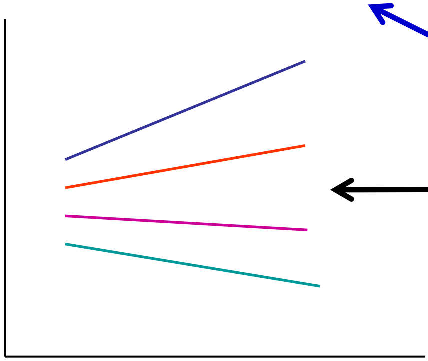
Describing Within-Person Change over Time

- Topics:
  - **Multilevel modeling notation and terminology**
  - **Fixed and random effects of linear time**
  - Predicted variances and covariances from random slopes
  - How random effects model dependency
  - Describing nonlinear change: polynomial models
  - Describing nonlinear change: other alternatives

Modeling Change vs. Fluctuation

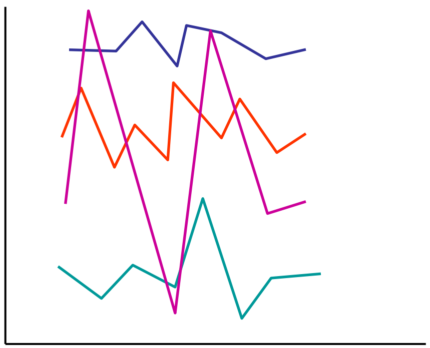
Pure WP Change



Time

Our focus for today using random effects models

Pure WP Fluctuation



Time

Uses alternative covariance structure models instead

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 Variance:

- **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

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*

Parsimony

<u>Empty Model:</u>
Predicts NO change over time

1 Fixed Effect

<u>In-between options:</u>
polynomial slopes,
piecewise slopes,
nonlinear slopes…

<u>Saturated Means:</u>
Reproduces mean at each occasion

 # Fixed Effects = # Occasions

Good fit

Name… that… Trajectory!

The Big Picture of Longitudinal Data: **Models for the Variance**

Compound Symmetry (CS)		Unstructured (UN)
$\begin{bmatrix} \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e \end{bmatrix}$	Most useful model: likely somewhere in between!	$\begin{bmatrix} \sigma^2_{T1} & \sigma_{T12} & \sigma_{T13} & \sigma_{T14} \\ \sigma_{T21} & \sigma^2_{T2} & \sigma_{T23} & \sigma_{T24} \\ \sigma_{T31} & \sigma_{T32} & \sigma^2_{T3} & \sigma_{T43} \\ \sigma_{T41} & \sigma_{T42} & \sigma_{T43} & \sigma^2_{T4} \end{bmatrix}$

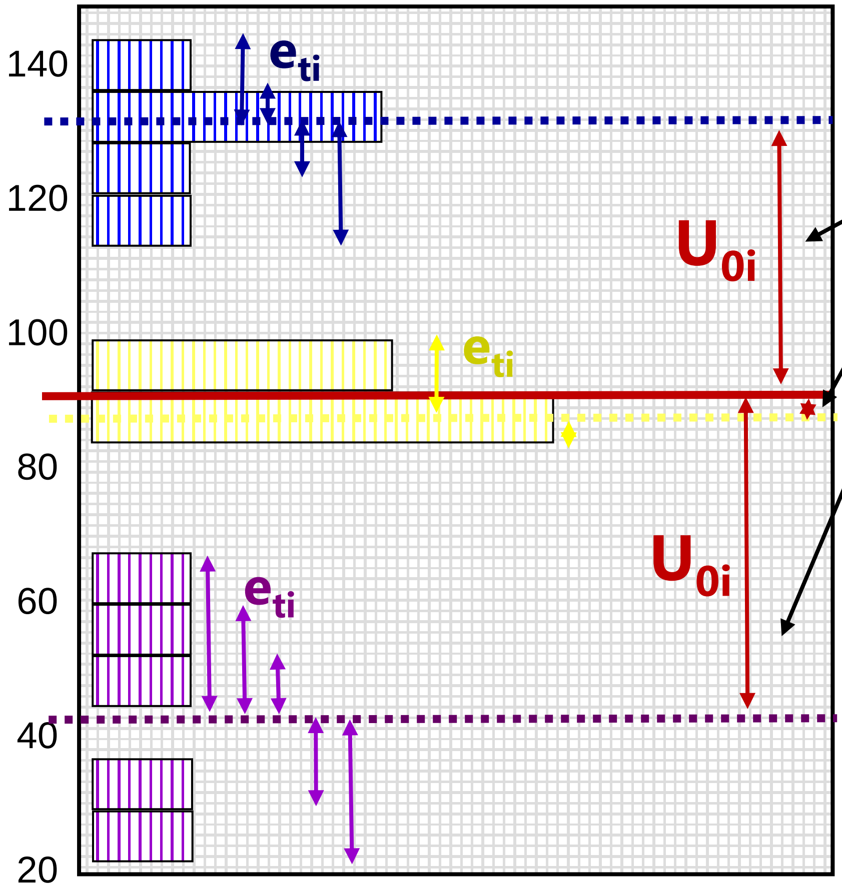
Parsimony ⟵⟶ **Good fit**

Univariate RM ANOVA | ***Name …that … Structure!*** | Multivariate RM ANOVA

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



y_{ti} variance → 2 sources:

Level 2 Random Intercept Variance (of U_{0i}, as $\tau^2_{U_0}$):

→ **Between**-Person Variance

→ Differences from **GRAND** mean

→ **INTER**-Individual Differences

Level 1 Residual Variance (of e_{ti}, as σ^2_e):

→ **Within**-Person Variance

→ Differences from **OWN** mean

→ **INTRA**-Individual Differences

Empty Means, Random Intercept Model

GLM Empty Model:

- $\mathbf{y_i = \beta_0 + e_i}$

MLM Empty Model:

- Level 1:

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

- Level 2:

