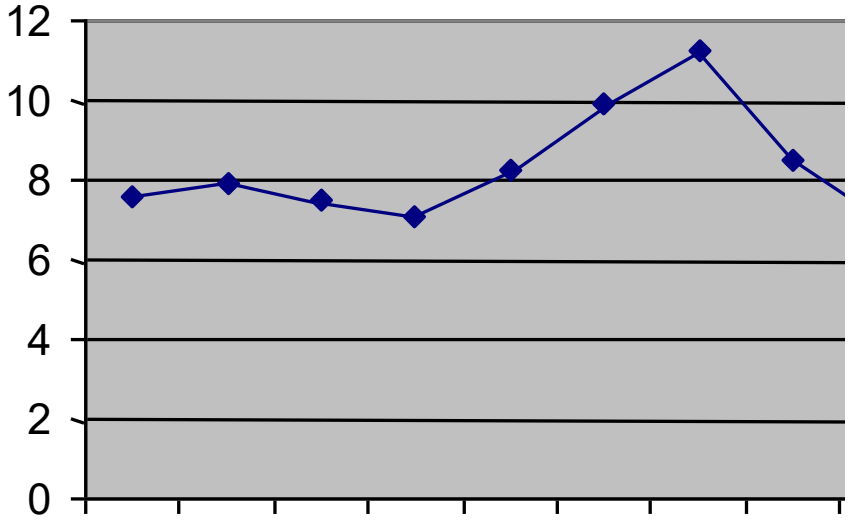
- **Within-Person (WP) Fluctuation**: Few *expected* effects of time in study
  - e.g., "Ecological Momentary Assessment" or "Intensive" LD → **Time in study is not the point!**
  - Outcomes are expected to just fluctuate over time (e.g., emotion, mood, stress), but time may still be relevant in other ways (e.g., day of week, circadian rhythms)
  - Lends itself to questions about WP effects of relative changes and inconsistency



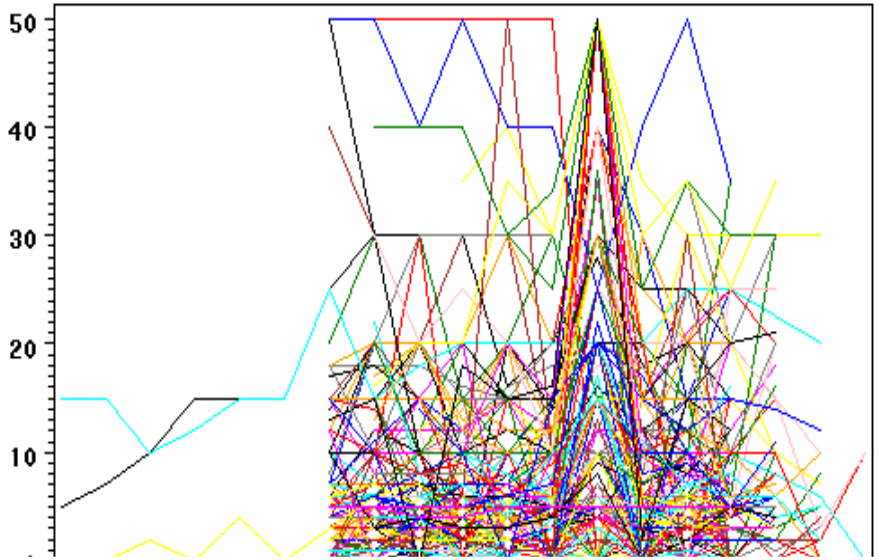
Unexpected Change over Time

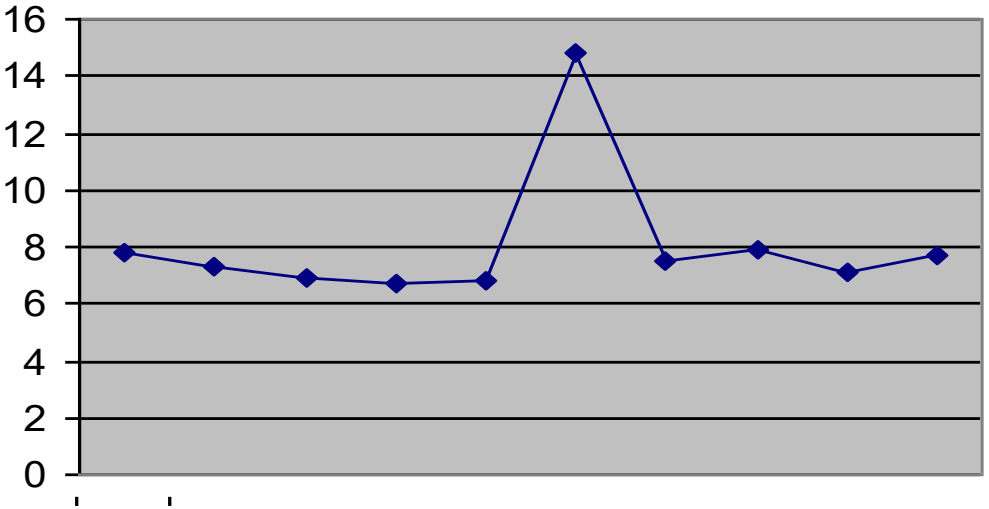
#Drinks by Interview Week Number





#Drinks by Time to Spring Break





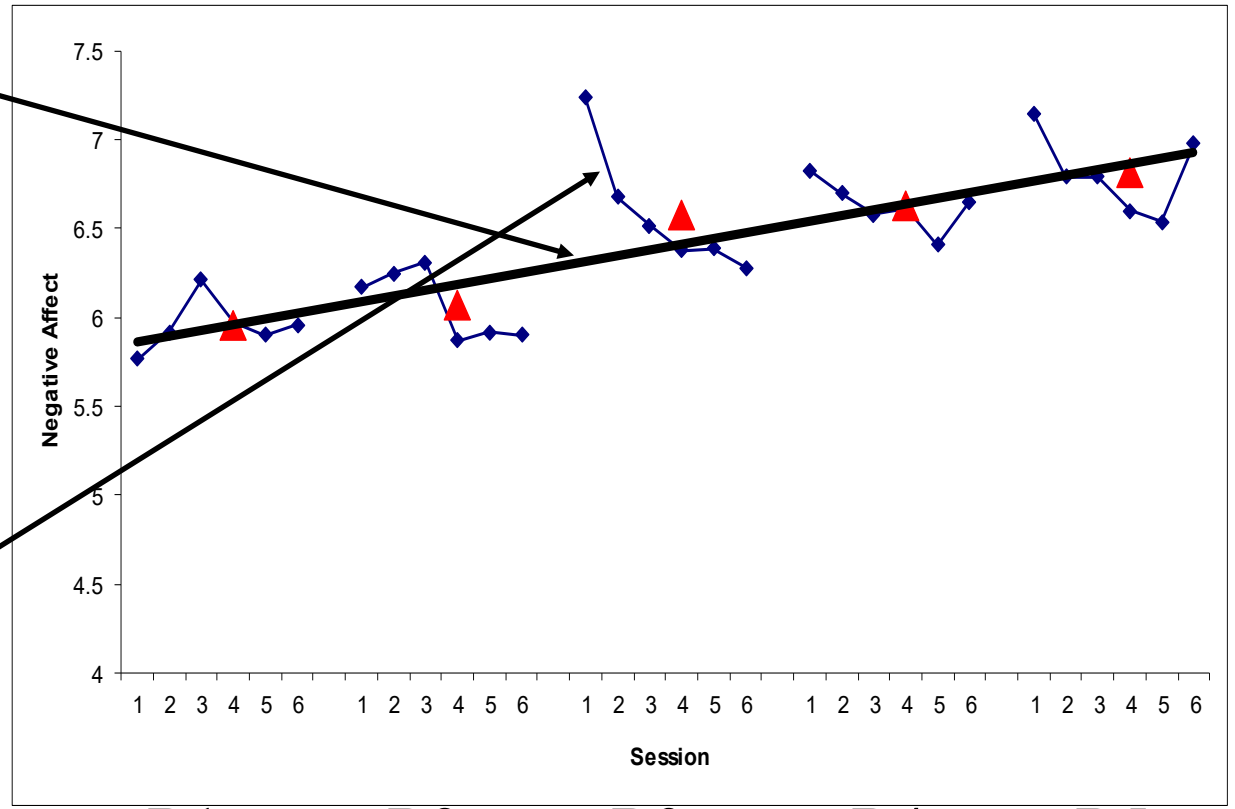
Different Sources of Change over Time

How does negative affect change over short (days) and long (months) intervals?

Stawski & Sliwinski, GSA 2005

Aging
.24/burst (6 mos.)
$p < .0001$

Reactivity
-.07/session
$p < .01$



Considerations of "Time" in Longitudinal Data

- What aspects of "**time**" are relevant?
  - **WP change**: e.g., time in study, age, grade, time to/from event
  - **WP fluctuation**: e.g., time of day, day of week, day in study

- Does time vary **within persons (WP)** AND **between persons (BP)**?
  - If people differ in time at the study beginning (e.g., accelerated designs), we will need to **differentiate BP time effects from WP time effects to avoid smushing time slopes!**
  - If there is more than one kind of WP "time" (e.g., occasions within days), we will need to **differentiate distinct sources of WP time effects**

- Is time *balanced* or *unbalanced*?
  - **Balanced** = **shared** measurement schedule (not necessarily equal interval) → T1, T2, T3…
    - Although some people may miss some occasions, making their data "incomplete"
  - **Unbalanced** = people have **different** possible time values → T1.2, T2.4, T3.7… T1.3, T2.1, T3.4…
    - By definition, the possible outcomes are at least partially "incomplete" across persons
    - This may be a consequence of using a time metric that also varies between persons

Options for Modeling Longitudinal Dependency

- Outcomes from the same sampling unit (i.e., person) will have one or more sources of **dependency → correlated residuals**
  - If ignored, dependency in a longitudinal outcome will result in incorrect fixed effect standard errors and *p*-values (well-known problem)
  - If ignored, dependency in a longitudinal predictor variable will result in incorrect fixed effect estimates, too (i.e., smushed effects, a relatively less well-known problem)
    - Because time-varying predictors have both BP and WP variation—same idea as in clustered data!

- The sources of residual correlation of occasions from same person can be captured by a model in three main ways:
  1. **Fixed effects:** Add Person ID as a predictor (via *L2n*−1 binary contrasts)
  2. **"Multivariate" single-level models:** Just allow correlation and/or non-constant variance over occasions directly (not a choice in clustered data)
  3. **Add a "level" (or more):** Use random effect variances, which then introduce constant correlation (random intercept) or non-constant correlation and variance (random slopes)

Modeling Longitudinal Dependency: Option 1

- **Fixed effects:** Add Person ID as a categorical ("factor variable") predictor (estimate fixed effects of *L2n* − 1 binary contrasts for person ID)
  - Person ID **main effects** capture dependency due to **mean differences** (intercept shifts)
  - **Interactions** of Person ID with time-varying predictors (like **time**) capture other predictor-specific sources of person dependency

- Pros: Does adequately control for person dependency
  - "Fixed effects" models are very common in econometrics and some other fields
  - Does a better job in studies with "few" persons (< 15ish)
  - Allows person-specific conclusions (i.e., as in aggregated N-of-1 randomized control trials)

- Con: Does not allow prediction of WHY any individual differences occurred ☹
  - Model would be saturated with respect to BP differences (so no time-invariant predictors)

Modeling Longitudinal Dependency: Option 2

- **Multivariate single-level models**: Change model variance–covariance structure to allow (non-constant) correlation over occasions (and any residual non-constant variance) to exist
  - These "alternative covariance structure" (ACS) models are the focus of chapter 4 of Hoffman (2015)
  - Only possible given **balanced data** (all people measured at same possible occasions) and conditionally normal outcomes (i.e., not when using generalized models)

- Is the basis of old-school repeated measures (RM) "ANOVA", of which there are 2 kinds
  - "**Univariate approach**": residuals have equal variance and equal correlations across all repeated measures outcomes—but this "**compound symmetry**" pattern can only possibly hold if all **people change the same**!
  - "**Multivariate approach**": all residual variances and correlations are separately estimated—but this "**unstructured**" (MANOVA) model becomes impossible given many outcomes (especially with few people)
  - The problem with old-school RM ANOVA is OLS estimation → person-listwise deletion of missing data ☹

- Switching to REML/ML estimation uses all complete occasions AND offers more choices for patterns of residual variance and correlation beyond just compound symmetry and unstructured
  - e.g., Compound Symmetry Heterogeneous (different variances by occasion, equal correlation)
  - Traditional options that use time-lagged covariances also require balanced, equal-interval occasions: e.g., First-order auto-regressive (AR1), moving average (MA), or antedependence; Toeplitz
  - For conditionally normal outcomes given unbalanced occasions, AR1 can be extended to continuous time using lme in R; also in SAS MIXED /REPEATED as a "spatial power" SP(time) structure

Modeling Longitudinal Dependency: Option 3

- **Add one or more "levels"** → Add random effect (latent variable) variances
- Random effect = model term that each person gets their own version of; included by estimating the variance of each random effect across persons → BP differences
  - Random intercept create constant expected variance and correlation; random slopes capture patterns of non-constant variance and correlation for individual differences
  - Random effects work for any kind of outcome and for balanced or unbalanced longitudinal data

