Review of Concepts and Terminology in Longitudinal Modeling

- Topics:
  - **Concepts and terminology in longitudinal models (and their estimation in current software)**
  - Fixed and random effects of time (and their analog through latent variables)
  - Significance testing and effect size for fixed and random effects (in MLM or SEM)
  - Modeling time-invariant predictors

Levels of Analysis in Longitudinal Data

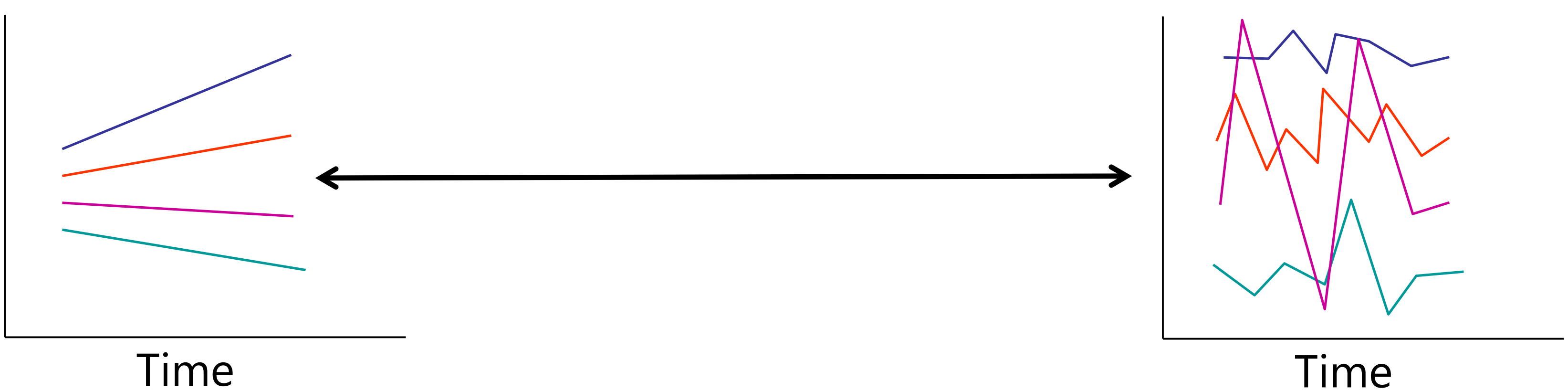
- **Between-Person (BP)** Variation:
  - "**INTER**-individual Differences" – **Time-Invariant** ("macro" level 2)
  - All longitudinal studies that begin as cross-sectional studies have this
- **Within-Person (WP)** Variation:
  - "**INTRA**-individual Differences" – **Time-Varying** ("micro" level 1)
  - 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
  - BP = more/less than other people; WP = more/less than one's average
- I use "person" here, but "between" units can be anything that is measured repeatedly (like animals, schools, countries…)

A Longitudinal Data Continuum

- **Within-Person (WP) Change**: Expect systematic effect(s) of time
  - Magnitude or direction of change can be different across individuals
  - e.g., "(Latent) Growth Curve Models" → **Time is meaningfully sampled**
- **Within-Person (WP) Fluctuation**: No expected effects of time*
  - Outcome just varies/fluctuates over time (e.g., emotion, mood, stress)
  - **Time is just a way to get lots of data per person** (e.g., EMA studies)
  - * Need to consider reactivity, day of the week, circadian/schedule effects

Pure WP Change

Pure WP Fluctuation



Characterizing Longitudinal Data

- What should "**time**" be?
  - **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), the model needs to **differentiate BP from WP time effects**
  - If there is more than one kind of WP "time" (e.g., occasions within days), the model needs to **differentiate levels of WP time effects**
- Is time *balanced* or *unbalanced*?
  - **Balanced** = **shared** measurement schedule (not necessarily equal interval)
    - Although some people may miss occasions, making their data "incomplete"
  - **Unbalanced** = people have **different** time values possible
    - By definition, observations are "incomplete" across persons
    - This may be a consequence of using a time metric that also varies between persons

The Two Sides of *Any* Model

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

- **Model for the Variance:**
  - *Aka* **Random Effects and Residuals**, Stochastic Part of Model
  - What you *were* used to **making assumptions about** instead
  - How **residuals are distributed and related** across sampling dimensions (persons, occasions) → these relationships are known as "dependency" and ***this is the primary way that longitudinal models differ from general linear models*** *(e.g., regression)*

Modeling Longitudinal Dependency

- **Sources of dependency** (reasons for the correlation of occasions from same person) can be captured by a model in three main ways:

- **Fixed effects:** Add Person ID as a predictor (*N*-1 dummy codes)
  - ID main effects capture dependency due to mean differences; interactions of ID with time-level predictors capture other predictor-specific types of person dependency
  - Does not allow prediction of why any of those person differences occurred ☹

- **Alternative multivariate variance–covariance structures (ACS):** Just allow/describe patterns over time (for unknown reasons)
  - e.g., Compound Symmetry Heterogeneous; Unstructured as the "answer key"
  - Only possible for balanced longitudinal data; those using time-lagged covariances also require equal interval occasions: AR1(H), ANTE(1), TOEP(H)

- **Add a level (or more):** Use random effect (latent factor) variances
  - Capture patterns of non-constant variance and covariance for different, testable reasons even with unbalanced longitudinal data → **LET'S TALK ABOUT THIS…**