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

3 Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

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

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

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

Random Intercept = individual-specific deviation from predicted intercept

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

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

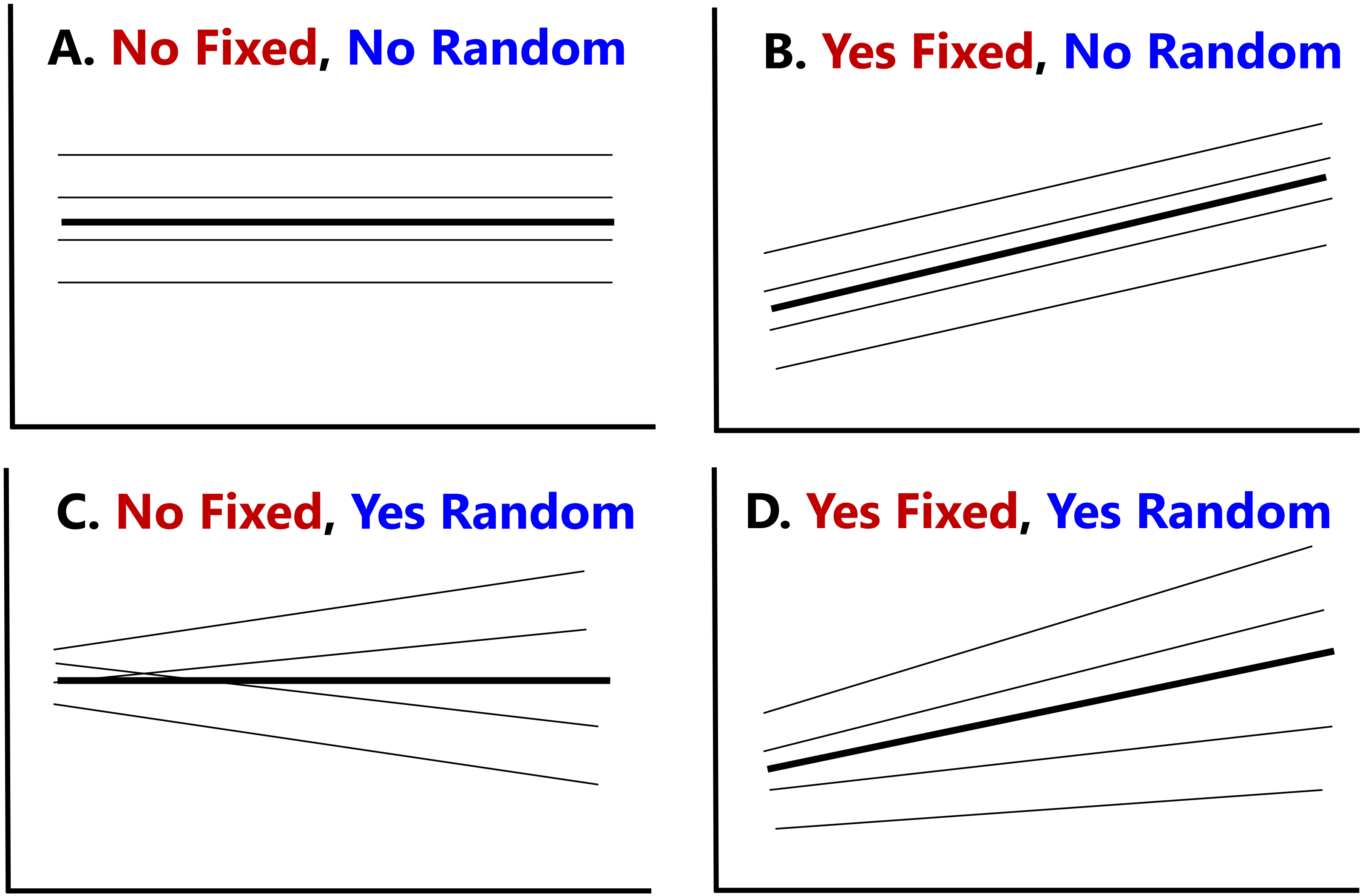
- 2 questions about the possible effects of *time*:

1. **Is there an effect of time on average?**
   - Is 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)



B. Fixed Linear Time, Random Intercept Model
(4 parameters: effect of 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 at exactly the same rate.

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:

$$\text{Pseudo } R_{U0}^2 = \frac{\text{random intercept variance}_{\text{fewer}} - \text{random intercept variance}_{\text{more}}}{\text{random intercept variance}_{\text{fewer}}}$$

  - 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 $\tau^2_{U_0}$ will often increase as a consequence of reducing level-1 residual variance σ^2_e
- Observed level-2 $\tau^2_{U_0}$ is NOT just between-person variance
  - Also has a small part of within-person variance (level-1 σ^2_e), or: **Observed $\boldsymbol{\tau^2_{U_0}}$ = True $\boldsymbol{\tau^2_{U_0}}$ + ($\boldsymbol{\sigma^2_e/n}$)**
    - As *n* occasions increases, bias of level-1 σ^2_e is minimized
  - Likelihood-based estimates of "true" $\tau^2_{U_0}$ use (σ^2_e/n) as correction factor: **True $\boldsymbol{\tau^2_{U_0}}$ = Observed $\boldsymbol{\tau^2_{U_0}}$ – ($\boldsymbol{\sigma^2_e/n}$)**
- For example: observed level-2 $\tau^2_{U_0}$=4.65, level-1 σ^2_e=7.06, *n*=4
  - True $\tau^2_{U_0}$= 4.65 –(**7.60**/4) = **2.88** in empty means model
  - Add fixed linear time slope → reduce σ^2_e from 7.06 to 2.17 (R^2 = .69)
  - But now True $\tau^2_{U_0}$= 4.65 –(**2.17**/4) = **4.10** in fixed linear time model

Variance Accounted for… For Real

- **Pseudo-R^2** 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!

- **A simpler alternative: Total R^2**
  - Generate model-predicted y's from fixed effects only (NOT including random effects) and correlate with observed y's
  - Then square correlation → total R^2
  - Total R^2 = total reduction in overall variance of y across levels
  - Can be "unfair" in models with large unexplained sources of variance

- Always specify EXACTLY which kind of pseudo-R^2 you used—give the formula and the reference!

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

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

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

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

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

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

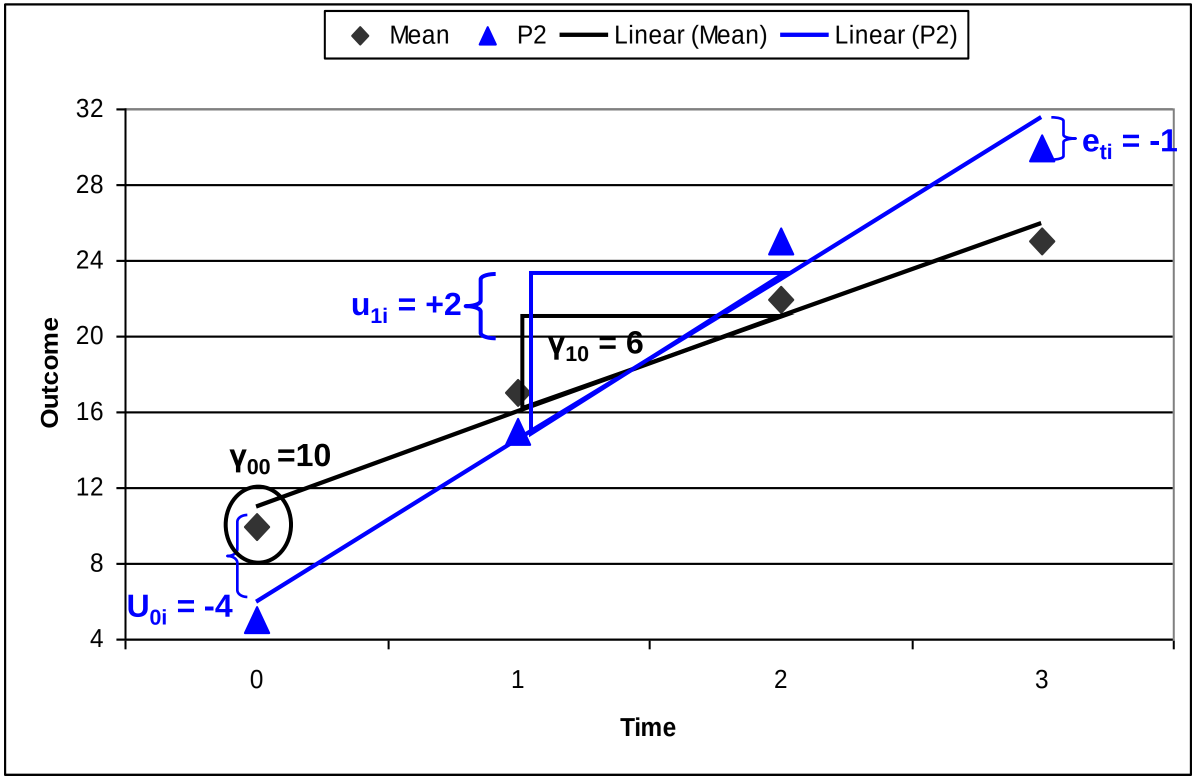
Composite Model

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

Random Linear Time Model

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

Fixed Intercept — Random **Intercept** Deviation — **Fixed Slope** — Random **Slope** Deviation — error for person i at time t