- More generally, a "**level**" is (still) a dimension of sampling that has **unexplained** outcome variability represented by 1+ random effects
  - e.g., Randomized Control Trial (RCT) of 5 occasions → 1. within-person, 2. between-person
  - e.g., EMA design of 4 observations/ day for 3 weeks → 1. within-day, 2. between-day, 3. between-person

- Occasion-specific residual correlation may still be needed beyond random effects!
  - e.g., AR1: occasions 1 "lag" apart are correlated r; 2 lags apart are correlated r^2; 3 lags apart → r^3…
  - This is the part we can safely ignore in MLMs for clustered data (no residual correlation logically possible)

Building Fixed and Random Effects of Time

- 2 questions about the possible effects of **"time"**
  (e.g., time in study in WP change; time of day or day of week in WP fluctuation):

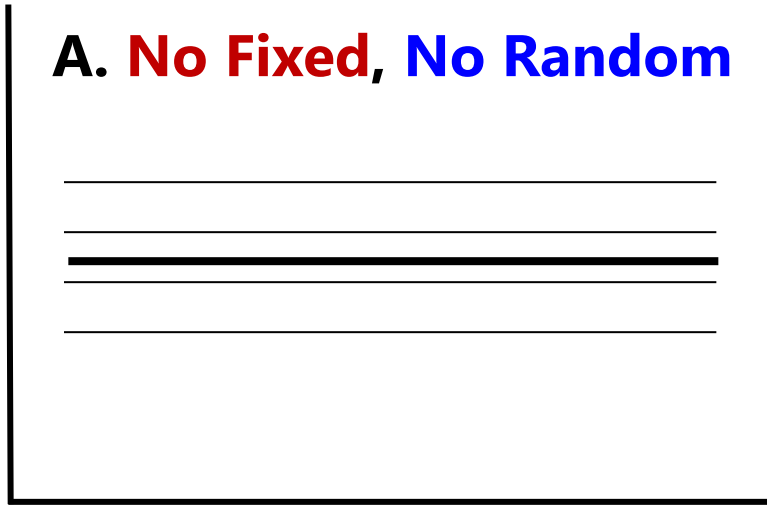
1. **Is there an effect of time on average?**
   - Is the line connecting the sample means not flat?
   - If so, you need **FIXED** effect(s) of time

2. **Does the average effect of time vary across individuals?**
   - Does each individual need their *own* version of that line?
   - If so, you need **RANDOM** effect(s) of time

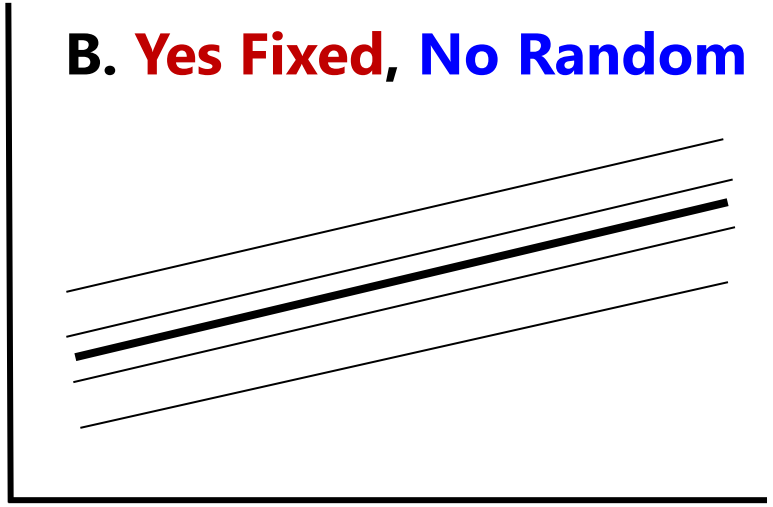
- Let's review examples using **linear time** effects to start…

Fixed and Random Effects of Time
(Note: The intercept is random in every figure)

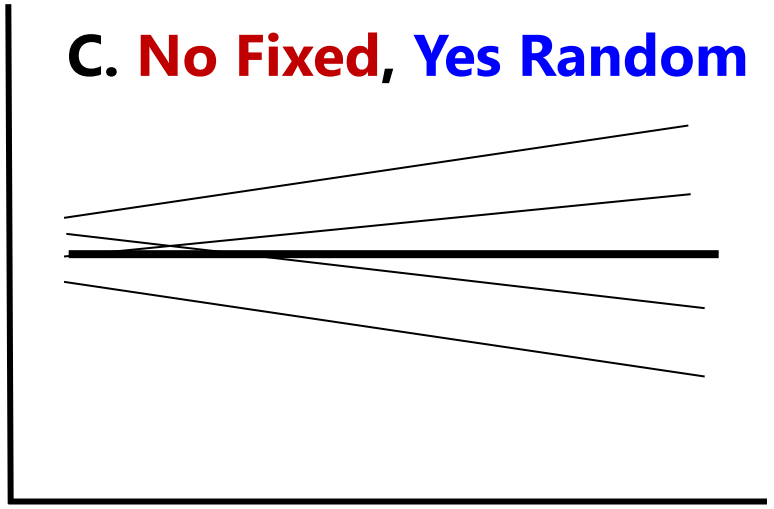
A. No Fixed, No Random



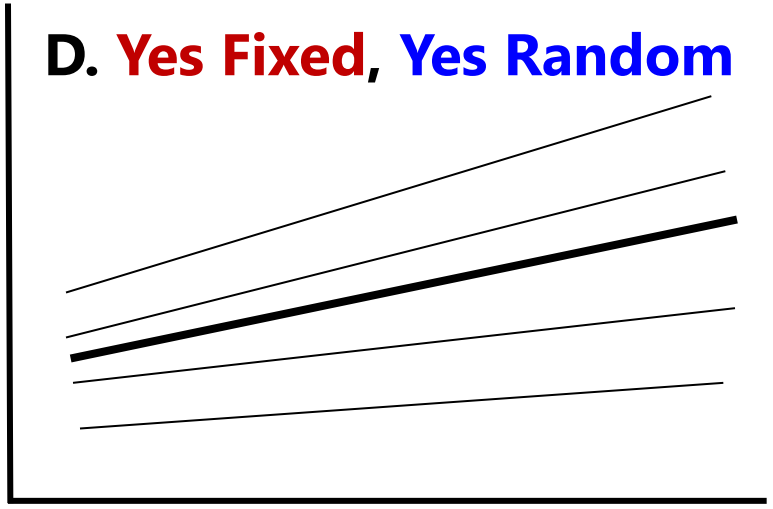
B. Yes Fixed, No Random



C. No Fixed, Yes Random



D. Yes Fixed, Yes Random



Two-Level Longitudinal Models:
A. Empty Means, Random Intercept Model

MLM Empty Model:

- Level 1 (L1):

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

- Level 2 (L2):

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

L1 Residual = occasion-specific deviation at time t from the outcome predicted by fixed and random effects for person i

Fixed Intercept = mean of person means (because no predictors yet)

L2 Random Intercept = person-specific deviation from predicted intercept

3 estimated parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

- Level-1 **Within-Person** Variance of e_{ti} → σ_e^2
- Level-2 **Between-Person** Variance of U_{0i} → $\tau_{U_0}^2$

"Composite" equation:

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

Intraclass Correlation ("ICC" = ICC_1 specifically)

$$ICC_1 = \frac{\text{Between}}{\text{Betweeen} + \text{Within}} = \frac{\text{L2 Intercept Var}}{\text{L2 Intercept Var} + \text{L1 Residual Var}} = \frac{\boldsymbol{\tau^2_{U_0}}}{\boldsymbol{\tau^2_{U_0}} + \boldsymbol{\sigma^2_e}}$$

$\boldsymbol{\tau^2_{U_0}}$ → *Why don't all **people** have the same mean outcome?*

$\boldsymbol{\sigma^2_e}$ → *Why don't all **occasions** from the same person have the same outcome?*

- ICC_1 = **Proportion** of total variance that is **between persons**
- ICC_1 = Average **correlation** of occasions from same person
- ICC_1 is a standardized expression of how much correlation there is due to **person** mean differences → $\boldsymbol{ICC_1}$ **is an effect size for *constant* person dependency**
  - Non-constant person dependency can still be created by differences across persons in the slopes of level-1 time-varying predictors (i.e., random slopes)
- Btw, **no variance has been "explained"** in an empty model → just 2 kinds of "error"
  (side note: others have disagreed, but this is a hill I will still die on)

A Random Intercept = CS Model for the Variance

Total "Marginal" data matrix is called V matrix, and each person gets their own!

$$\begin{bmatrix} \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e \end{bmatrix}$$

***N* = total outcome observations**
***L1n* = # occasions (5 here)**

In a standardized V matrix, the off-diagonal becomes ICC1

Level 2, BP Variance

Random effect variances and covariances live in the **G matrix**, shared across people

Here, 1 random intercept variance only; random slope variances could be included $\begin{bmatrix} \tau^2_{U_0} \end{bmatrix}$

To combine **G** and **R** to form **V**, **G** is pre- and post-multiplied by an ***N* x 1 Z** matrix that holds the values of the predictors with random effects (just the intercept here): $\mathbf{V}_i = \mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i^T + \mathbf{R}_i$

Level 1, WP Variance

Residual variances and covariances live in the **R matrix**, whose ingredients are shared across people (each person has own ***L1n* x *L1n* R** matrix)

$$\begin{bmatrix} \sigma^2_e & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_e \end{bmatrix}$$

Here, a "diagonal" form of 1 constant residual variance, no correlations

Diagonal is the only choice in lmer in lme4; lme in nlme has more (like AR1)

B. Fixed Linear Time, Random Intercept Model

(4 parameters: slope for change over time is **FIXED** only)

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

Multilevel Model

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

Fixed Intercept = predicted mean outcome *at time 0*

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

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

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

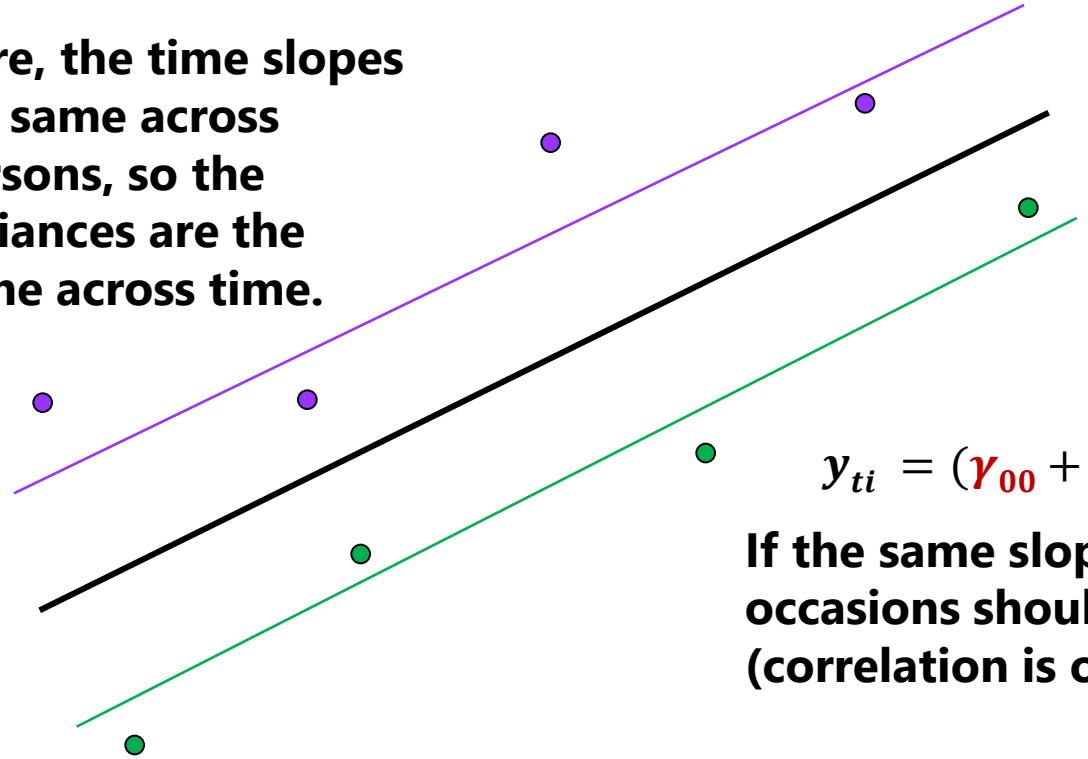
Composite Model

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

Because the effect of time is **fixed**, everyone is predicted to change exactly the same → parallel lines!

Choices in Modeling Variances: Random Intercept Only (Compound Symmetry)

Here, the time slopes are same across persons, so the variances are the same across time.



Combined V matrix:

$$\begin{bmatrix} \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e \end{bmatrix}$$

$$y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10})(Time_{ti}) + e_{ti}$$

If the same slope fits all persons, then all occasions should be equally correlated (correlation is only due to U_{0i} variance).

If the time slopes are the same across people, then people differ from each other systematically in only 1 way (i.e., their U_{0i} level) → THIS IS COMPOUND SYMMETRY.