Two* Analytic Frameworks for the Estimation of Longitudinal Models

** Possible because random effects and latent variables are the same thing*

- **"Multilevel/Mixed/Hierarchical Linear Models: MLM"**
  - MLM builds directly from regression, <u>so I always start here for longitudinal modeling</u>
  - Dependency is captured primarily by **random effects** (through "**levels**" in **stacked/long data**, so occasions can be unbalanced and cover multiple types of WP time)
  - Software for **univariate MLMs** (single outcome over time) is common (SAS, SPSS, STATA MIXED), has REML and denominator DF for **small samples**, but can't do some variants
  - Software for **multivariate MLMs** (2+ outcomes over time) is more flexible, less common (Mplus), and is more likely to break down in small samples (**no REML**; more parameters)
    - What I call "multivariate MLM" is "multilevel SEM" to others, even with no measurement models!
- **"Structural Equation Models: SEM"**
  - Dependency is captured primarily by **latent variables** (through **multivariate outcomes** in **wide, single-level data**, so univariate or multivariate occasions are treated as boxes)
  - General SEM software is common, but **no REML** or denominator DF for small samples
  - SEM is flexible with respect to model variants, but **may not work for unbalanced data** or designs with **more than one level of time** (e.g., occasions within days within persons)

A Statistician's World View

- Outcome type: General (normal) vs. General*ized* (not normal)
- Dimensions of sampling: One (so one variance term per outcome) vs. **Multiple** (so multiple variance terms per outcome) → **OUR WORLD**

- **General Linear Models:** conditionally normal outcome distribution, **fixed effects** (identity link; only one dimension of sampling)

Note: Least Squares is only for GLM

- **General*ized* Linear Models:** **any conditional outcome distribution**, **fixed** effects through **link functions**, no random effects (one dimension)
- **General Linear Mixed Models:** conditionally normal outcome distribution, **fixed and random effects** (identity link, but multiple sampling dimensions)
- **General*ized* Linear Mixed Models:** **any conditional outcome distribution**, **fixed and random effects** through **link functions** (multiple dimensions)
  - Many of the same concepts, but with more complexity in estimation

- "Linear" means fixed effects predict the *link-transformed* conditional mean of DV in a linear combination of (effect*predictor) + (effect*predictor)…

For Example: A Single-Level (BP) Model

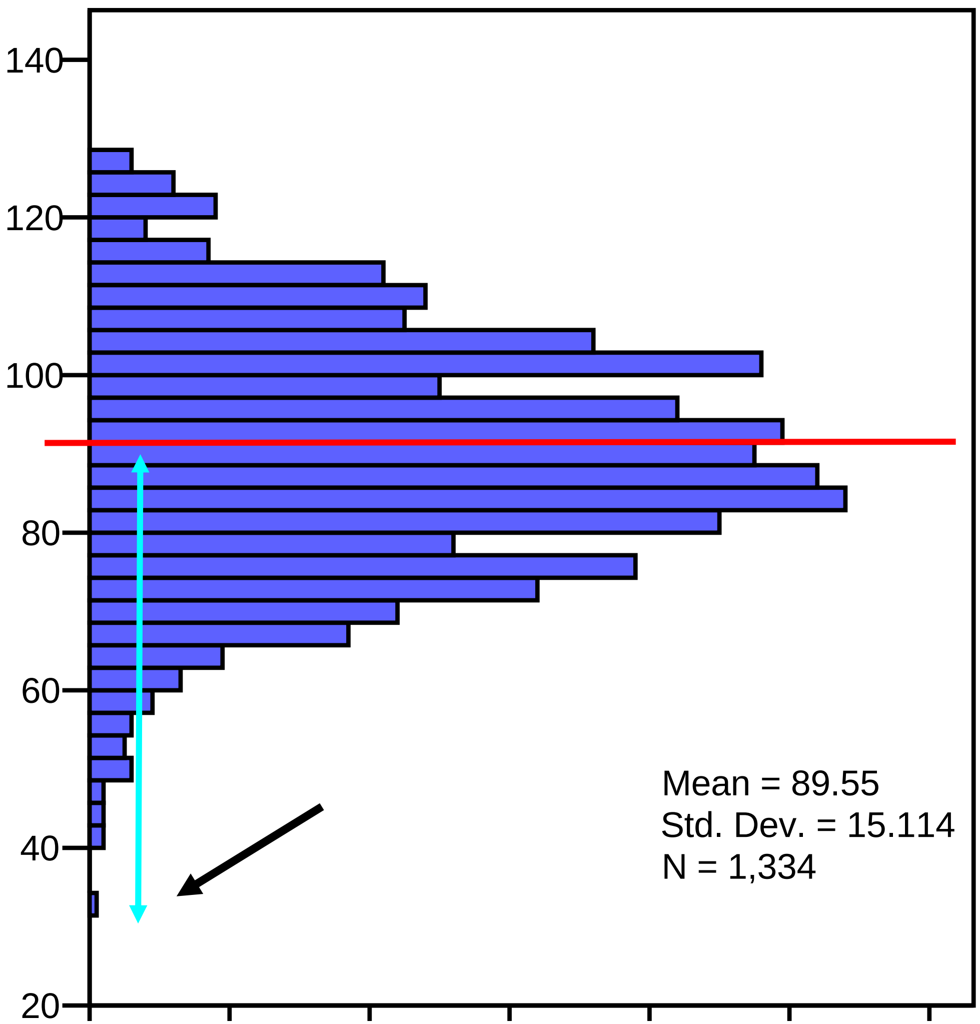
$$y_i = \boxed{\beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i} + \boxed{e_i}$$

= Single-Level

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

- **Model for the Variance (→"Piles" of Variance):**
  - $e_i \sim N(0, \sigma_e^2)$→ ONE residual (unexplained) deviation, so **estimated parameter is residual variance in single-level (BP) model**
  - e_i residuals have a mean of 0 with some estimated **constant variance** σ_e^2, are **normally distributed**, are unrelated to X and Z, and are **independent** across all observations
  - *We should change models when any of these assumptions do not hold…*