6 Parameters:

2 Fixed Effects:

γ_{00} Intercept, γ_{10} Slope

2 Random Effects Variances:

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

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

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

1 e_{ti} Residual Variance $= \sigma^2_e$

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 γ_{10} **= 1.72**, so people increase by 1.72/time on average. I also have a significant random linear time slope variance of $\tau^2_{U_1}$ **= 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:

$$\text{Random Effect 95\% CI} = \text{fixed effect} \pm \left(1.96 * \sqrt{\text{Random Variance}}\right)$$

$$\text{Linear Time Slope 95\% CI} = \gamma_{10} \pm \left(1.96 * \sqrt{\tau^2_{U_1}}\right) \rightarrow 1.72 \pm \left(1.96 * \sqrt{0.91}\right) = -0.15 \text{ to } 3.59$$

  - 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 over Time

- Topics:
  - Multilevel modeling notation and terminology
  - Fixed and random effects of linear time
  - **Predicted variances and covariances from random slopes**
  - **How random effects model dependency**
  - Describing nonlinear change: polynomial models
  - Describing nonlinear change: other alternatives

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 $\tau^2_{U_0}$ in the **G** matrix), the $\mathbf{e_{ti}}$ **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

$$\begin{bmatrix} \tau^2_{U_0} \end{bmatrix}$$

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

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

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

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

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:

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_e & 0 & 0 & 0\\ 0 & \sigma^2_e & 0 & 0\\ 0 & 0 & \sigma^2_e & 0\\ 0 & 0 & 0 & \sigma^2_e\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e\end{bmatrix}$$

VCORR then provides the intraclass correlation, calculated as:

ICC = $\boldsymbol{\tau^2_{U_0} / (\tau^2_{U_0} + \sigma^2_e)}$

$$\begin{bmatrix}1 & \text{ICC} & \text{ICC} & \text{ICC}\\ \text{ICC} & 1 & \text{ICC} & \text{ICC}\\ \text{ICC} & \text{ICC} & 1 & \text{ICC}\\ \text{ICC} & \text{ICC} & \text{ICC} & 1\end{bmatrix}$$

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

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 $\tau^2_{U_0}$ and $\tau^2_{U_1}$ 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

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

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

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

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

Random Linear Time Model
(6 parameters: effect of time is now RANDOM)

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

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

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

$\mathbf{\beta_{1i} = \gamma_{10} + U_{0i}}$

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

Predicted *Time-Specific* Variance:

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

Random Linear Time Model

(6 parameters: effect of time is now RANDOM)

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

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

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

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

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

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

$$\mathrm{Cov}\left[y_{Ai}, y_{Bi}\right] = \mathrm{Cov}\left[\left\{\left(\gamma_{00}+U_{0i}\right)+\left(\gamma_{10}+U_{1i}\right)\left(A_i\right)+e_{Ai}\right\},\left\{\left(\gamma_{00}+U_{0i}\right)+\left(\gamma_{10}+U_{1i}\right)\left(B_i\right)+e_{Bi}\right\}\right]$$

$$= \mathrm{Cov}\left[\left\{U_{0i}+\left(U_{1i}A_i\right)\right\},\left\{U_{0i}+\left(U_{1i}B_i\right)\right\}\right]$$

$$= \mathrm{Cov}\left[U_{0i},U_{0i}\right]+\mathrm{Cov}\left[U_{0i},U_{1i}B_i\right]+\mathrm{Cov}\left[U_{0i},U_{1i}A_i\right]+\mathrm{Cov}\left[U_{1i}A_i,U_{1i}B_i\right]$$

$$= \left\{\mathrm{Var}\left(U_{0i}\right)\right\}+\left\{\left(A_i+B_i\right)*\mathrm{Cov}\left(U_{0i},U_{1i}\right)\right\}+\left\{\left(A_iB_i\right)\mathrm{Var}\left(U_{1i}\right)\right\}$$

$$= \left\{\tau^2_{U_0}\right\}+\boxed{\left\{\left(A_i+B_i\right)\tau_{U_{01}}\right\}+\left\{\left(A_iB_i\right)\tau^2_{U_1}\right\}}$$

Random Linear Time Model

(6 parameters: effect of time is now RANDOM)

- Scalar "mixed" model equation per person:

$$\mathbf{Y}_i = \boxed{\mathbf{X}_i * \boldsymbol{\gamma}} + \boxed{\mathbf{Z}_i * \mathbf{U}_i + \mathbf{E}_i}$$

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

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

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

$\mathbf{X}_i$ = *n* x *k* values of **predictors with fixed effects**, so can differ per person (*k* = 2: intercept, linear time)

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

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

$\mathbf{U}_i$ = *u* x 2 estimated individual **random effects**, so can differ per person

$\mathbf{E}_i$ = *n* x *n* time-specific residuals, so can differ per person

Random Linear Time Model
(6 parameters: effect of time is now 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 = complicated ☺

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

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

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

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

$\mathbf{Z}_i$ = *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 with **unbalanced time** observations:

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

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

- 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:

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

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

- 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 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 differing variance and covariance due to other predictors, too

Compound Symmetry (CS)

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

Unstructured (UN)

$$\begin{bmatrix} \sigma^2_{T1} & \sigma_{T12} & \sigma_{T13} & \sigma_{T14} \\ \sigma_{T21} & \sigma^2_{T2} & \sigma_{T23} & \sigma_{T24} \\ \sigma_{T31} & \sigma_{T32} & \sigma^2_{T3} & \sigma_{T43} \\ \sigma_{T41} & \sigma_{T42} & \sigma_{T43} & \sigma^2_{T4} \end{bmatrix}$$

Parsimony ← → **Good fit**

Univariate RM ANOVA

Random effects models use **G** and **R** to predict something in-between!