C or D: Random Linear Time Model (6 parms)

Multilevel Model

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

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

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

Fixed Intercept = predicted mean outcome at time 0

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

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

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

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

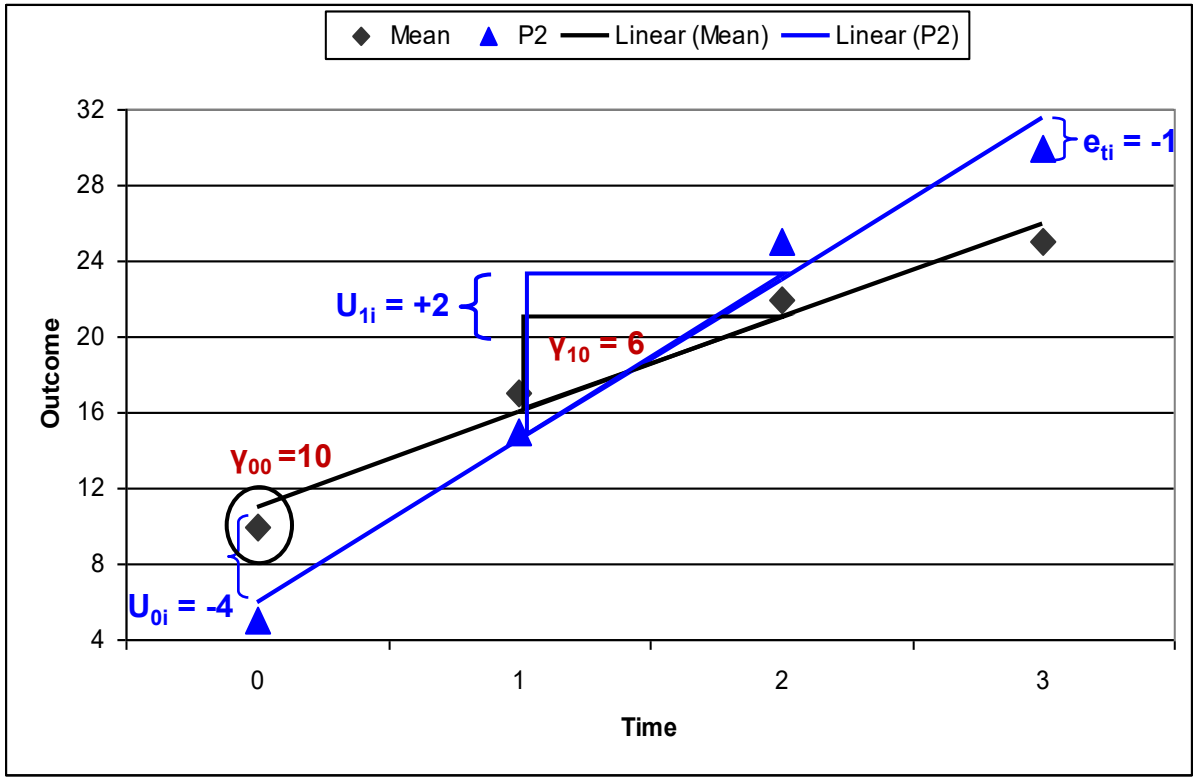
Composite Model

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

Random Linear Time Model

$$y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(Time_{ti}) + e_{ti}$$

- γ_{00}: **Fixed Intercept**
- U_{0i}: Random **Intercept** Deviation
- γ_{10}: **Fixed Slope**
- U_{1i}: Random **Slope** Deviation
- e_{ti}: error for person *i* at time *t*



6 Parameters:

2 Fixed Effects: γ_{00} Intercept, γ_{10} Slope

U_{0i} Random Intercept Variance $= \tau^2_{U_0}$

U_{1i} Random Slope Variance $= \tau^2_{U_1}$

Random Int-Slope Covariance $= \tau_{U_{01}}$

e_{ti} Residual Variance = σ^2_e

Please pretend the black line is red ☺

Random Linear Time Models Imply:

- **People differ from each other systematically in TWO ways**—in intercept (U_{0i}) and time slope (U_{1i}), which implies **TWO kinds of BP variance → 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 $\boldsymbol{\tau^2_{U_0}}$ and $\boldsymbol{\tau^2_{U_1}}$ variances in the **G** matrix), the $\boldsymbol{e_{ti}}$ **residuals** (whose variance and covariance are estimated in the **R** matrix) should be **uncorrelated with homogeneous variance across time**, as shown (or else a different type of **R** matrix is needed):

Level-2 **G** matrix (always unstructured)

$$\begin{bmatrix} \tau^2_{U_0} & \tau_{U_{10}} \\ \tau_{U_{01}} & \tau^2_{U_1} \end{bmatrix}$$

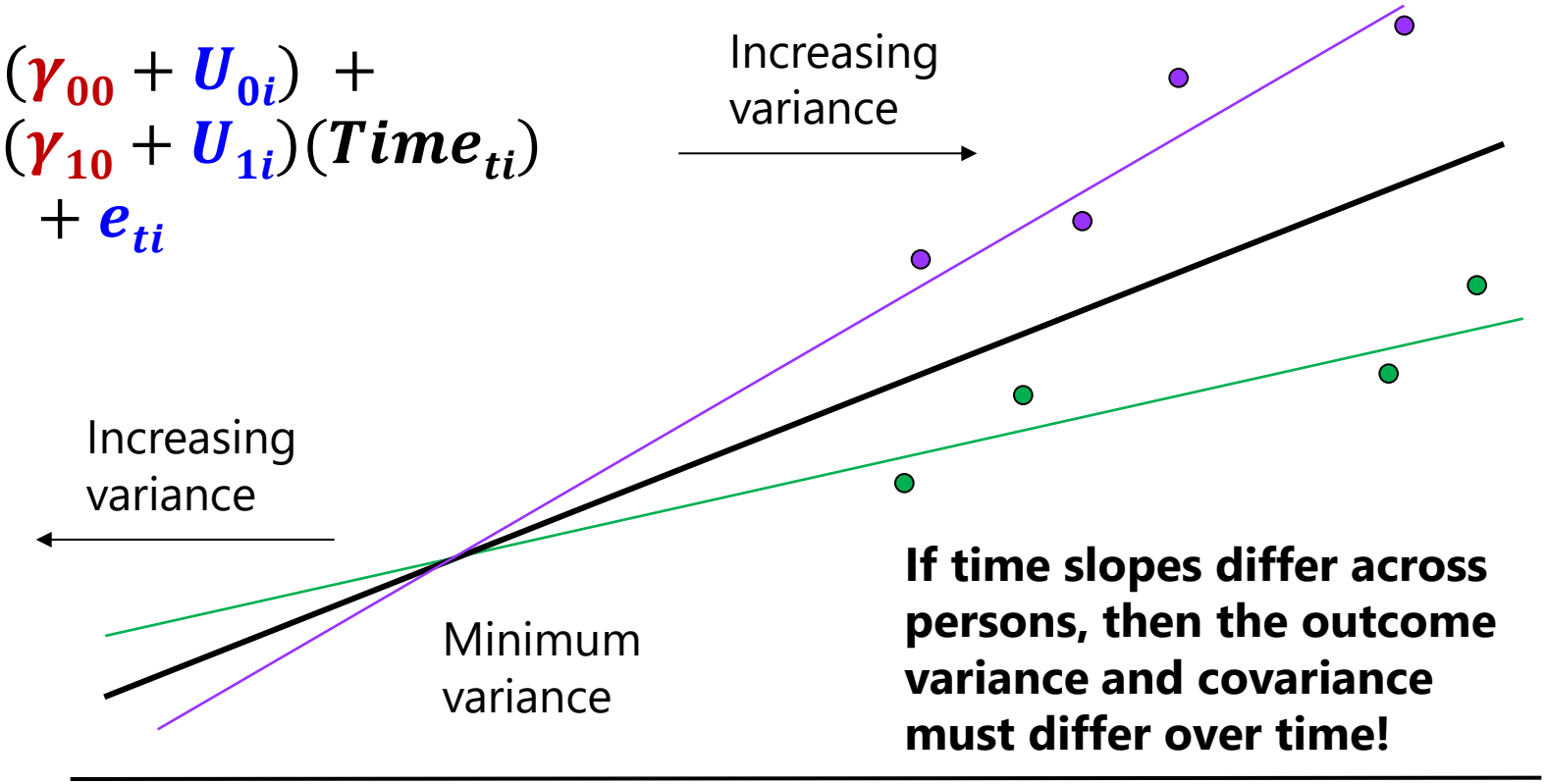
Level-1 **R** matrix (diagonal here, other kinds possible)

$$\begin{bmatrix} \sigma^2_e & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & \sigma^2_e \end{bmatrix}$$

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

Choices in Modeling Variance: Random Intercepts and Time Slopes

$$y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(Time_{ti}) + e_{ti}$$



If slopes are different across people, then people differ from each other systematically in 2 ways (U_{0i} and U_{1i})
→ this implies compound symmetry will NOT hold…

Random Linear Time Model

(6 parameters: effect of time is RANDOM)

Predicted total variances and covariances per person:

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

$$\mathbf{V}_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \tau^2_{U_0} & \tau_{U_{01}} \\ \tau_{U_{01}} & \tau^2_{U_1} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} \sigma^2_e & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & \sigma^2_e \end{bmatrix}$$

$\mathbf{V}_i$ matrix: $\text{Variance}[y_{\text{time}}]$

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

$\mathbf{V}_i$ matrix: $\text{Covariance}[y_A, y_B]$

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

$\mathbf{V}_i$ matrix is time-specific!

$\mathbf{Z}_i$ = n x u values of **predictors with random effects**, so can differ per person (u = 2: int., time slope)

$\mathbf{Z}_i^T$ = u x n values of predictors with random effects (just $\mathbf{Z}_i$ transposed)

$\mathbf{G}_i$ = u x u estimated **random effects variances and covariances**, so will be the same for all persons ($\tau^2_{U_0}$ = int. var., $\tau^2_{U_1}$ = slope var.)

$\mathbf{R}_i$ = n x n **time-specific residual variances and covariances**, so will be same for all persons (here, just diagonal σ^2_e)

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

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

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

$$\mathbf{V} = \begin{bmatrix} 1 & 0.0 & 0 & 0 \\ 1 & 1.0 & 0 & 0 \\ 1 & 2.0 & 0 & 0 \\ 1 & 3.0 & 0 & 0 \\ 0 & 0 & 1 & 0.2 \\ 0 & 0 & 1 & 1.4 \\ 0 & 0 & 1 & 3.5 \end{bmatrix} \begin{bmatrix} \tau^2_{U_0} & \tau_{U_{01}} & 0 & 0 \\ \tau_{U_{01}} & \tau^2_{U_1} & 0 & 0 \\ 0 & 0 & \tau^2_{U_0} & \tau_{U_{01}} \\ 0 & 0 & \tau_{U_{01}} & \tau^2_{U_1} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0.0 & 1.0 & 2.0 & 3.0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0.2 & 1.4 & 3.5 \end{bmatrix} + \begin{bmatrix} \sigma^2_e & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_e & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_e \end{bmatrix}$$

- 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 correlation or differential residual variance across time (in lme, but not lmer in R)

Strategies for Longitudinal MLMs: Part 1

- Estimate empty models for all lower-level variables; compute ICC1
  - Determine the **levels of variance** each variable contains → guide model inclusion

- Are **effects of** "**time**" relevant: for change, yes; for fluctuation, maybe?
  - For accelerated designs (e.g., people start at different ages), create WPtime as time in study (WPtime = current age − starting age) to avoid smushed change effects!
    - For more, see Hoffman (2015) chapter 10
  - Estimate a model treating "**time**" as (rounded) a **categorical** predictor to see the mean trend
  - Evaluate **fixed and random effects** of "time" in **G** matrix as warranted by the mean trend and the design of the study (i.e., consider continuous vs discontinuous; linear vs nonlinear)
    - For lots of examples, see Hoffman (2015) chapter 6
  - Evaluate need for **residual correlation** in **R** matrix (must use lme, not lmer, for this)
    - For more, see Hoffman (2015) chapter 4–5
  - **Do all of this for each level-1 variable, regardless of outcome or predictor!**

Strategies for Longitudinal MLMs: Part 2

- Examine L2 time-invariant and/or other L1 time-varying predictors
  - L2 predictors can have fixed effects moderate the change patterns (as main effect + interactions with the time predictors), explaining random effect variances
  - L1 predictors can have fixed effects and random effects, but don't smush them!
    - Person-mean-centering (up next) into L1 within + L2 between is most common!
    - Contextual effects don't make as much sense (an occasion is not an entity like a person is)

- 4 Examples to Follow… stay tuned
  - 1. Time-Varying Stress Predicting Time-Varying Disease Severity
  - 2: Random Quadratic Change Moderated by Person Reasoning
  - 3: Three-Occasion Piecewise Change Moderated by Treatment Group
  - 4. Four-Phase Multivariate Change Moderated by Treatment Group

Formative Assessment #5 Questions

1. In two-level models for longitudinal data, what are the "levels" of variance at play in any variable measured repeatedly, and what kind of information do they each contain?