An Empty Means, BP-Only Model (i.e., Single-Level)



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

Filling in values:

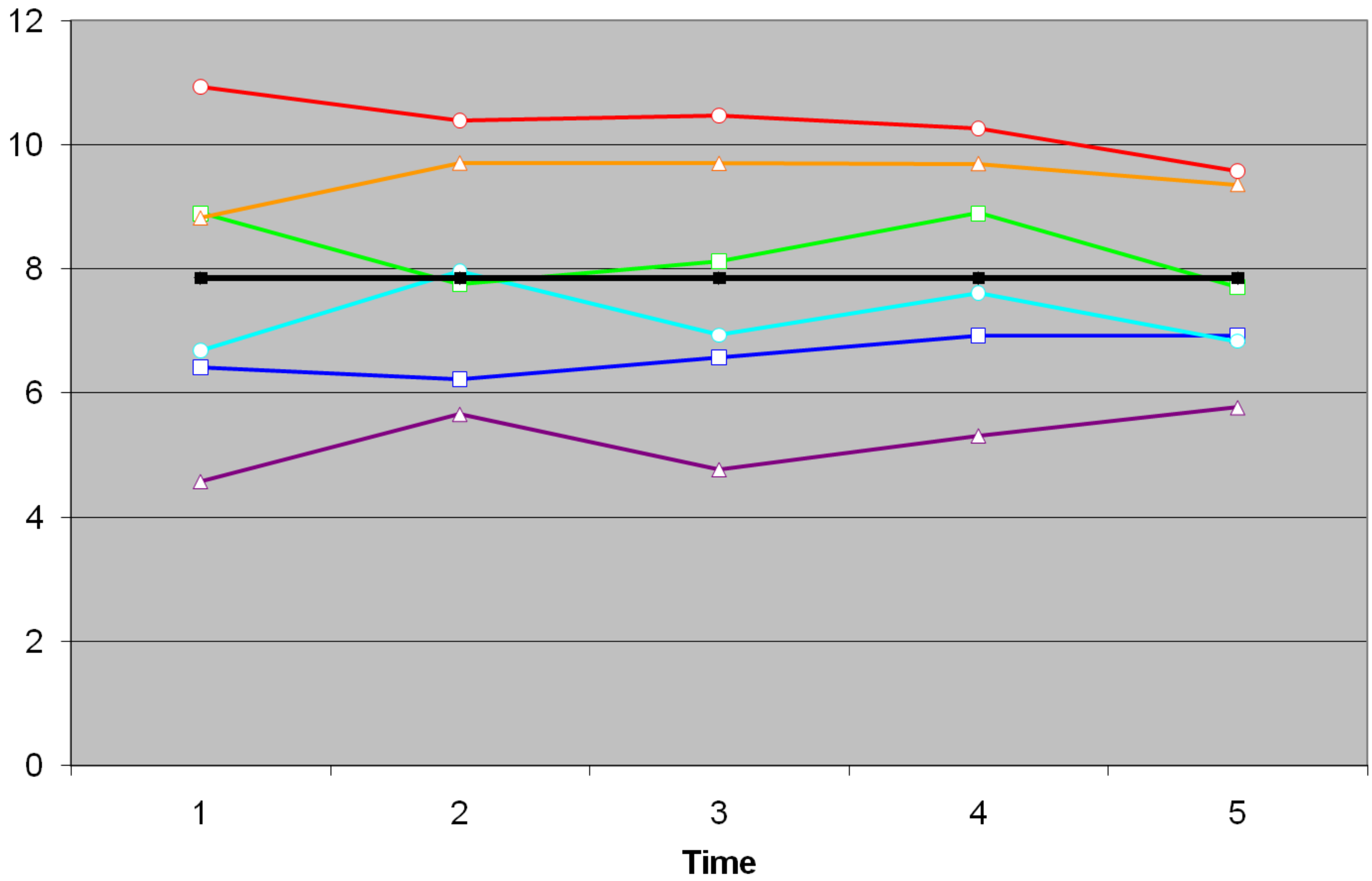
$$\mathbf{32 = \underbrace{90}_{y_{pred}} + -58}$$