Multivariate RM ANOVA

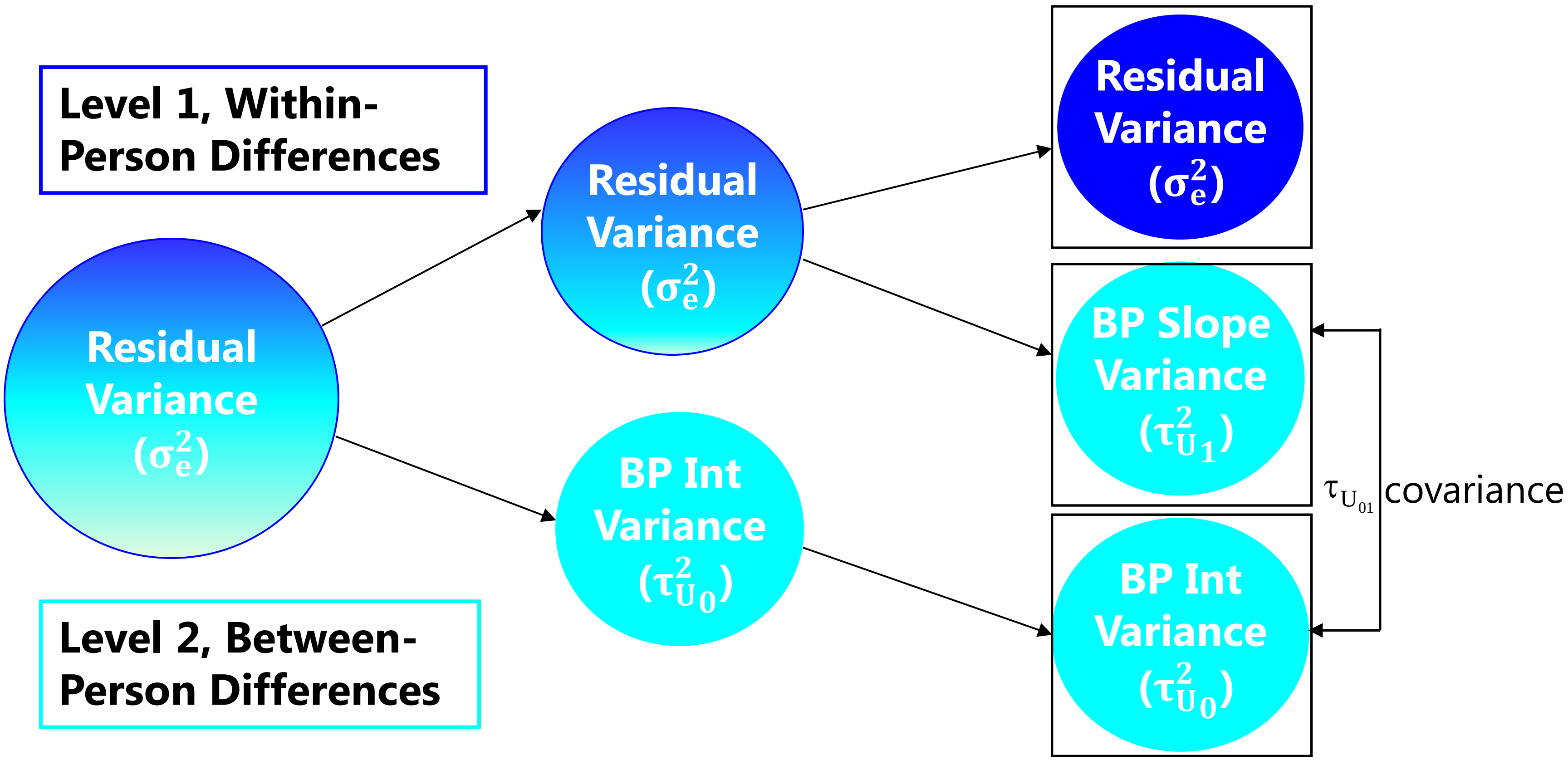
How MLM “Handles” Dependency

- 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 alternative covariance structure 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 just 1 pile of "error variance" in Univariate RM ANOVA

- **But making piles does NOT make error variance go away…**

Level 1 (one source of) <u>Within-Person Variation</u>: gets accounted for by time-level predictors

Residual Variance (σ_e^2)

FIXED effects make variance go away (explain variance).

RANDOM effects just make a new pile of variance.

Level 2 (two sources of) <u>Between-Person Variation</u>: gets accounted for by person-level predictors

BP Int Variance ($\tau_{U_0}^2$)

BP Slope Variance ($\tau_{U_1}^2$)

$\tau_{U_{01}}$ **covariance**

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 variance for U_{0i}**
  - Individual time slope differences can be included as:
    - **N-1*time person dummy code interactions OR 1 random variance for U_{1i}**
  - 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**

Describing Within-Person Change over Time

- Topics:
  - Multilevel modeling notation and terminology
  - Fixed and random effects of linear time
  - Predicted variances and covariances from random slopes
  - How random effects model dependency
  - **Describing nonlinear change: polynomial models**
  - Describing nonlinear change: other alternatives

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 Variance**
  - 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?

The Big Picture of Longitudinal Data: Model for the Means (Fixed Effects)

- What kind of change occurs on average over "time"?
  - What is the most appropriate **metric of time**?
    - Time in study (with predictors for BP differences in time)?
    - Time since birth (age)? Time to event (time since diagnosis)?
    - Measurement occasions need not be the same across persons or equally spaced (code time as exactly as possible)
  - What kind of **theoretical process** generated the observed trajectories, and thus what kind of model do we need?
    - Linear or nonlinear? Continuous or discontinuous? Does change keep happening or does it eventually stop?
    - Many options: polynomial, piecewise, and nonlinear families

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 <u>together</u> to describe curved trajectories

- **Can have polynomial fixed time slopes UP TO: *n* – 1***
  - 3 occasions = 2nd order ($time^2$)= Fixed Quadratic Time or less
  - 4 occasions = 3rd order ($time^3$) = 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: $y_T = 50.0 + 4.0T + 0.3T^2$

First Derivative (Velocity) at Time T: $\dfrac{d\,y_T}{d(T)} = 4.0 + 0.6T$

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

Interpreting Quadratic Fixed Effects

<u>A Quadratic time effect is a two-way interaction: time*time</u>

- Fixed quadratic time = "**<u>half</u> 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:

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

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

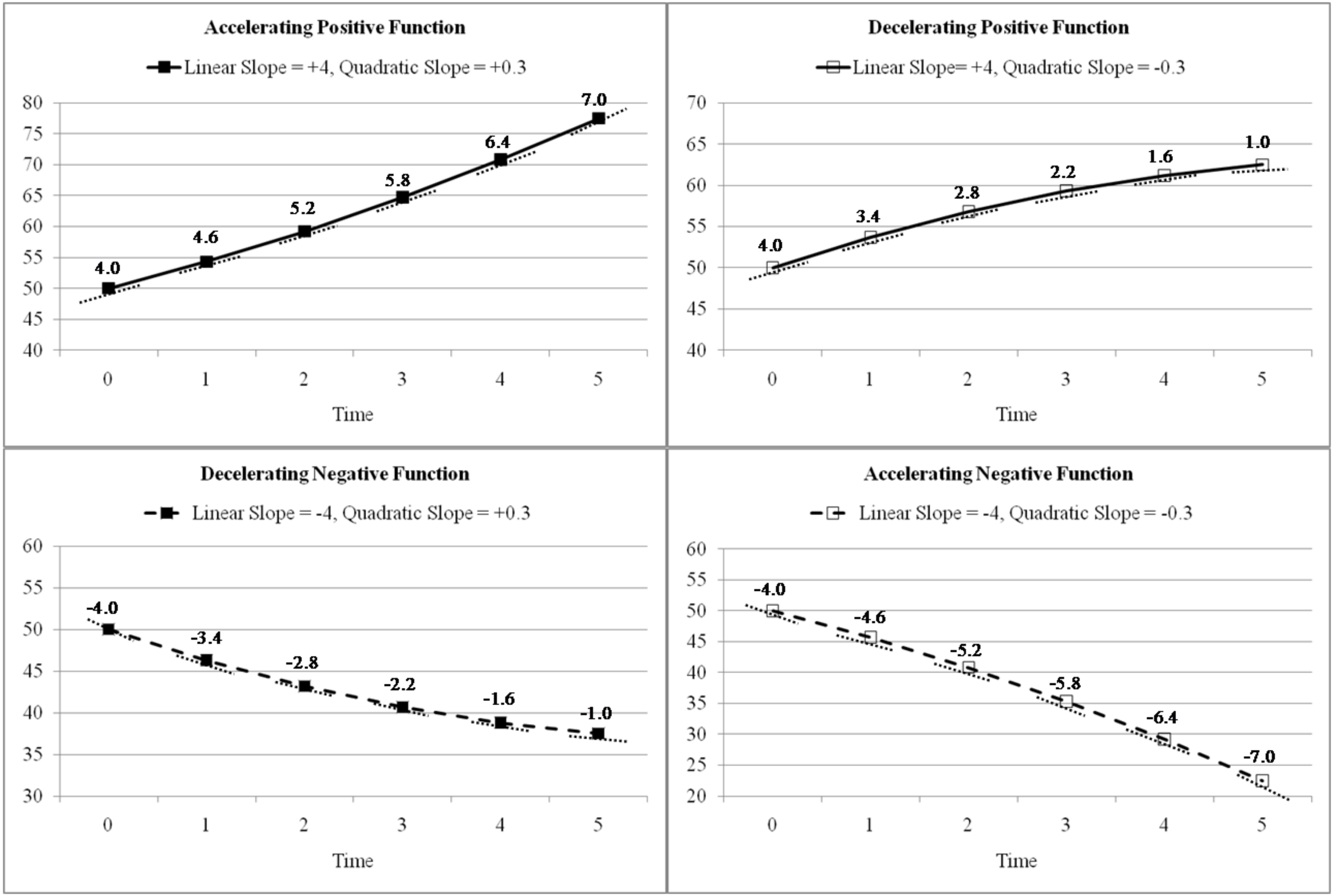
$$\text{Effect of } Z = \beta_2 + \beta_3 X$$