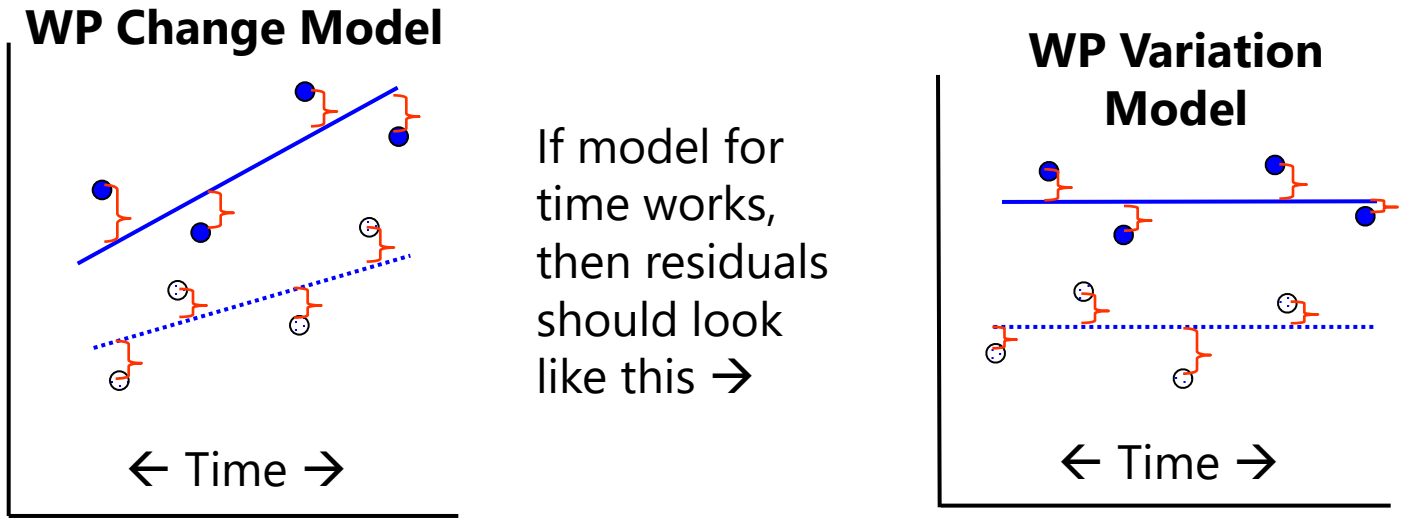
2. Contrast designs for measuring within-person change with within-person fluctuation with respect to their purpose and the role of "time in study" in each. Give an example of a research question that would be relevant for each.

3. What is the difference between fixed and random slopes for time (or for any time-varying predictor more generally)?

4. You submit a paper analyzing three-occasion data for 20 participants using multilevel model with REML estimation in which time is treated as a factor variable (i.e., categorical predictor), and the only random effect is the intercept. Reviewer 2 insists that your sample size is too small for a procedure as complex as multilevel modeling, and that you should have used repeated measures ANOVA instead. How do you respond?

The Joy of Time-Varying (TV) Predictors

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



- TV predictors are complicated because they represent an **aggregated effect**:
  - Effect of the *between-person* variation in the predictor x_{ti} on y_{ti}
  - Effect of the *within-person* variation in the predictor x_{ti} on y_{ti}
  - For now, we are assuming the predictor x_{ti} only **fluctuates** over time…
    - *We will need a **different model** when x_{ti} changes individually over time!*

The Joy of Time-Varying Predictors

- 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 **ICC1**
  - ICC1 = (BP variance) / (BP variance + WP variance)
  - **ICC1 > 0?** TV predictor has **BP** variation (so it *could* have a BP effect)
  - **ICC1 < 1?** TV predictor has **WP** variation (so it *could* have a WP effect)

Between-Person vs. Within-Person Slopes

- Between-person and within-person slopes 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 slopes 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!

Person-Mean-Centering (P-MC)

- Person-MC splits the TV predictor x_{ti} into 2 variables to directly represent its BP (level-2) and WP (level-1) sources of variation and **include these variables as the predictors**:

- **Level-2, PM predictor = person mean of x_{ti}**
  - $\boldsymbol{PMx_i = \bar{x}_i - C_2}$
  - PMx_i is centered at constant C_2, chosen for meaningful 0 (e.g., sample mean)
  - 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}**
  - $\boldsymbol{WPx_{ti} = x_{ti} - \bar{x}_i}$ *(note: uncentered person mean $\bar{x}_i$ is used to center x_{ti})*
  - WPx_{ti} is NOT centered at a constant – it 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 slopes directly through **separate** parameters

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

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

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

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

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

$PMx_i = \bar{x}_i - C_2$ → it has only Level-2 BP variation

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

γ_{01} = BP main effect of having *more $\bar{x}_i$ than other people*

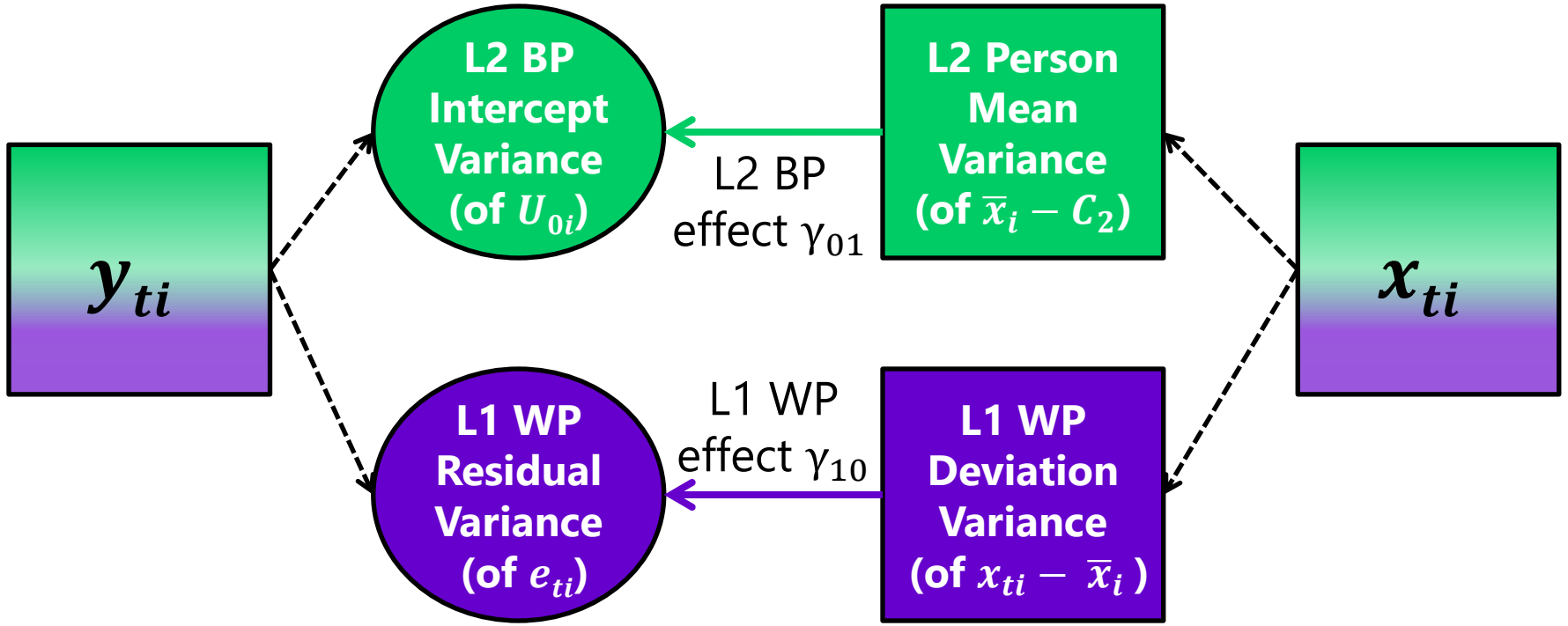
Random slopes and cross-level interactions can be added to predict β_{1i} to allow BP differences in the WPx_{ti} slope (and moderation thereof)

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

Univariate MLM: Variable-Centering

Model-based partitioning of level-1 y_{ti} outcome variance into estimated **variance components**:

Brute-force partitioning of level-1 x_{ti} predictor variance into **observed variables**:

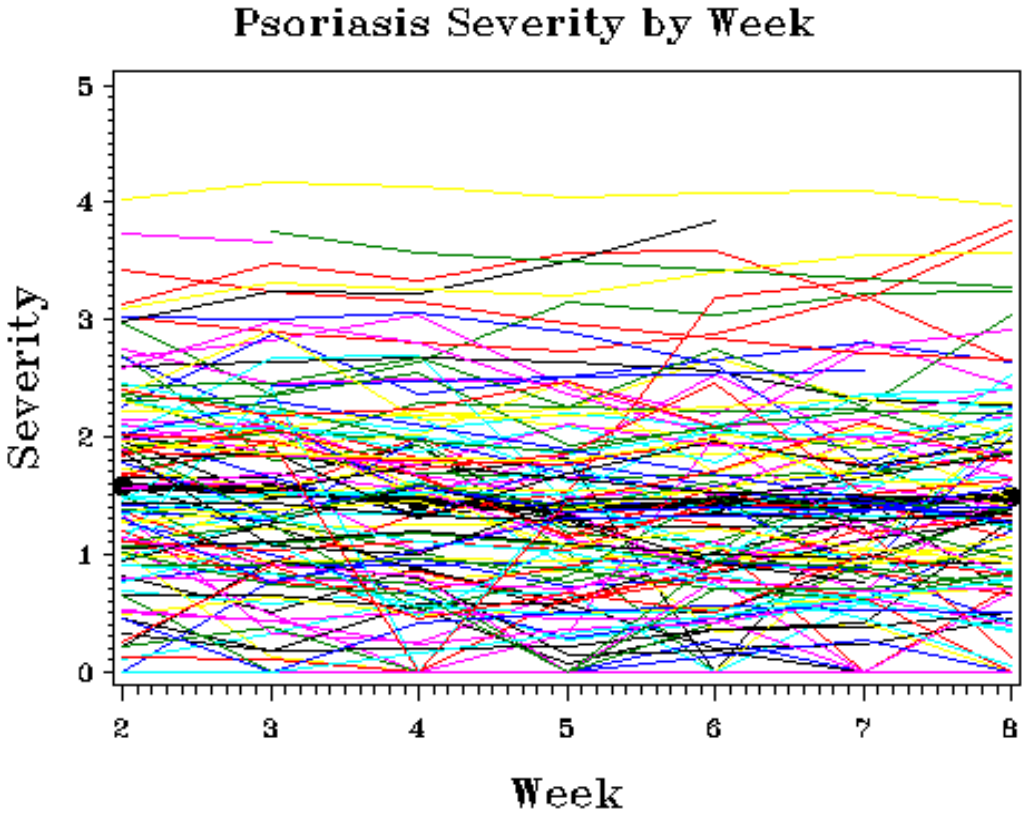


Why not let the model estimate variance components for x_{ti}, too?
This is the basis of multivariate MLM (or "multilevel SEM" = M-SEM).

Example 1: Weekly Stress and Psoriasis

- 127 psoriasis patients, 8 weekly assessments (only last 7 used)
- How does perceived stress predict psoriasis severity? At what time lag?
- No change in treatment → only fluctuation over time

- Analysis plan:
  - ICCs for stress and severity—how much variance is at each level?
  - Assess pattern of variance and covariance in severity over time
  - Evaluate prediction of severity by stress at lag 0 and lag 1 weeks… without smushing!



Example 1: Weekly Stress and Psoriasis

- Empty means, random intercept model to get ICCs → proportion of total variance due to BP mean differences
  - For each variable: $y_{ti} = \gamma_{00} + U_{0i} + e_{ti}$, $\mathbf{ICC_1} = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{\mathbf{BP}}{\mathbf{BP+WP}}$
  - Severity outcome: ICC1 = .83; stress predictor: ICC1 = .56

- For the severity outcome, the best-fitting unconditional time model for the variance had a level-2 random intercept (in the G matrix), along with heterogeneous level-1 residual variances and a Toeplitz (banded) correlation structure up to lag 3 (in the R matrix, below)

Estimated R Correlation Matrix for ID 1 → WP residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5115	0.3566	0.1112			
2	0.5115	1.0000	0.5115	0.3566	0.1112		
3	0.3566	0.5115	1.0000	0.5115	0.3566	0.1112	
4	0.1112	0.3566	0.5115	1.0000	0.5115	0.3566	0.1112
5		0.1112	0.3566	0.5115	1.0000	0.5115	0.3566
6			0.1112	0.3566	0.5115	1.0000	0.5115
7				0.1112	0.3566	0.5115	1.0000

Example 1: Weekly Stress and Psoriasis

Level 1: $severity_{ti} = \beta_{0i} + \beta_{1i}(WPstressL0_{ti}) + \beta_{2i}(WPstressL1_{ti}) + e_{ti}$

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

$\beta_{1i} = \gamma_{10}$ $\beta_{2i} = \gamma_{20}$

$WPx_{ti} = x_{ti} - \bar{x}_i$ → **it has only Level-1 WP variation**

$PMx_i = \bar{x}_i - C_2$ → **it has only Level-2 BP variation**

WP effects are <u>fixed</u> only (no random slopes) → same for everyone

Model for the Means:

γ_{00} → expected severity for someone with person mean stress = 2, and who had severity = 2 last week and currently