Model for the Means

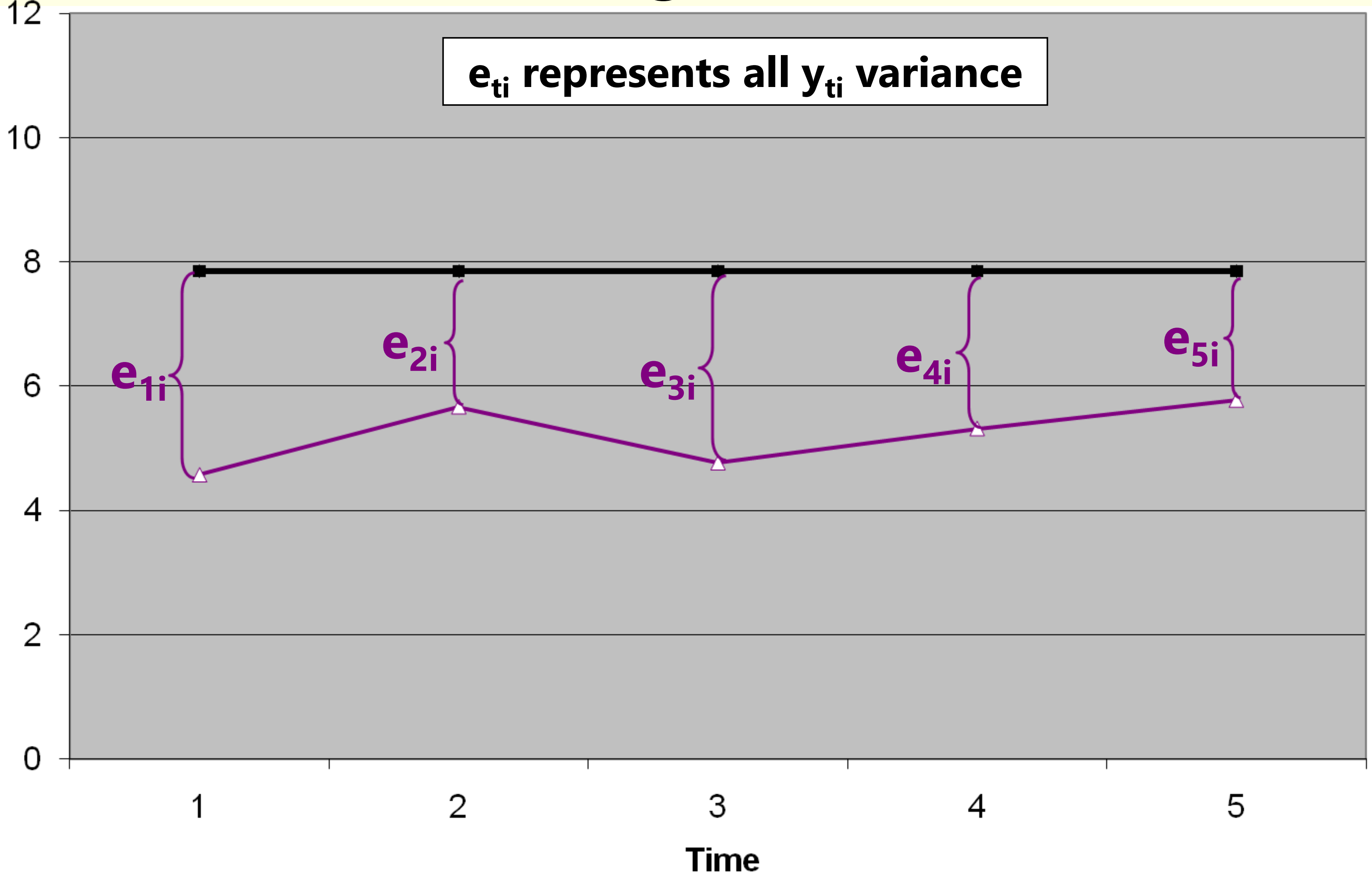
y_i error variance:

$$\frac{\sum\left(y_i - y_{pred}\right)^2}{N-1}$$

Hypothetical Longitudinal Data

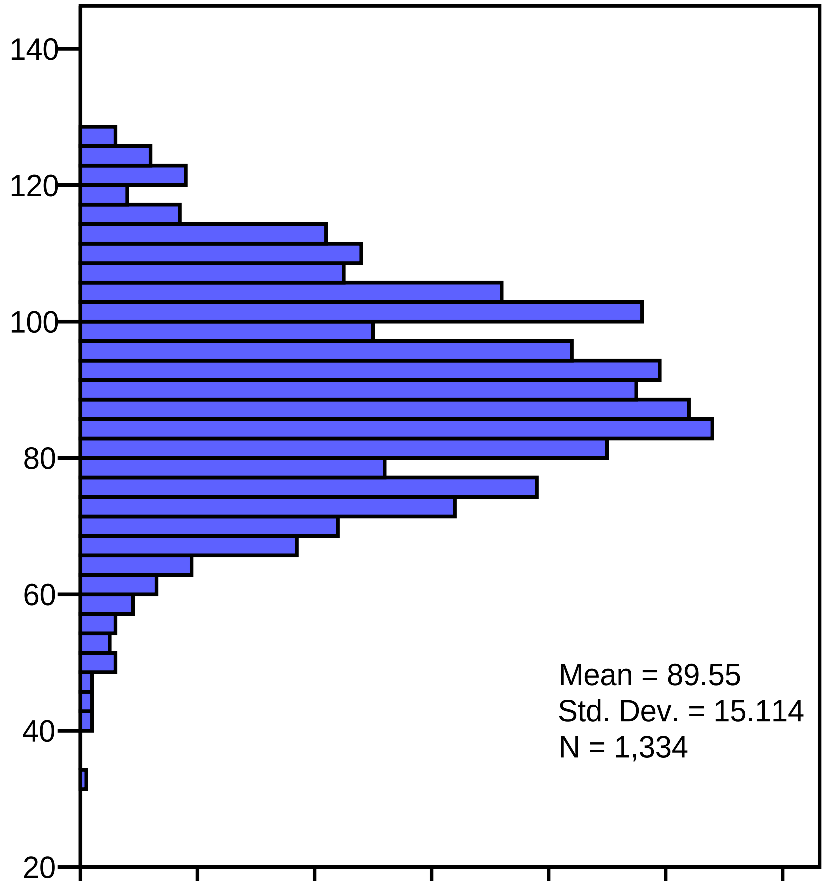


"Error" in a BP-Only Model for the Variance: Single-Level Model

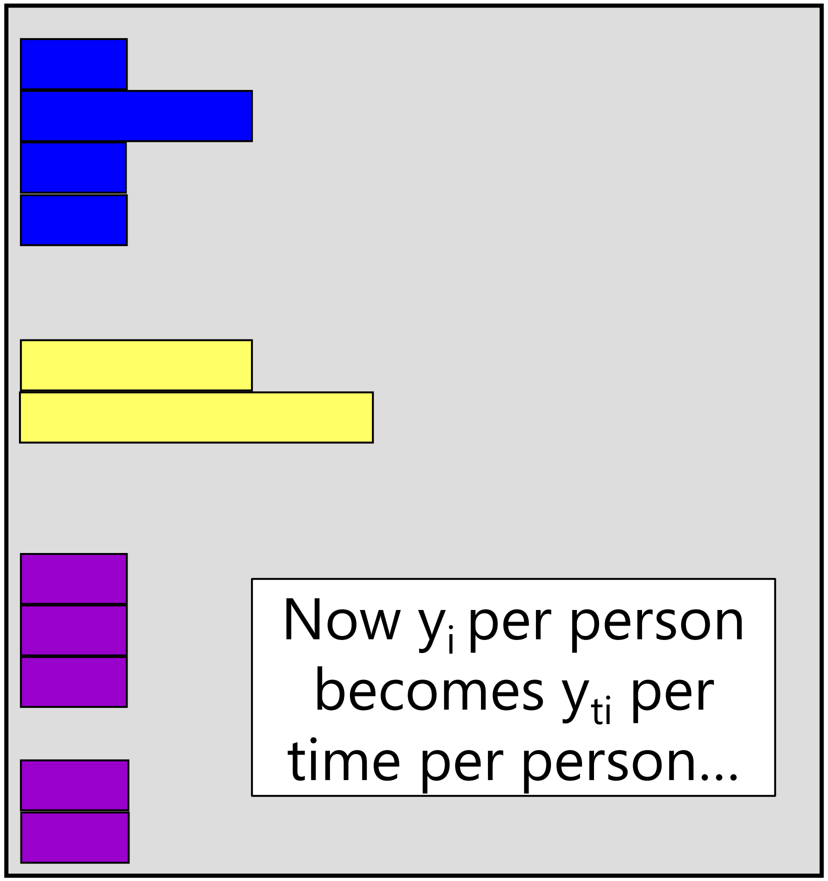


Adding Within-Person Information… (i.e., to become a Two-Level Model)

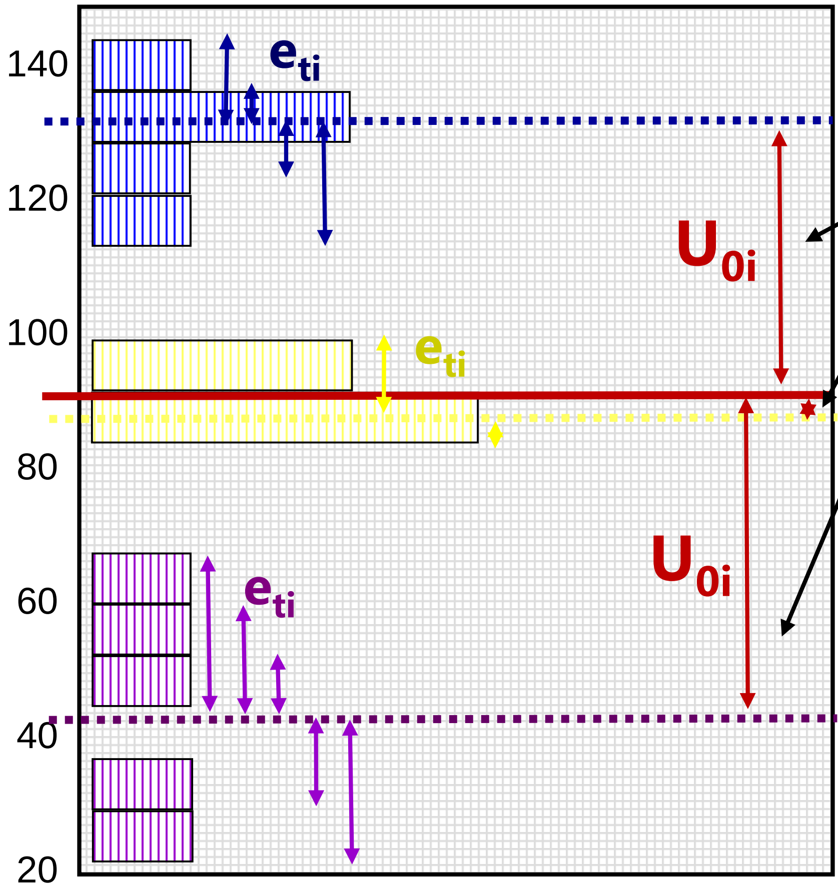
Full Sample Distribution



3 People, *n*=5 Occasions each



Empty Means, BP+WP (Two-Level) Model



y_{ti} variance (V) → 2 sources:

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

→ **Between**-Person Variance (in **G**)