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

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

- 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 <u>twice</u> to the <u>one</u> (main) linear effect of time.

Examples of Fixed Quadratic Time Effects



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)

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)

Random Quadratic Time Model

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

Level 2 Equations (one per β):

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

β_{0i}: **Intercept** for person *i*

γ_{00}: Fixed (mean) **Intercept**

U_{0i}: Random (Deviation) **Intercept**

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

β_{1i}: **Linear Slope** for person *i*

γ_{10}: Fixed (mean) **Linear Slope**

U_{1i}: Random (Deviation) **Linear Slope**

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

β_{2i}: **Quad Slope** for person *i*

γ_{20}: Fixed (mean) **Quad Slope**

U_{2i}: Random (Deviation) **Quad Slope**

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 = $n - 2$

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

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?***
    - <u>Random Intercept</u> = BP variance *for any occasion* in predicted mean Y (= variance in grand mean because individual lines are parallel)
  - ***Add Random Linear Time?***
    - <u>Random Intercept</u> = **now** BP variance *at time=0* in predicted mean Y *(would be different if time was centered elsewhere)*
    - <u>Random Linear Time</u> = BP variance *across all occasions* in linear rate of change *(would be the same if time was centered elsewhere)*
  - ***Add Random Quadratic Time?***
    - <u>Random Intercept</u> = still BP variance *at time=0* in predicted mean Y
    - <u>Random Linear Time</u> = **now** BP variance *at time=0* in linear rate of change *(would be different if time was centered elsewhere)*
    - <u>Random Quadratic Time</u> = 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

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

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…

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 −2ΔLL(df≈2) for random linear variance
4. Fixed Quadratic, Random Linear → check fixed quadratic *p*-value
5. Random Quadratic → check −2ΔLL(df≈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 **−2ΔLL**

→ Also see if AIC and BIC are smaller when adding random effects

Likelihood-Based Model Comparisons for Deciding Between Alternative Models for the Variance (again)

- **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**
  - Two flavors (labeled as −2 log likelihood in SAS, SPSS, but given as LL instead in STATA and Mplus): Maximum Likelihood (**ML**) or Restricted (Residual) ML (**REML**)

- **Nested models are compared using their deviance values: −2ΔLL Test** (i.e., Likelihood Ratio Test, Deviance Difference Test)

1. Calculate −2ΔLL: $(-2LL_{fewer}) - (-2LL_{more})$
2. Calculate Δdf: $(\#\ Parms_{more}) - (\#\ Parms_{fewer})$
3. Compare −2ΔLL to χ^2 distribution with df = Δdf
   CHIDIST in excel will give exact p-values for the difference test; so will STATA

1. & 2. must be positive values!

- **Add** parameters? Model fit can be **BETTER** (signif) or **NOT BETTER**
- **Remove** parameters? Model fit can be **WORSE** (signif) or **NOT WORSE**

Summarizing so far…

- Modeling within-person change involves specifying effects of time for both sides of the model
  - **Fixed effects in model for the means:**
    - What kind of change am I observing on average?
    - What kind of trajectory will reproduce those means?
  - **Random effects (and residuals) in model for the variance:**
    - What kind of individual differences in change am I observing?
    - How many random effects do I need to reproduce the observed pattern of variances and covariances over time?

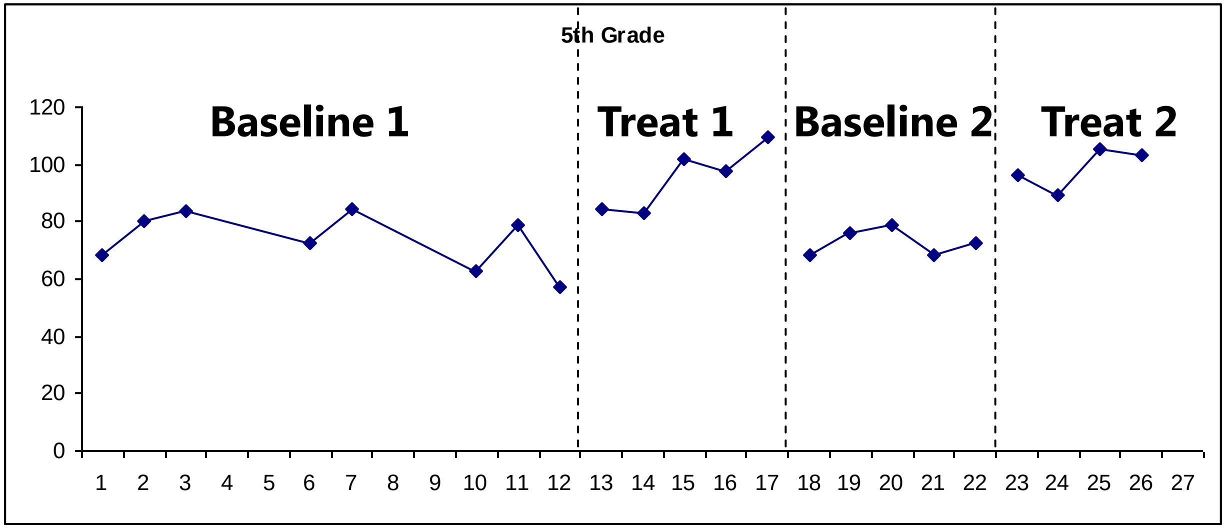
- One option: Polynomial models (linear, quadratic, cubic)
  - Terms work together to describe non-linear trajectories
  - Careful with the covariances among random effects, though
- Other options: Piecewise slopes and “truly” nonlinear change…

Describing Within-Person Change over Time

- Topics:
  - Multilevel modeling notation and terminology
  - Fixed and random effects of linear time
  - Predicted variances and covariances from random slopes
  - How random effects model dependency
  - Describing nonlinear change: polynomial models
  - **Describing nonlinear change: other alternatives**