γ_{01} → BP difference in *average* severity per unit person mean stress

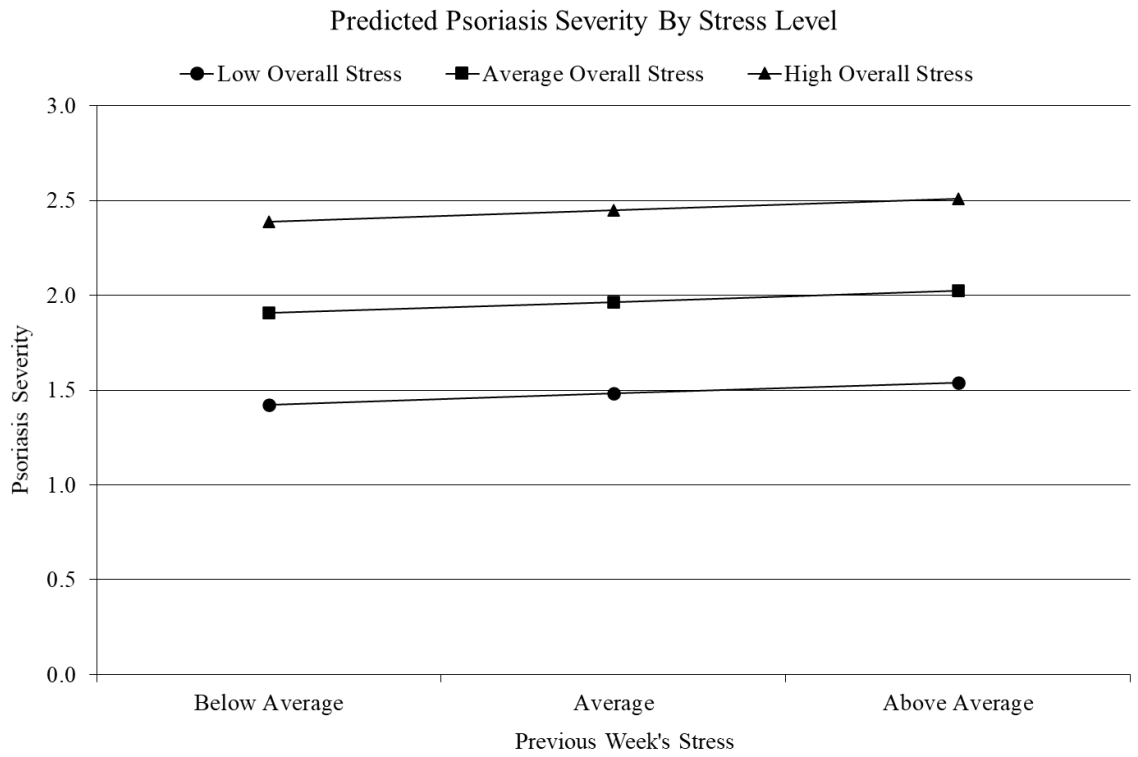
γ_{10} and γ_{20} → WP change in *current* severity per unit more stress than usual this week (lag 0) and last week (lag 1)

Example 1: Weekly Stress and Psoriasis

Level 1: $severity_{ti} = \beta_{0i} + \beta_{1i}(WPstressL0_{ti}) + \beta_{2i}(WPstressL1_{ti}) + e_{ti}$

Level 2: $\beta_{0i} = 1.96 + 0.48^{*}(PMstress_i - 2) + U_{0i}$

$\beta_{1i} = 0.02 \qquad \beta_{2i} = 0.06^{*}$



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

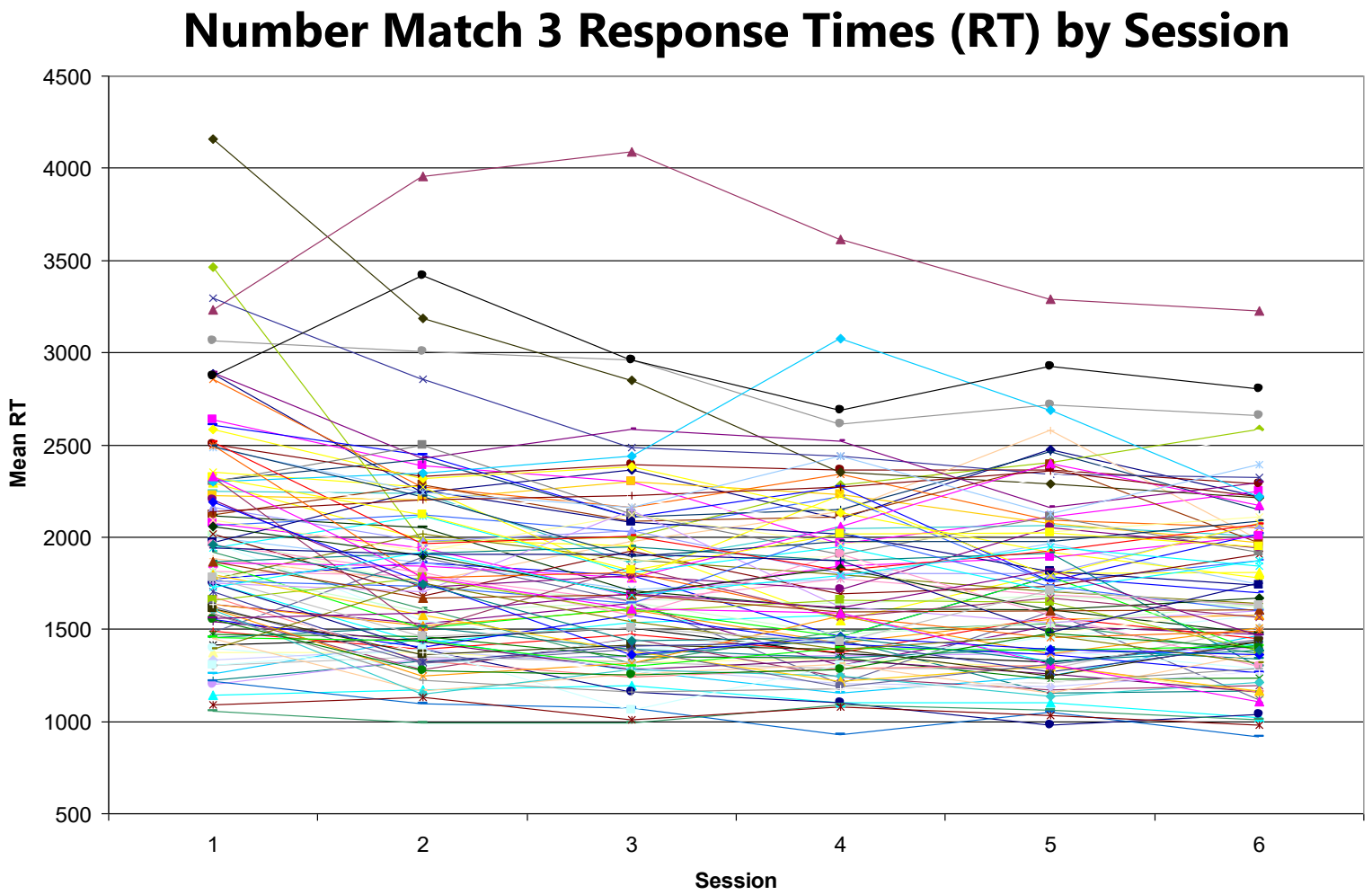
$PMx_i = \bar{x}_i - C_2$ → it has only Level-2 BP variation

WP effects are <u>fixed</u> only (no random slopes) → same for everyone

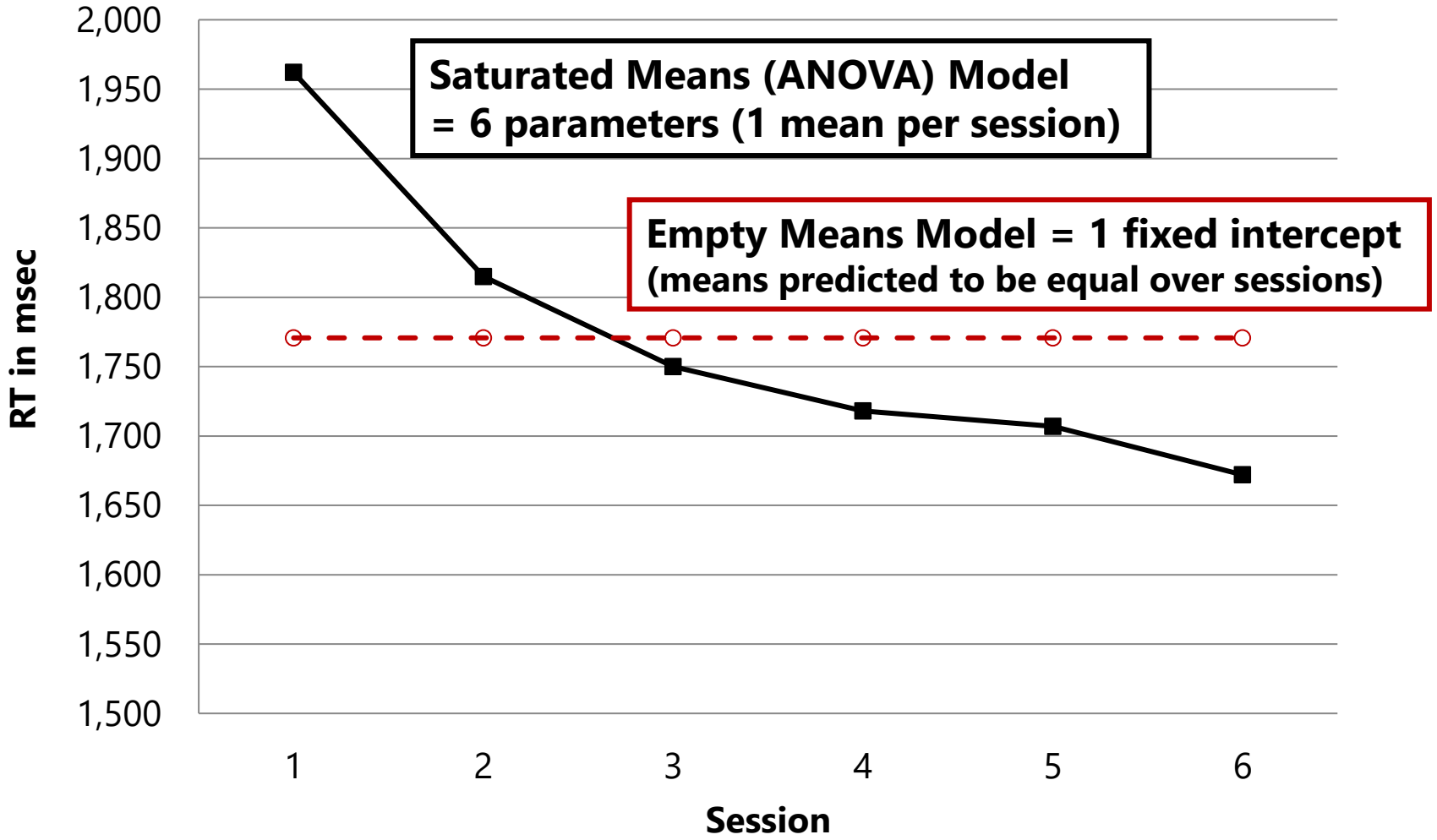
Families of Nonlinear Change

- Polynomial functions (e.g., time^2, time^3) → see details on next slides and Example 2
  - Best suited for time slopes that should change directions (in which time is treated as continuous)
- Piecewise (linear spline) functions (see examples 3 and 4)
  - Best suited for distinct phases of time (known "knot" points)
  - Otherwise, location of "latent" knots can be model parameters
- Linear effect of log(time) → exponential-ish
  - Good for time slopes that should level off (hit upper or lower asymptote)
  - Adding quadratic log(time) adjusts how fast the time slope levels off
- Latent basis → single slope with estimated nonlinearity
  - In SEM software, for random time slope factor: fix first loading to 0, last loading to 1, and estimate the other loadings to capture proportion of change by each occasion
- Truly nonlinear models (e.g., logistic, exponential, time-to-onset/offset)
  - Harder to estimate, particularly for random effects variances