→ Differences from **GRAND** mean

→ **INTER**-Individual Differences

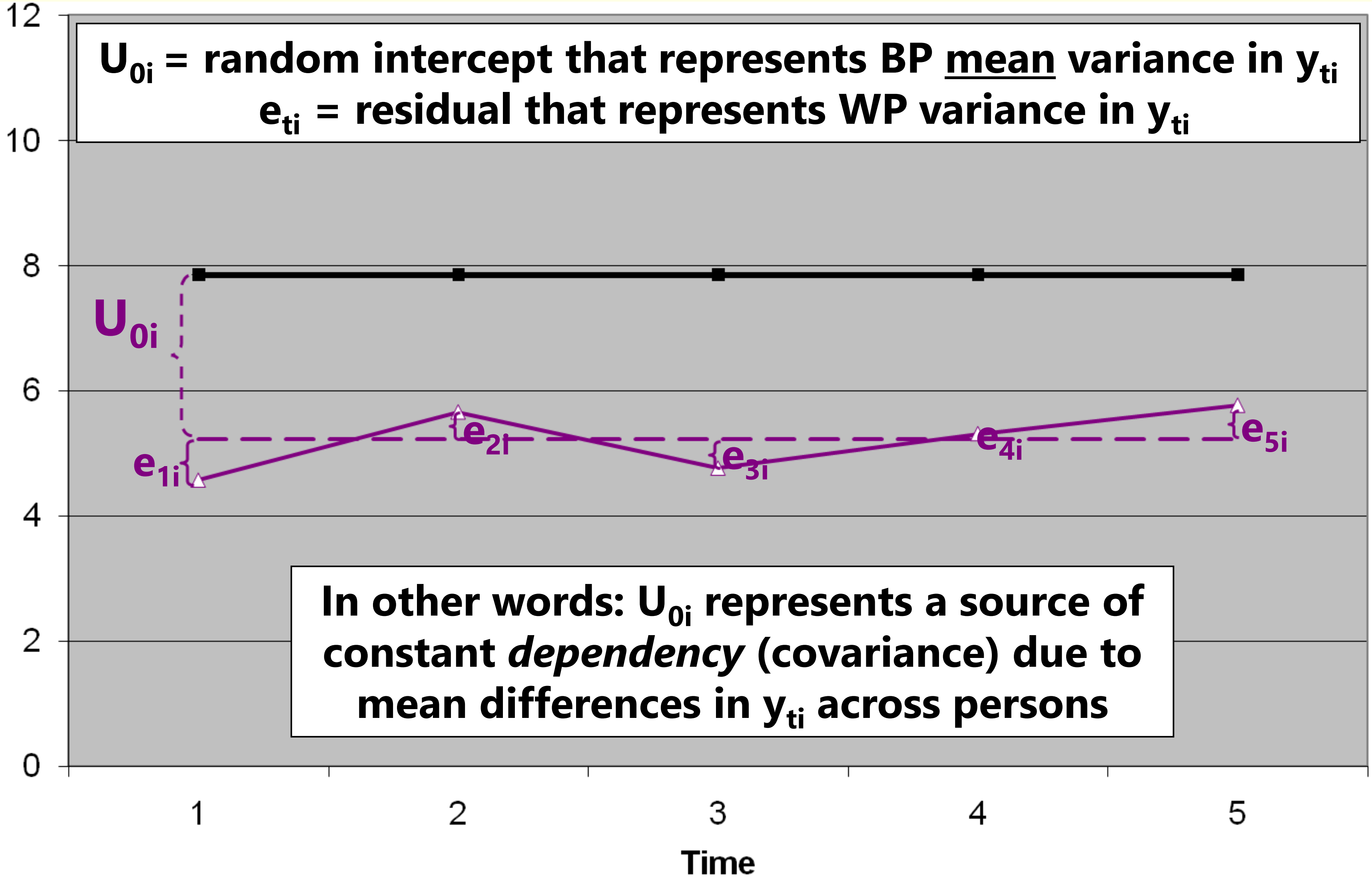
<u>Level-1 Residual Variance</u> (of e_{ti}, as σ^2_e):

→ **Within**-Person Variance (in **R**)

→ Differences from **OWN** mean

→ **INTRA**-Individual Differences

"Error" in a BP+WP Model for the Variance: Multilevel Model (two levels)



BP-Only vs. BP+WP Empty Models

- Empty Means, **BP-Only** Model (used for 1 occasion):

$$y_i = \beta_0 + e_i$$

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

- Empty Means, **BP+WP** Model (for 2+ occasions):

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

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

Same Model Using Multilevel Notation: Empty Means, Random Intercept Model

GLM Empty Model:

- $y_i = \beta_0 + e_i$

MLM Empty Model:

- Level 1:

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

- Level 2:

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

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:

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

A "Random Intercept" Model for the Variance

Total Predicted Data Matrix is called V Matrix, of which 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 obs**
n **= # occasions (5 here)**

<u>Level 2, BP Variance</u>

Unstructured **G Matrix** (**RANDOM** statement)

Each person has same **1 x 1 G** matrix (no covariance across persons in two-level model)

1 Random Intercept Variance only

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

To be added to **R** in order 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): $\mathbf{V}_i = \mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i^T + \mathbf{R}_i$

<u>Level 1, WP Variance</u>

Diagonal (VC) **R Matrix** (**REPEATED** statement)

Each person has same ***n* x *n* R** matrix → **equal variances and 0 covariances** across time (and no covariance across persons)

1 Residual Variance only

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

Intraclass Correlation (ICC)

ICCs are frequently used for longitudinal data:

$$\mathrm{ICC} = \frac{\mathrm{BP}}{\mathrm{BP}+\mathrm{WP}} = \frac{\text{Intercept Var.}}{\text{Intercept Var.} + \text{Residual Var.}} = \frac{\boldsymbol{\tau^2_{U_0}}}{\boldsymbol{\tau^2_{U_0}} + \boldsymbol{\sigma^2_e}}$$

$$\mathrm{Corr}(y_1,y_2) = \frac{\mathrm{Cov}(y_1,y_2)}{\sqrt{\mathrm{Var}(y_1)} * \sqrt{\mathrm{Var}(y_2)}}$$

V matrix **VCORR Matrix**

$$\begin{bmatrix} \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0} \\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_e \end{bmatrix} \begin{bmatrix} 1 & \mathrm{ICC} & \mathrm{ICC} & \mathrm{ICC} & \mathrm{ICC} \\ \mathrm{ICC} & 1 & \mathrm{ICC} & \mathrm{ICC} & \mathrm{ICC} \\ \mathrm{ICC} & \mathrm{ICC} & 1 & \mathrm{ICC} & \mathrm{ICC} \\ \mathrm{ICC} & \mathrm{ICC} & \mathrm{ICC} & 1 & \mathrm{ICC} \\ \mathrm{ICC} & \mathrm{ICC} & \mathrm{ICC} & \mathrm{ICC} & 1 \end{bmatrix}$$

- ICC = Proportion of total variance that is **between persons**
- ICC = **Correlation of occasions** from same person (in VCORR)
- ICC is a standardized way to express *dependency due to person mean differences* → **effect size for <u>constant</u> person dependency**