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!**



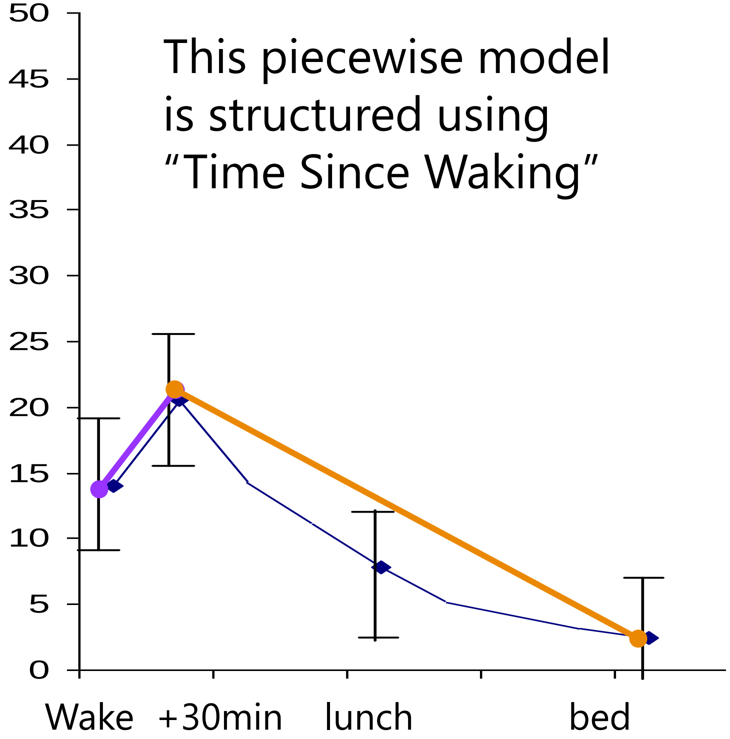
Piecewise Model:

4 slopes
(one per phase)

3 "jumps"
(shift in intercept between phases)

Example of Daily Cortisol Fluctuation: Morning Rise and Afternoon Decline

Average Trajectories

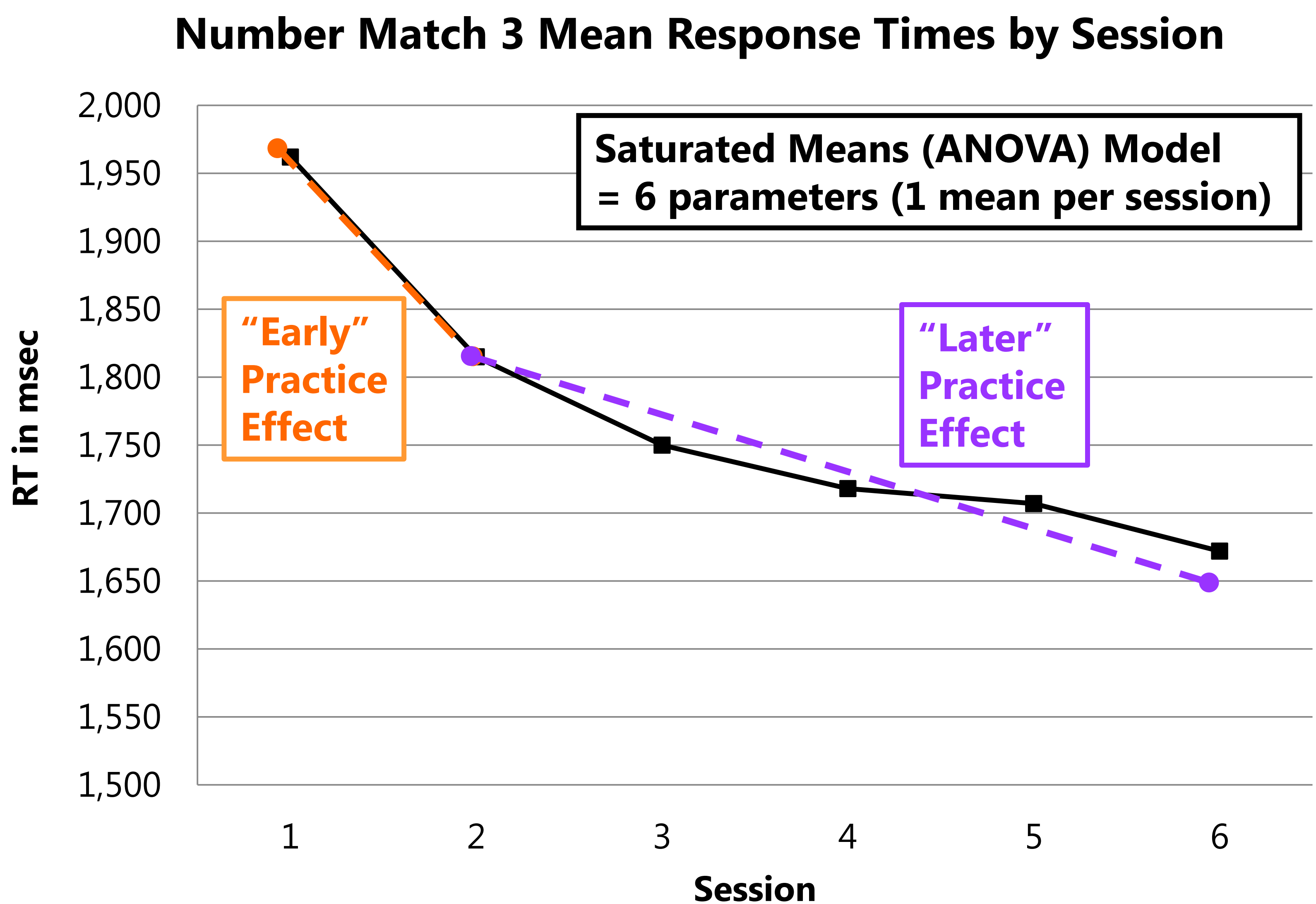


SAS Code to create two piecewise slopes from continuous time of day in stacked data:

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;

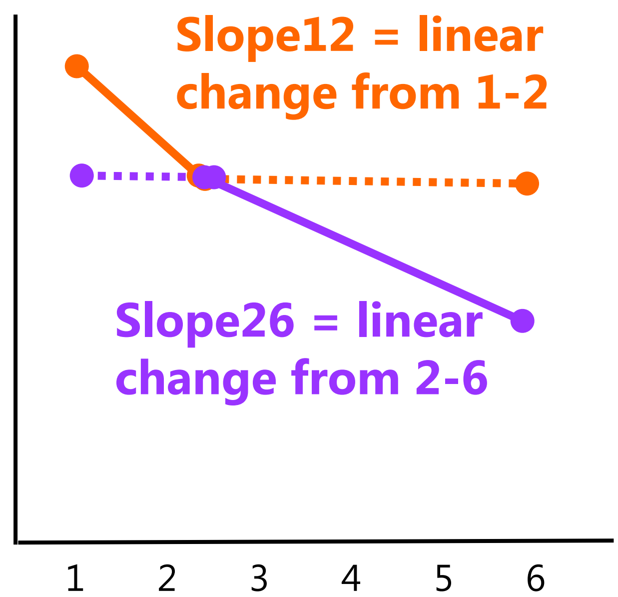
Note that a quadratic slope may be necessary for the afternoon decline slope!

What kind of piecewise model could predict our example data mean change across sessions?