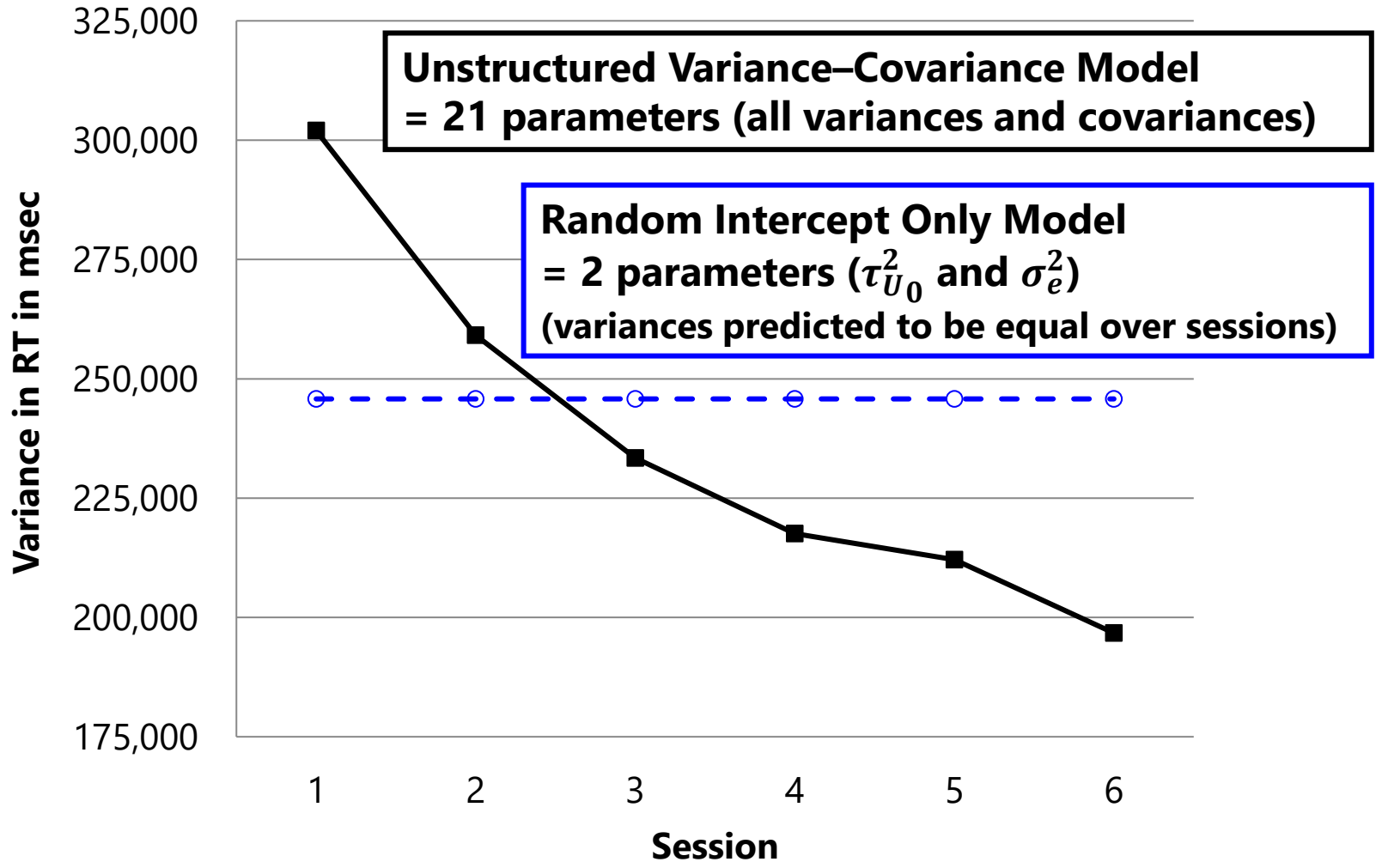
Example 2: Individual Trajectories
101 older adults, 6 occasions within 2 weeks



Example 2 Mean RT by Session:
Baseline Models for the Means



Example 2 Variance in RT by Session:
Baseline Models for the Variance



Interpreting Quadratic Fixed Effects

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

- Fixed quadratic slope = "**half the rate of acceleration/deceleration**"
- To interpret the quadratic slope as how the linear time slope changes per unit time, **you must multiply the quadratic slope 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, at time 2 = 5.2

$$\text{Intercept (Position) at Time T: } y_T = 50.0 + 4.0T + 0.3T^2$$

$$\text{First Derivative (Velocity) at Time T: } \frac{dy_T}{d(T)} = 4.0 + 0.6T$$

$$\text{Second Derivative (Acceleration) at Time T: } \frac{d^2 y_T}{d(T)} = 0.6$$

$$y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ$$

$$\text{Effect of } X = \beta_1 + \beta_3 Z$$

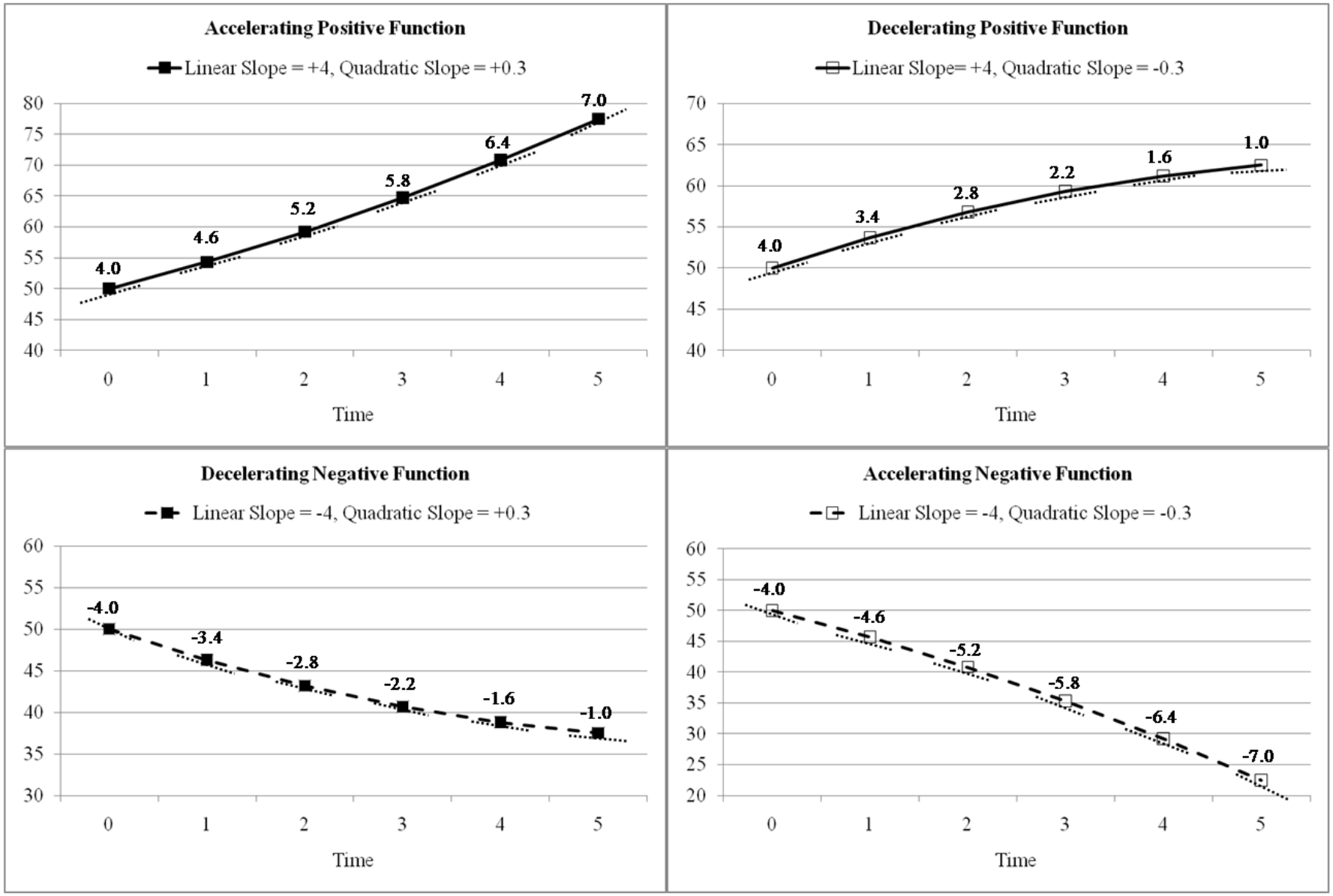
$$\text{Effect of } Z = \beta_2 + \beta_3 X$$

$$y_T = \beta_0 + \beta_1 \text{Time}_T + ____ + \beta_3 \text{Time}_T^2$$

$$\text{Effect of Time}_T = \beta_1 + 2\beta_3 \text{Time}_T$$

- **Why?** Left → because derivatives; Right → because time^2 is an interaction!
- Because time is interacting with itself, there is no second "main effect" in the model for the interaction to modify. So the quadratic time slope gets applied twice when added to the one (main) linear time slope

Examples of Fixed Quadratic Time Trends



Random Quadratic Time Unconditional Model

Level 1: $\boldsymbol{RT_{ti} = \beta_{0i} + \beta_{1i}(Time_{ti}) + \beta_{2i}(Time_{ti}^2) + e_{ti}}$

Level 2 Equations (one per β):

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

- $\boldsymbol{\beta_{0i}}$: **Intercept** for person *i*
- $\boldsymbol{\gamma_{00}}$: Fixed (mean) **Intercept**
- $\boldsymbol{U_{0i}}$: Random (Deviation) **Intercept**

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

- $\boldsymbol{\beta_{1i}}$: **Linear Time Slope** for person *i*
- $\boldsymbol{\gamma_{10}}$: Fixed (mean) **Linear Slope**
- $\boldsymbol{U_{1i}}$: Random (Deviation) **Linear Slope**

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

- $\boldsymbol{\beta_{2i}}$: **Quadratic Time Slope** for person *i*
- $\boldsymbol{\gamma_{20}}$: Fixed (mean) **Quad Slope**
- $\boldsymbol{U_{2i}}$: Random (Deviation) **Quad Slope**

Time = session – 1
random effect covariances estimated

Fixed Effect Subscripts:
1st = which level-1 term
2nd = which level-2 term

# of Possible Time-Related Slopes by # of Occasions (*n*):

Fixed time slopes = $n - 1$
Random time slopes = $n - 2$

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

Example 2: Adding Reasoning (0=22) as a Time-Invariant Predictor Is RT Improvement Predicted by Fluid Intelligence?

Level 1: $RT_{ti} = \beta_{0i} + \beta_{1i}(Time_{ti}) + \beta_{2i}(Time_{ti}^2) + e_{ti}$

Level 2 Equations (one per β):

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

- β_{0i}: **Intercept** for person *i*
- γ_{00}: Fixed Intercept when Time=0 and Reas=22
- γ_{01}: Δ in Intercept per unit Δ in Reas
- U_{0i}: Random (Deviation) Intercept after controlling for Reas

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

- β_{1i}: **Linear Slope** for person *i*
- γ_{10}: Fixed Linear Time Slope when Time=0 and Reas=22
- γ_{11}: Δ in Linear Time Slope per unit Δ in Reas (=Reas*time)
- U_{1i}: Random (Deviation) Linear Time Slope after controlling for Reas

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

- β_{2i}: **Quad Slope** for person *i*
- γ_{20}: Fixed Quad Time Slope when Reas=22
- γ_{21}: Δ in Quad Time Slope per unit Δ in Reas (=Reas*time^2)
- U_{2i}: Random (Deviation) Quad Time Slope after controlling for Reas

Example 2: Reasoning (0=22) as a Time-Invariant Predictor Is RT Improvement Predicted by Fluid Intelligence?

Level 1: $\boldsymbol{RT_{ti} = \beta_{0i} + \beta_{1i}(Time_{ti}) + \beta_{2i}(Time_{ti}^2) + e_{ti}}$

Level 2 Equations (one per β):

$$\boldsymbol{\beta_{0i} = \gamma_{00} + \gamma_{01} Reas_i + U_{0i}}$$
$$\boldsymbol{\beta_{1i} = \gamma_{10} + \gamma_{11} Reas_i + U_{1i}}$$
$$\boldsymbol{\beta_{2i} = \gamma_{20} + \gamma_{21} Reas_i + U_{2i}}$$

$\boldsymbol{\gamma_{11}}$ and $\boldsymbol{\gamma_{21}}$ are known as "**cross-level**" interactions (level-1 predictor by level-2 predictor)

Each fixed slope of reasoning will predict the random $\boldsymbol{U_i}$ variance in its level-2 equation if present, or $\boldsymbol{e_{ti}}$ residual variance otherwise. That's why random slopes should be tested **before** adding cross-level interactions!