Review of Concepts and Terminology in Longitudinal Modeling

- Topics:
  - Concepts and terminology in longitudinal models (and their estimation in current software)
  - **Fixed and random effects of time (and their analog through latent variables)**
  - Significance testing and effect size for fixed and random effects (in MLM or SEM)
  - Modeling time-invariant predictors

Augmenting the Empty Means, Random Intercept model with **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):

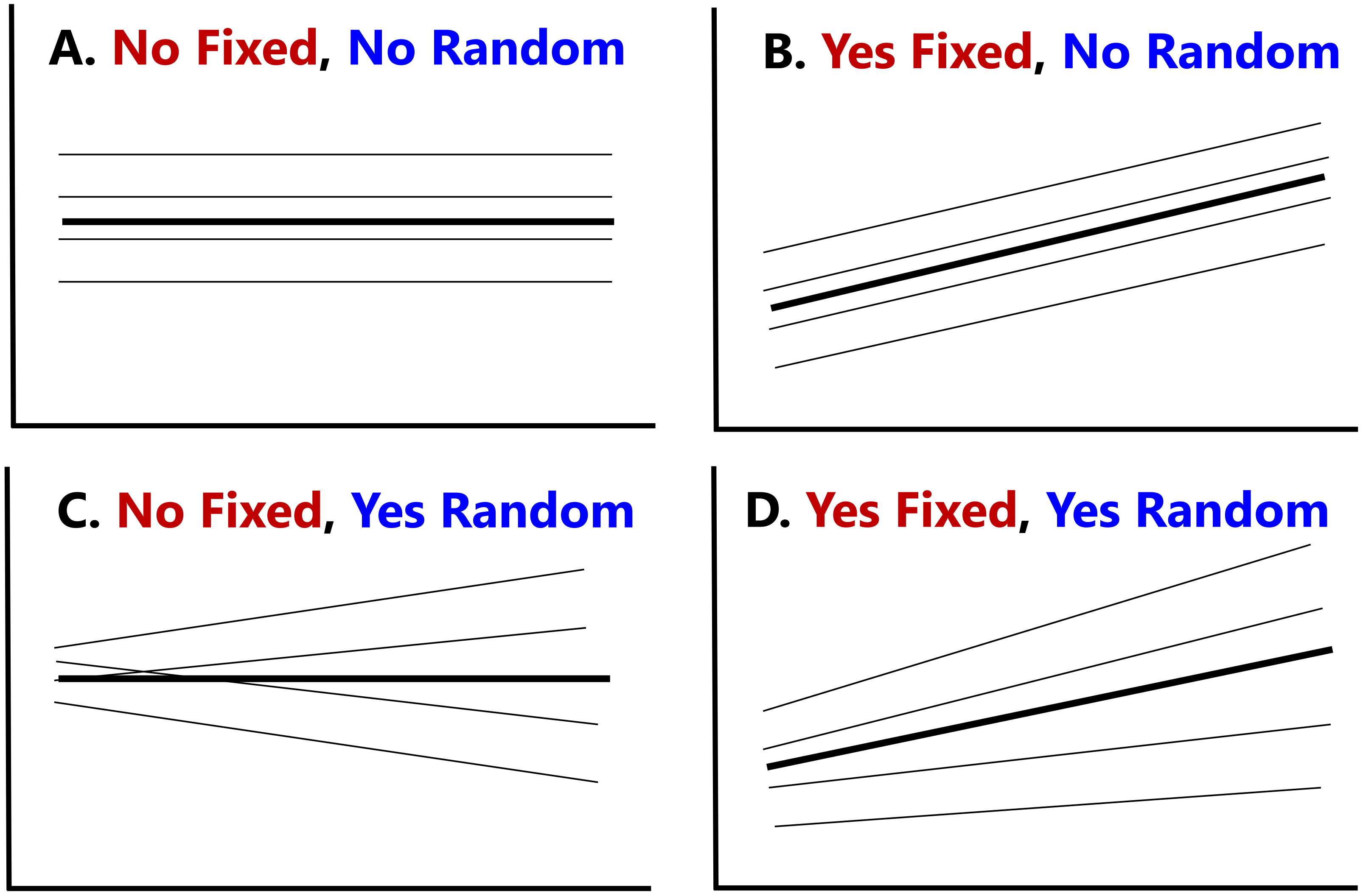
1. **Is there an effect of time on average?**
   - Is the line connecting the sample means over time something other than flat?
   - If so, you need **FIXED** effect(s) of time

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

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

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

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.