Piecewise Models: Two Direct Slopes

- "Early Practice Slope" and "Later Practice Slope"
- Use to specify slopes through each discrete phase directly (can request test of difference)
- Session (1-6) gets recoded into 2 new time predictor variables, as shown below:



Session	**1**	**2**	**3**	**4**	**5**	**6**
Early Practice → Slope12 =	**0**	**1**	**1**	**1**	**1**	**1**
Later Practice → Slope26 =	**0**	**0**	**1**	**2**	**3**	**4**

2 Direct Slopes Model: Random Effects

- Parameters directly **represent each part** of trajectory:

- **Fixed effects** for mean change over time (3 fixed effects):
  - Fixed Intercept = expected Y when both slopes = 0 (Session 1)
  - Fixed Slope12 = expected linear rate of change from 1 to 2
  - Fixed Slope26 = expected linear rate of change from 2 to 6

- Leads to possible **random effects** (up to 3 var+3 cov):
  - Random Intercept = BP variance in expected level
    when both slopes = 0 (at Session 1)
  - Random Slope12 = BP variance in linear slope from 1 to 2
  - Random Slope26 = BP variance in linear slope from 2 to 6

Random Two-Slope Piecewise Model

Level 1: $\mathbf{y_{ti} = \beta_{0i} + \beta_{1i}Slope1_{ti} + \beta_{2i}Slope2_{ti} + e_{ti}}$

Level 2 Equations (one per β):

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

Intercept for person *i*; Fixed (mean) **Intercept**; Random (Deviation) **Intercept**

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

Slope1 for person *i*; Fixed (mean) **Slope1**; Random (Deviation) **Slope1**

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

Slope2 for person *i*; Fixed (mean) **Slope2**; Random (Deviation) **Slope2**

Fixed Effect Subscripts:

1st = which Level 1 term

2nd = which Level 2 term

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

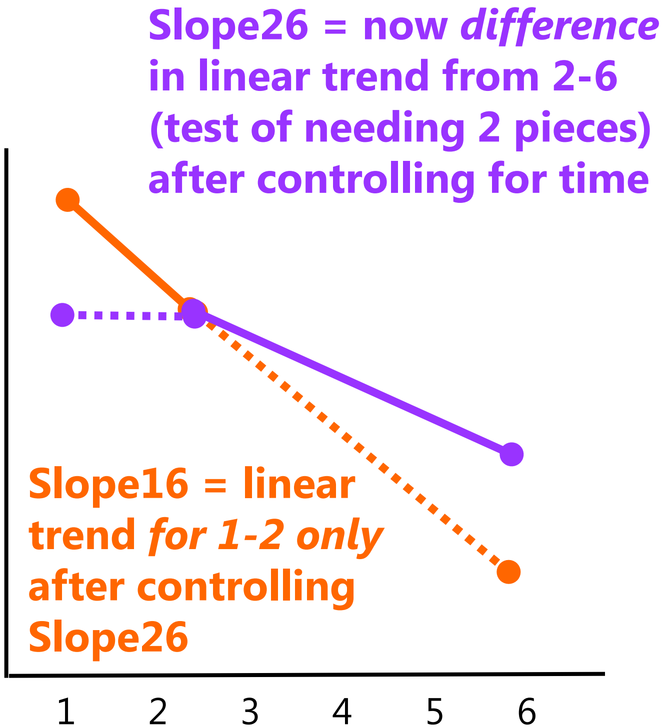
\# Fixed slopes = $n - 1$

\# Random slopes = $n - 2$

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

Piecewise Models: Slope + Deviation Slope

- "Linear Time Slope" and "Deviation Slope"
- Use to test if multiple slopes are needed directly in model
- Initial slope predictor is coded differently, second slope predictor is same:



Session		1	2	3	4	5	6
Time	→ Slope16 =	0	1	2	3	4	5
Deviation	→ Slope26 =	0	0	1	2	3	4

Slope + Deviation Slope: Random Effects

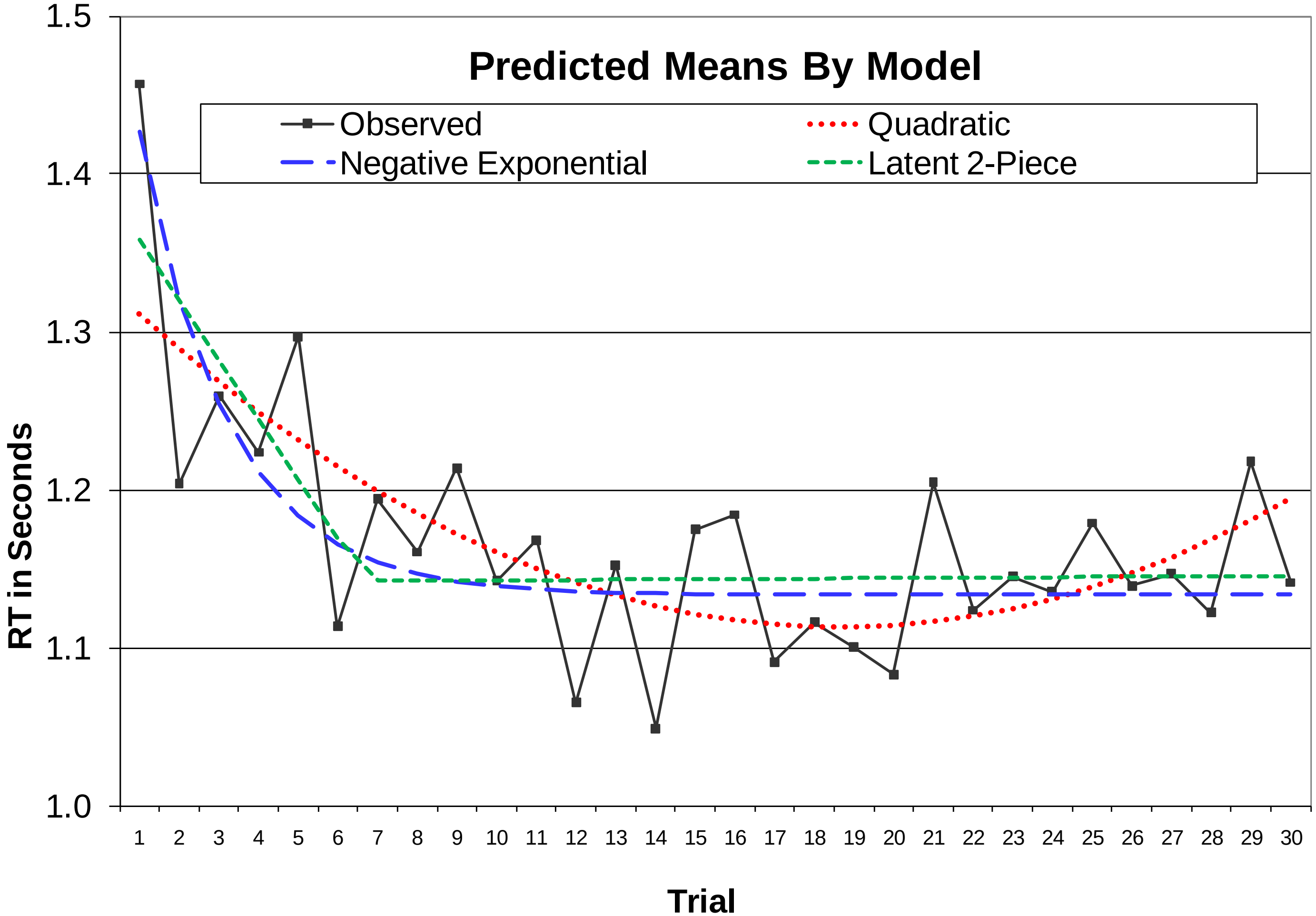
- Parameters directly **differences across parts** of trajectory:

- **Fixed effects** for mean change over time (3 fixed effects):
  - Fixed Intercept = expected Y when both slopes = 0 (Session 1)
  - Fixed Slope16 = expected linear rate of change from 1 to 2
    (after controlling for slope26)
  - Fixed Slope26 = expected **extra** linear rate of change from 2 to 6
    (after controlling for slope16, which is just time)

- Leads to possible **random effects** (up to 3 var+3 cov):
  - Random Intercept = BP variance in expected level
    when both slopes = 0 (at Session 1)
  - Random Slope16 = BP variance in linear slope from 1 to 2
  - Random Slope26 = BP variance in **extra** linear slope from 2 to 6

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)



Nonlinear Models

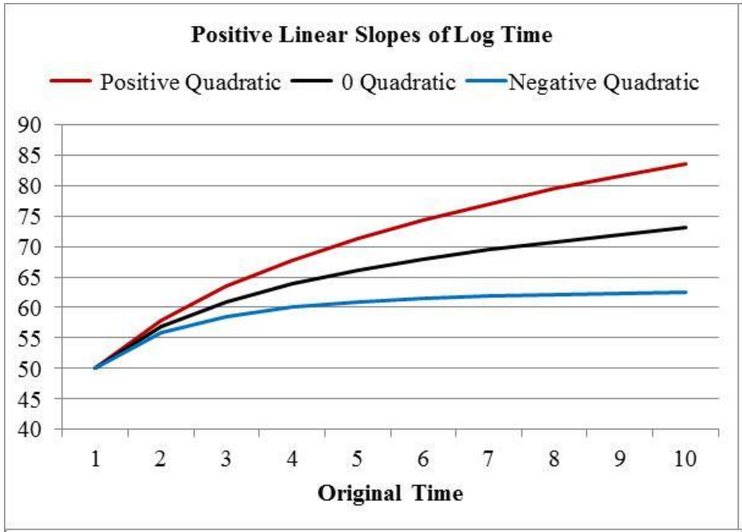
- Not all forms of change fit polynomial models
  - What goes up must come back down (and vice-versa)
  - Sometimes change needs to "shut off" (need asymptotes)

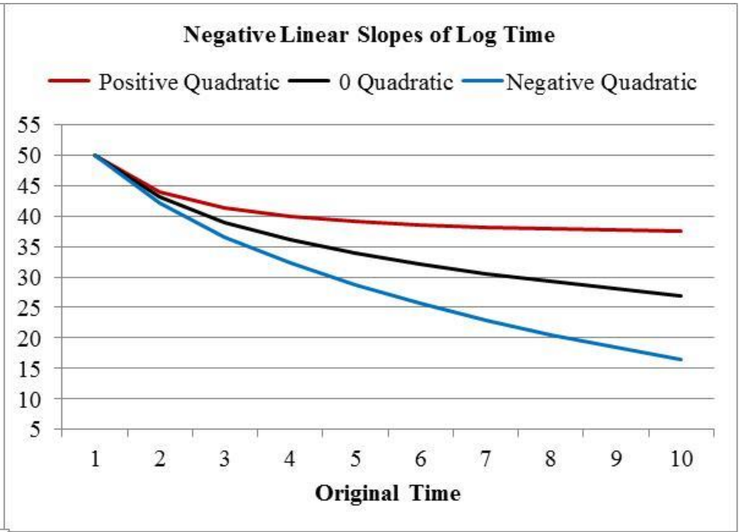
- Many kinds of truly nonlinear models can be used for longitudinal data
  - Linear in variables vs. linear in parameters (exp → nonlinear)
  - Logistic, power, exponential… see end of chapter 6 for ideas

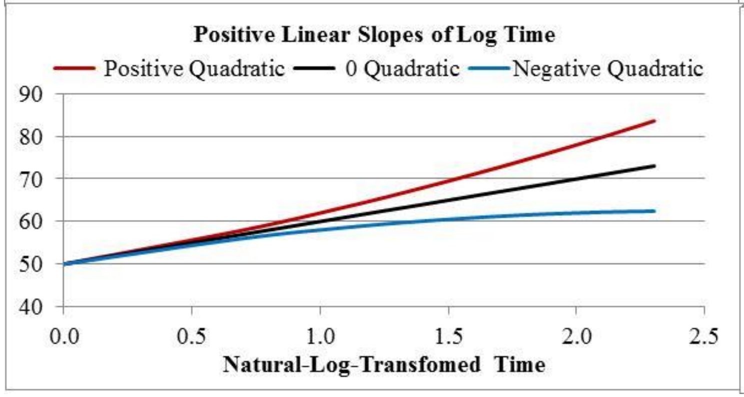
- Require extra steps to evaluate estimation quality
  - Start values are needed, especially for random variances
  - Check that "gradient" values are as close to 0 as possible (partial first derivative of that parameter in LL function)

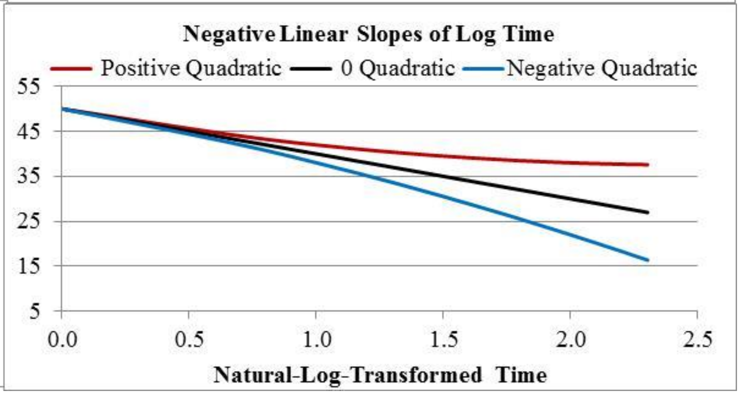
How to Mimic an Exponential Model

If you need to use REML, a predictor of natural-log-transformed time may be a good substitute for a truly nonlinear model









A <u>linear</u> effect of log time (black lines) predicts an <u>exponential</u> curve across *original* time.

Quadratic effects of log time (red or blue lines) can speed up or slow down the curve.

Bottom: There is a linear relationship between log-time and the outcome.

Which change family should I choose?

- Within a given family, nested models can usually be compared to judge the need for each parameter
  - e.g., linear vs. quadratic? One slope vs. two slopes?
  - Usual nested model comparison rules apply (*p*-values for fixed effects, $-2\Delta LL$ tests for random effects)
  - When using REML, you can test absolute fit of each side separately if you have balanced data to see if you are "there yet"

- Between families, however, alternative models of change may not be nested, so deciding among them can be tricky
  - e.g., quadratic vs. two-slope vs. log time vs. exponential?
  - Use ML AIC and BIC to see what is "preferred" across the families
  - In balanced data, you can also compare each alternative to a saturated means, UN model using ML as test of absolute fit
  - Also consider plausibility of alternative models in terms of both data predictions and theoretical predictions in deciding

New Material: Absolute Fit in REML

- Answer key model (possible only for balanced data):
  - Means Model = Saturated Means
  - Variance Model = Unstructured R, or RI+UN($n-1$) equivalent

- Tests of absolute fit of any simpler means model against saturated means can only be done via $-2\Delta LL$ when using ML, but what if you need to use REML given small level-2 N?
  - Use a multivariate Wald test instead: add enough contrasts for occasion-specific mean differences to create saturated means, then test that group of contrasts (see CLDP 944 example 6 for how to do so)

- Tests of absolute fit of any nested variance model against UN can be done using REML $-2\Delta LL$ if same means side (so keep the same fixed effects for time in each comparison model)