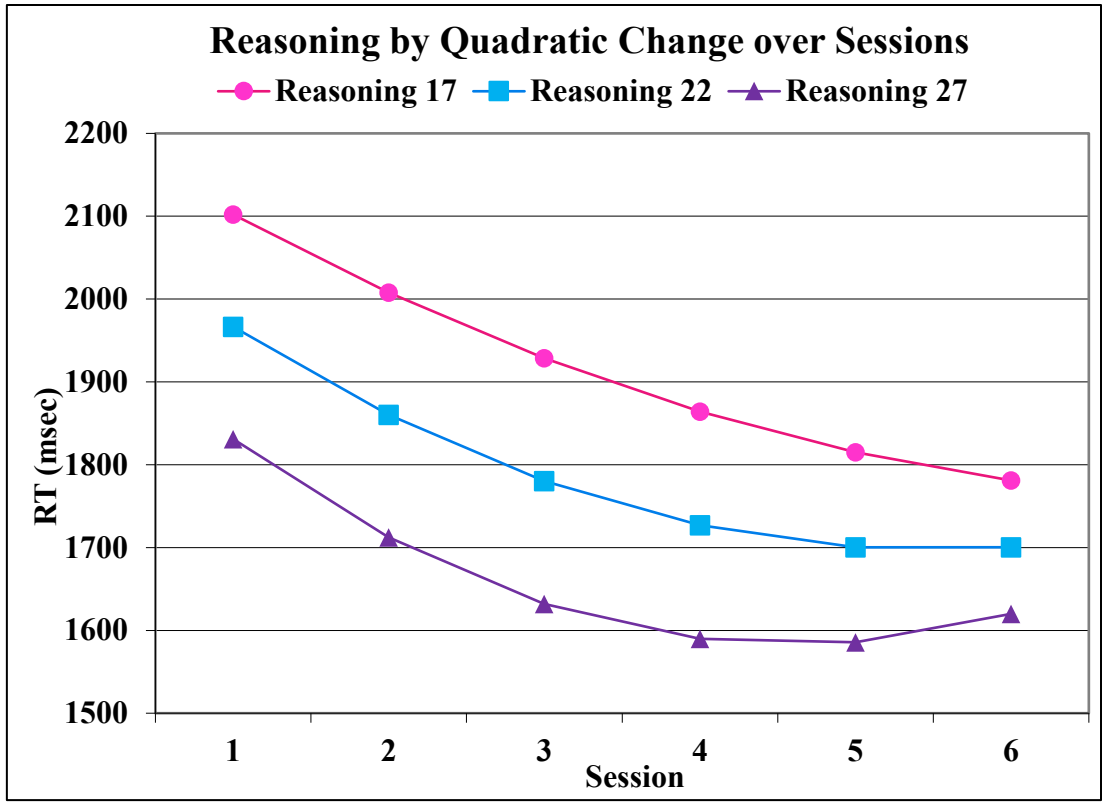
Composite equation:

$$\begin{aligned} \boldsymbol{y_{ti}} = {} & \boldsymbol{(\gamma_{00} + \gamma_{01} Reas_i + U_{0i}) +} \\ & \boldsymbol{(\gamma_{10} + \gamma_{11} Reas_i + U_{1i}) Time_{ti} +} \\ & \boldsymbol{(\gamma_{20} + \gamma_{21} Reas_i + U_{2i}) Time_{ti}^2 + e_{ti}} \end{aligned}$$

Example 2: Reasoning (0=22) as a Time-Invariant Predictor
Is RT Improvement Predicted by Fluid Intelligence?

$$RT_{ti} = (1966 + -27 * Reas_i + U_{0i}) + (-120 + -3.6 * Reas_i + U_{1i})Time_{ti} + (13 + 1.2 * Reas_i + U_{2i})Time_{ti}^2 + e_{ti}$$

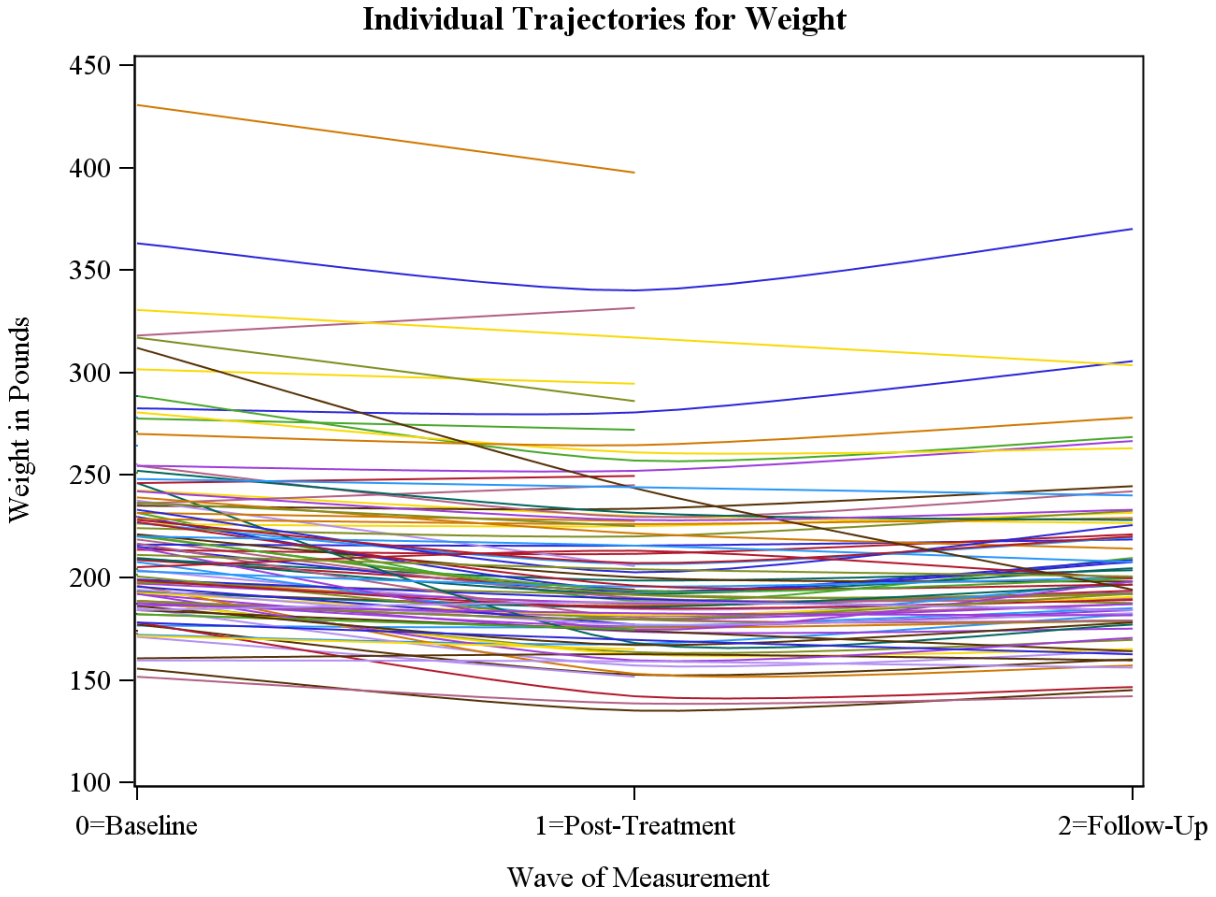
BP Pseudo-R^2 Values:
Intercept U_{0i} = .049
Linear Time U_{1i} = −.006
Quadratic Time U_{2i} = .024
WP Residual e_{ti} **= 0**



People with better reasoning:
- started out faster/lower *(intercept at session 1)*,
- improved more initially *(linear slope at session 1)*,
- and had a greater rate of deceleration with practice *(quadratic slope*2!)*

Example 3: Differences in Weight Loss

- Randomized Control Trial for differences in weight loss (and then maintenance of weight loss) between control and treatment groups
  - 105 persons measured up to 3 occasions → balanced but incomplete data



Example 3: Piecewise Model for the Means

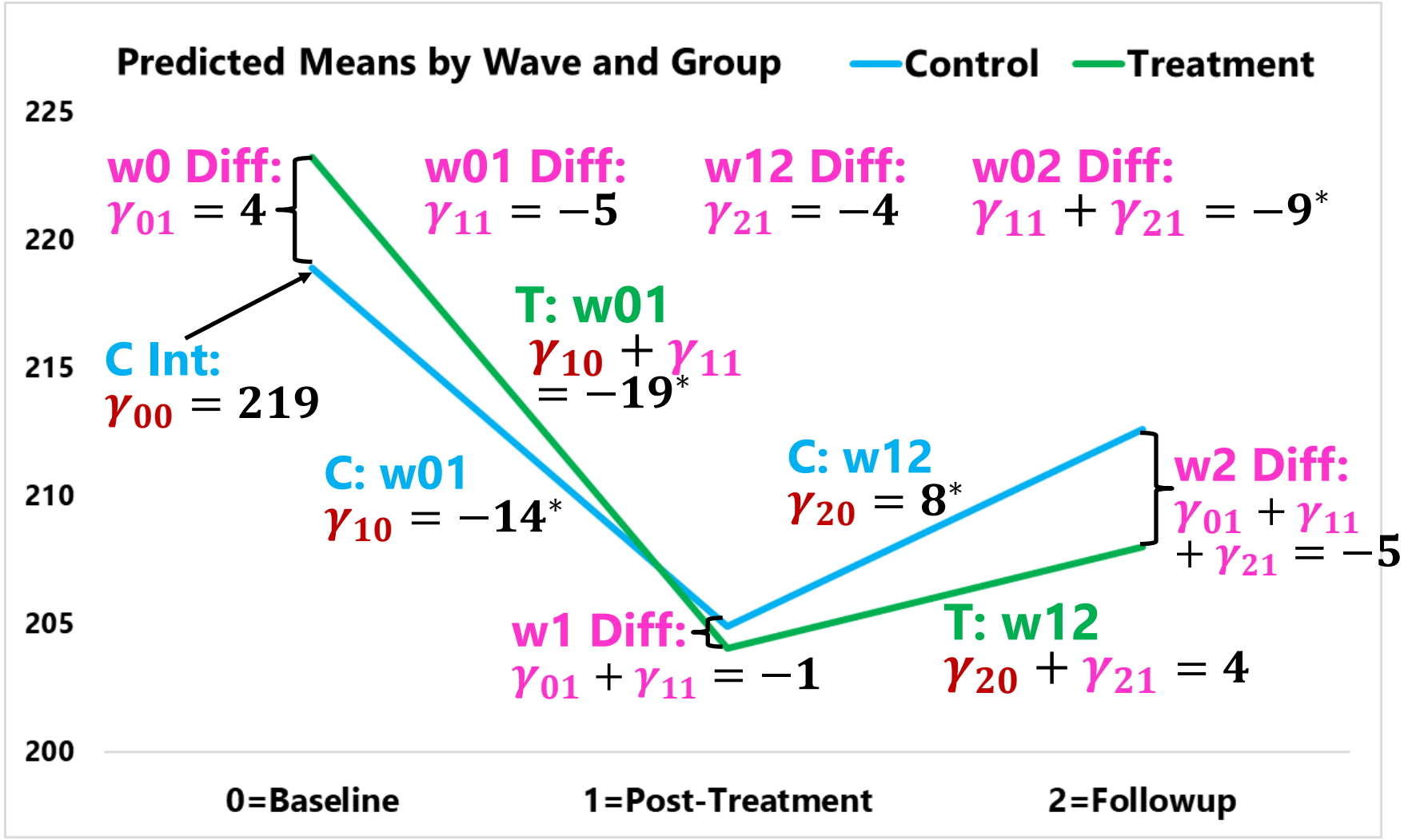
$$\widehat{weight}_{ti} = \boldsymbol{\gamma_{00}} + \boldsymbol{\gamma_{10}}(w01_{ti}) + \boldsymbol{\gamma_{20}}(w12_{ti}) + \boldsymbol{\gamma_{01}}(Tx_i) + \boldsymbol{\gamma_{11}}(w01_{ti})(Tx_i) + \boldsymbol{\gamma_{21}}(w12_{ti})(Tx_i)$$

Wave	w01	w12
0. Baseline	0	0
1. Post-Treat	**1**	0
2. Follow-Up	**1**	**1**

- Estimated fixed effects for control group ($Tx_i = 0$):
  - $\boldsymbol{\gamma_{00}}$ for intercept: Control weight at baseline
  - $\boldsymbol{\gamma_{10}}$ for w01 slope: Control change from baseline to post-treatment
  - $\boldsymbol{\gamma_{20}}$ for w12 slope: Control change from post-treatment to follow-up
- Estimated fixed effects for differences by treatment ($Tx_i = 0\ vs\ 1$):
  - $\boldsymbol{\gamma_{01}}$ for Tx slope: difference in weight at baseline
  - $\boldsymbol{\gamma_{11}}$ for w01*Tx slope: difference in change from baseline to post-treatment
  - $\boldsymbol{\gamma_{21}}$ for w12*Tx slope: differences in change from post-treatment to follow-up
- Model-implied fixed effects for treatment group ($Tx_i = 1$):
  - $\boldsymbol{\gamma_{00}} + \boldsymbol{\gamma_{01}}$: Tx weight at baseline
  - $\boldsymbol{\gamma_{10}} + \boldsymbol{\gamma_{11}}$: Tx change from baseline to post-treatment
  - $\boldsymbol{\gamma_{20}} + \boldsymbol{\gamma_{21}}$: Tx change from post-treatment to follow-up

Example 3: Piecewise Model for the Means

$$\widehat{weight}_{ti} = \gamma_{00} + \gamma_{10}(w01_{ti}) + \gamma_{20}(w12_{ti}) + \gamma_{01}(Tx_i) + \gamma_{11}(w01_{ti})(Tx_i) + \gamma_{21}(w12_{ti})(Tx_i)$$

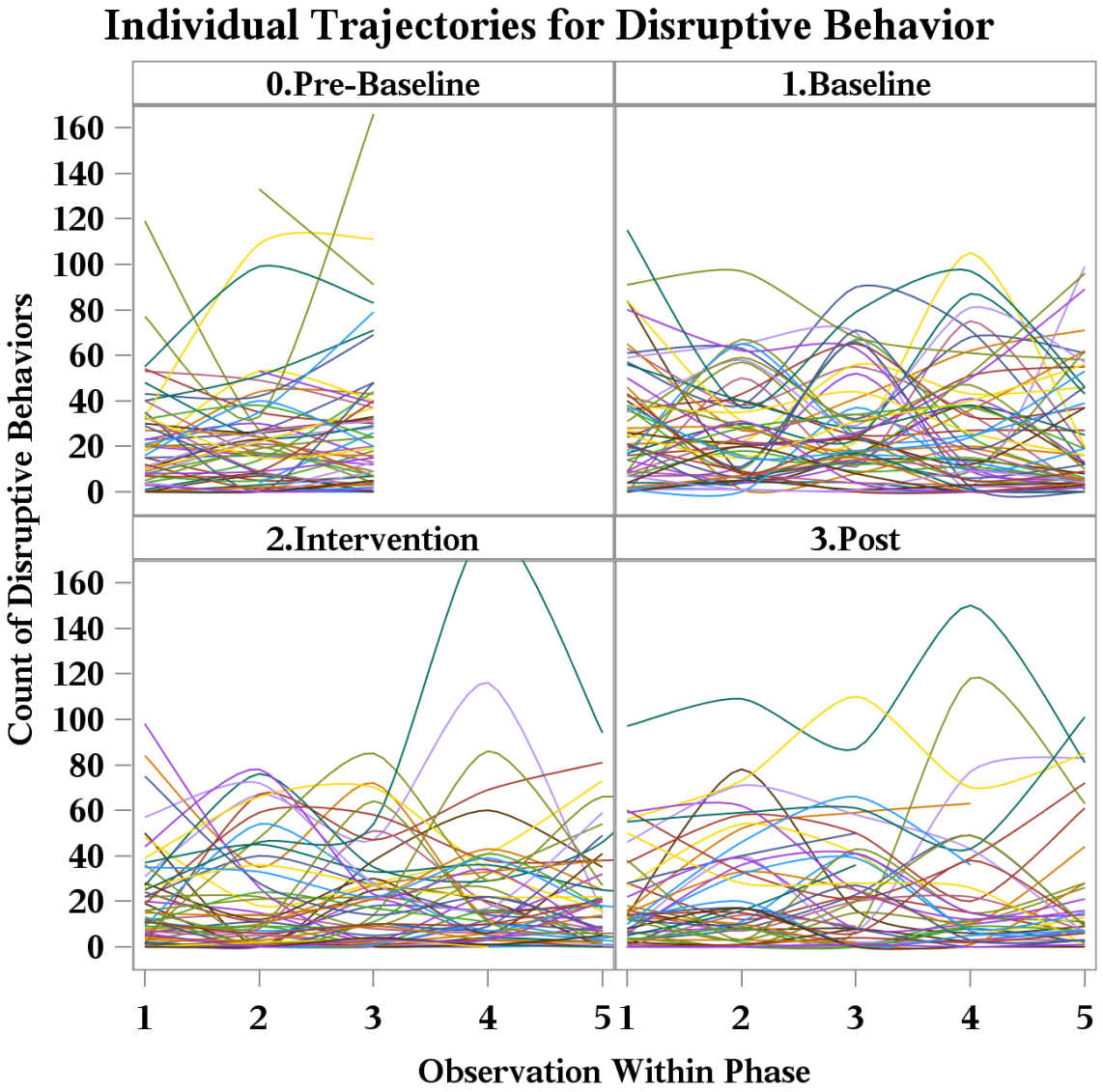


Example 3: Model for the Variance

- Random effects were NOT used to model dependency
  - Why? Given 2 slopes across 3 occasions, only 1 of the slopes could have a random effect → assumes parallel change for other time period!
  - The only way to get 2 random slopes from 3 occasions is to remove the WP residual variance, which assumes BP differences in change are measured perfectly!

- Instead, an unstructured variance–covariance matrix was estimated across occasions, separately for each group (option 2, ACS multivariate)
  - Unstructured = different variances (and covariances) by occasion
  - Still implies differences in individual change (they just can't be quantified separately from individual intercept differences)
  - Separate matrices by group allows treatment differences in individual *heterogeneity* of mean levels and change

Example 4: Multiple Condition RCT



- RCT for effectiveness of MoBeGo self-monitoring app for improving kids' challenging behavior
  - 57 kids measured multiple times under 4 conditions → unbalanced time and incomplete data
  - Multivariate (by condition) MLM on log-transformed count in stacked data
  - Avoided generalized model in order to use REML estimation and KR DDF
  - Modeled nonconstant variance by condition

Example 4: Multivariate MLM

Level 1: $DB_{tic} = \beta_{0i1}(Pre_c) + \beta_{0i2}(Base_c) + \beta_{0i3}(Inter_c) + \beta_{0i4}(Post_c) + e_{tic}$

Level 2:

- **Pre-Baseline:** $\beta_{0i1} = \gamma_{001} + \gamma_{011}(Tx_i) + U_{0i1}$
- **Baseline:** $\beta_{0i2} = \gamma_{002} + \gamma_{012}(Tx_i) + U_{0i2}$
- **Intervention:** $\beta_{0i3} = \gamma_{003} + \gamma_{013}(Tx_i) + U_{0i3}$
- **Post-Intervention:** $\beta_{0i4} = \gamma_{004} + \gamma_{014}(Tx_i) + U_{0i4}$

- **Model for the Means:**
  - Pre, Base, Inter, and Post are dummy-coded condition indicators (0=no, 1=yes)
  - Fixed effects for control group ($Tx_i = 0$): $\gamma_{001} \ldots \gamma_{004}$
  - Differences by treatment ($Tx_i = 0$ *vs* 1): $\gamma_{011} \ldots \gamma_{014}$
  - Model-implied fixed effects for treatment group ($Tx_i = 1$): $(\gamma_{001} + \gamma_{011}) \ldots (\gamma_{004} + \gamma_{014})$

- **Model for the Variance—Multivariate* Two-Level Model:**
  - Unstructured level-2 G matrix for 4 random intercepts → separate variances and covariances by condition (*which makes it a multivariate MLM, reviewer 2)
  - Heterogeneity of level-1 residual variance examined by condition
  - Could have also examined individual differences *in level-1 variance* in "scale" model!

Example 4: Multivariate MLM

Level 1: $DB_{tic} = \beta_{0i1}(Pre_c) + \beta_{0i2}(Base_c) + \beta_{0i3}(Inter_c) + \beta_{0i4}(Post_c) + e_{tic}$

Level 2:

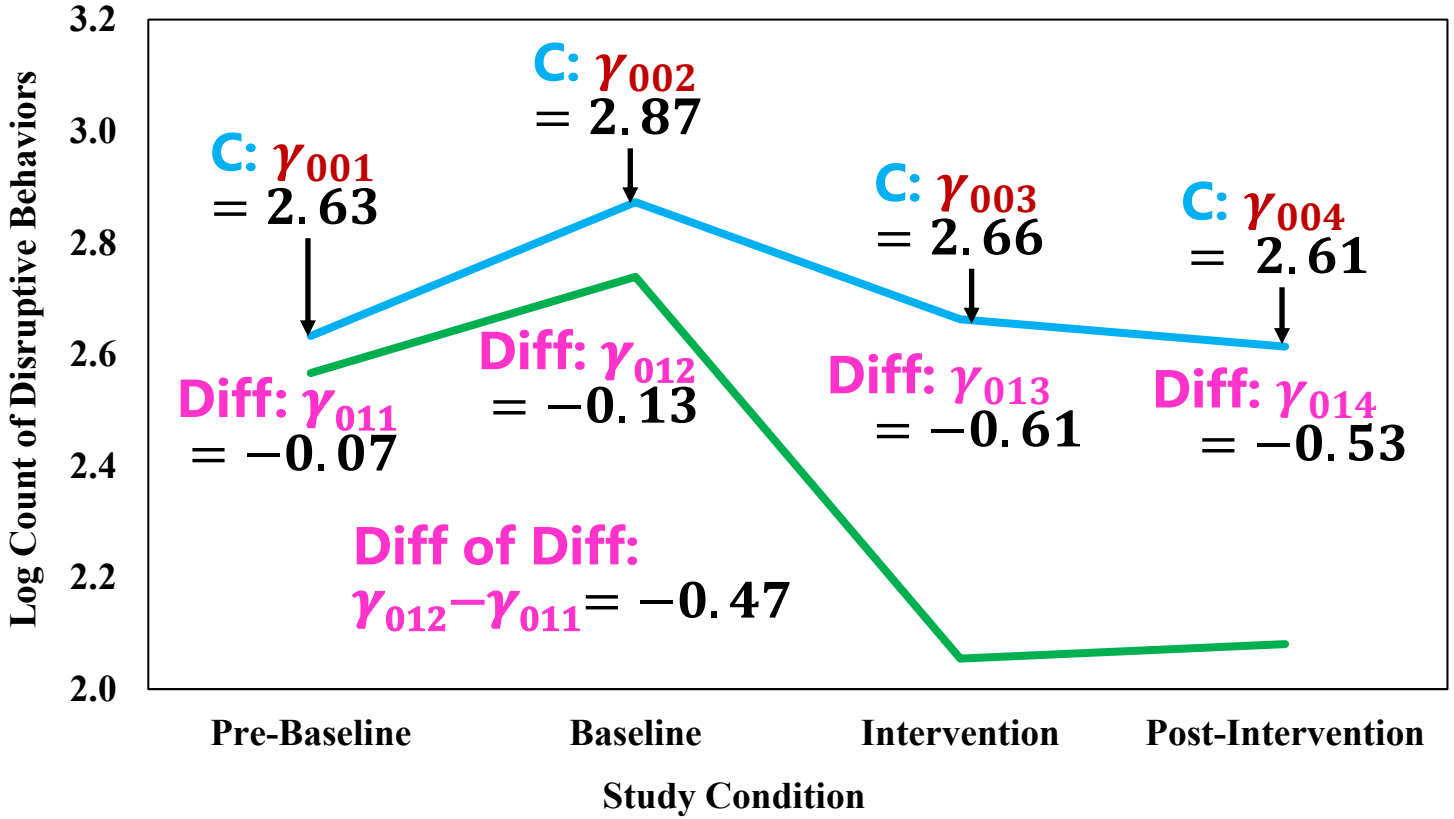
Pre-Baseline: $\beta_{0i1} = \gamma_{001} + \gamma_{011}(Tx_i) + U_{0i1}$

Baseline: $\beta_{0i2} = \gamma_{002} + \gamma_{012}(Tx_i) + U_{0i2}$

Intervention: $\beta_{0i3} = \gamma_{003} + \gamma_{013}(Tx_i) + U_{0i3}$

Post-Intervention: $\beta_{0i4} = \gamma_{004} + \gamma_{014}(Tx_i) + U_{0i4}$

All other differences and interactions (differences of differences) are linear combinations of model fixed effects…



R Code for Each Example

Example 1: Time-Varying Stress Predicts Time-Varying Severity (using lme in nlme, AR3 in R instead)

```
Model1 = lme(data=Example1, method="REML", fixed=severity~1+PMstress2+WPstressL0+WPstressL1,
             random=~1|ID,                                          # level-2 random intercept (G)
             correlation=corARMA(form=~wave|ID, p=3, q=0),    # AR(3) residual correlation
             weights=varIdent(form=~1|factor(wave)),    # heterogeneous residual variances by wave
```

Example 2: Random Quadratic Model using lmer in lme4 (+ lmerTest)

```
Model2 = lmer(data=Example2, REML=TRUE,
              formula=RT~1+time+timesq+reas
              +time:reas +timesq:reas +(1+time+timesq|ID))
```

Example 3: Three-Occasion Piecewise Model with Separate Unstructured Variances and Covariances by Group (using mmrm package)

```
Model3 = mmrm(data=Example3, REML=TRUE,
              method="Kenward-Roger" # KR denominator DF and adjusted SEs
              formula=weight ~ 1 + w01 + w12 + Tx + w01:Tx + w12:Tx +
                         us(wave | Tx / ID)) # separate UN R by Tx group
```

Example 4: Multivariate MLM for 4 Phases using lmer in lme4 (+ lmerTest)

```
Model4 = lmer(data=Example4, REML=TRUE, formula=logDB~0+Pre+Base+Inter+Post
              +Pre:Tx+Base:Tx+Inter:Tx+Post:Tx+(0+Pre+Base+Inter+Post|ID)) # 0 removes intercepts
```

When to Eject to Multivariate MLM

- Multivariate MLM (via SEM or M-SEM): predict multiple time-varying outcomes simultaneously *and predict each other* may be needed when:
  - Nontrivial amounts of missing data for predictors
    - All rows must be complete in univariate MLM (lme/lmer), but not multivariate (joint likelihood)
    - Bad news for person-MC, solvable through "latent centering" (TVx → outcome like TVy)
  - You want "cross-lagged" L1 effects in both directions (WPx → WPy; WPy → WPx)
    - Only WPx → WPy in univariate MLM, and can't use first occasion (lagged WPx is missing)
  - Time-varying predictors contain individual differences in change!
    - Person-MC still yields a separate "time-smushed" effect, as described in this paper
  - You want latent measurement models, mediation, or random effects as predictors

Summary

- Longitudinal MLMs are based on the same concepts as clustered MLMs:
  - Distinguishing and predicting level-specific outcome variance
  - Using random effects to account for the correlation of observations from same unit
  - Separating predictor fixed and random effects across all applicable levels (no smushing allowed, either as within/contextual or within/between models)

- The "time" element of longitudinal data requires extra consideration:
  - For WP change, for time in study; for WP variation, for relevant time (e.g., day of week; time of day): What pattern do the occasion-specific means and variances show? Let that guide you to a family of fixed and random time effects to consider
  - Do you also need residual correlation and/or heterogeneous variance by time?
  - Use person-mean-centering for TV predictors that do not have individual change; otherwise, plan on using a multivariate MLM (in which TV predictors are outcomes)