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

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

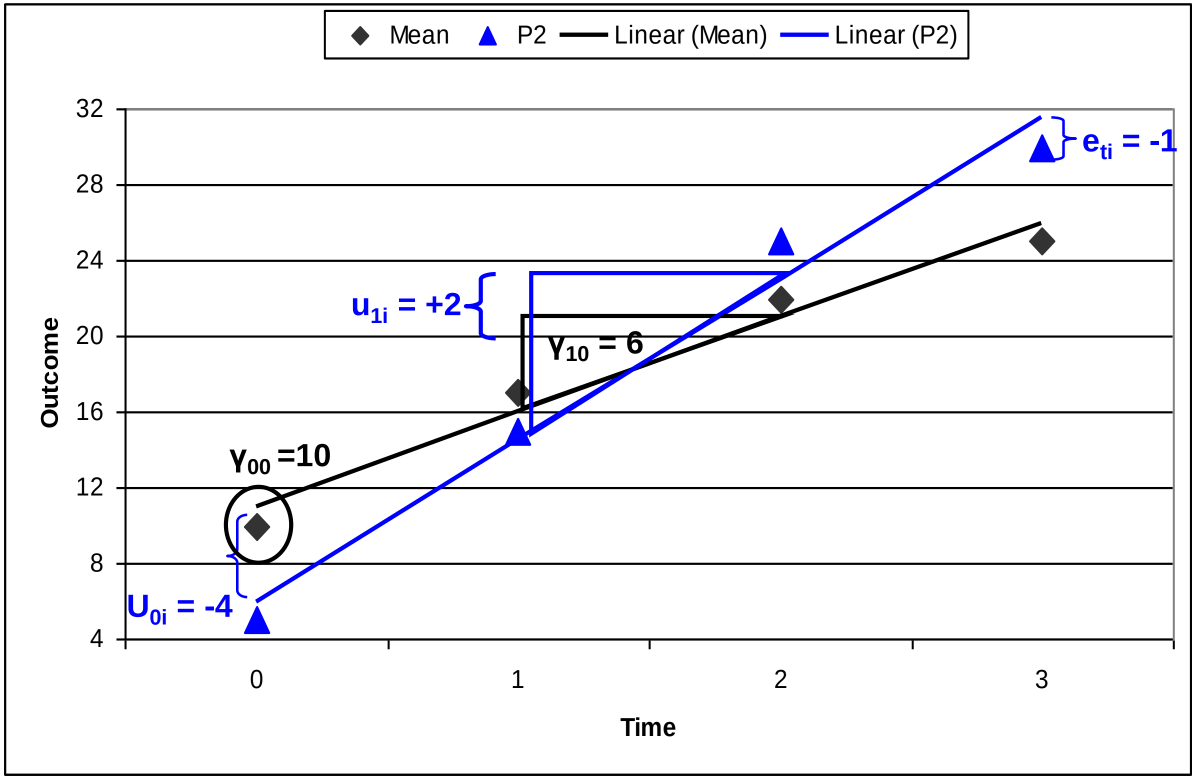
Fixed Intercept

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 $= \tau^2_{U0}$

U_{1i} Slope Variance $= \tau^2_{U1}$

Int-Slope Covariance $= \tau_{U\,01}$

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

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 (or else a different **R** matrix is needed):

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 for each person in **Z** (flexible for unbalanced time)

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

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

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

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

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

$\mathbf{Z}_i$ = 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 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

Translating Variance Partitioning in MLM into Single-Level SEM

- "**Random effects**" = "pile of variance" = "variance components"
  - Random effects represent **person*something interaction terms** that create person-caused sources of covariance over time
  - Random intercept → person*intercept (person "main effect")
  - Random linear time slope → person*time interaction

- Random effects are the same thing as **latent variables**
  - Latent variable = unobservable ability or trait, created by sources of **common variance** across items (or time-specific outcomes here)
  - Latent variables for BP differences can be interpreted as "general tendency" (random intercept) and "propensity to change" (random time slope)
  - Model-based way of de-trending longitudinal outcomes to distinguish BP from WP sources of information (and examine all kinds of relations)
  - Uses "wide" data structure in which each occasion = separate variable

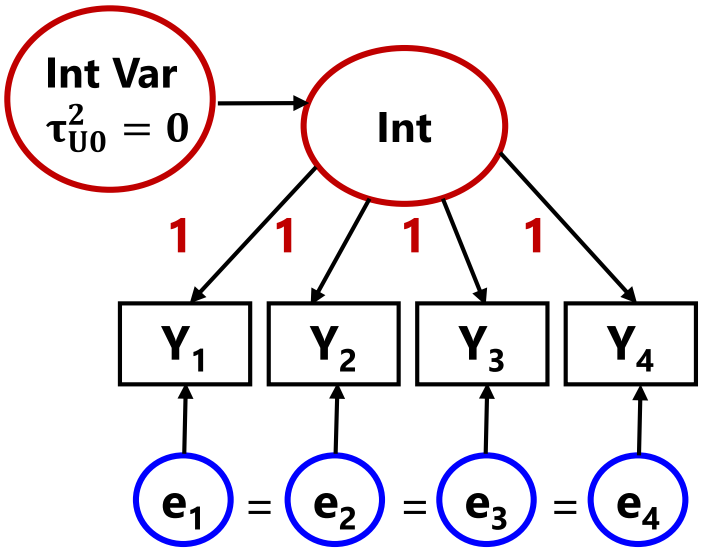
MLM as seen through Confirmatory Factor Analysis (CFA)

- **CFA model: $y_{is} = \mu_i + \lambda_i F_s + e_{is}$** (SEM is just relations among F's)
  - Observed response for item *i* (→ outcome at time *t*) and subject *s*
    = intercept of item *i* (μ)
    + subject *s*'s latent trait/factor (F), item-weighted by λ
    + error (e) of item *i* and subject *s*

- Four big differences when using CFA/SEM for longitudinal change:
  - Usually two factors for "level" and "change" (intercept and slope):
    $y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})time_{ti} + e_{ti}$ → so the U's are the F's
  - The separate **item (time-specific outcome) intercepts μ_i** cannot be identified separately from the "intercept" factor and therefore **must be fixed to 0**
  - The **factor loadings λ_i** for how each outcome is predicted by the latent factor are usually pre-determined by **how much time as passed**, and are fixed to the difference in time that corresponds to the **type of change** (e.g., linear, quadratic, piecewise)
  - Item (time-specific outcome) **residual variances should be constrained equal** (not default, but changes in variance over time should be captured by random slopes)

Random Effects as Latent Variables

- **BP-only model: e_{ti} = model for the variance**
  - $y_{ti} = \gamma_{00} + e_{ti}$



Mean of the intercept factor = fixed intercept γ_{00}

Loadings of intercept factor = 1 (all occasions contribute equally)

Item intercepts = 0 (always)

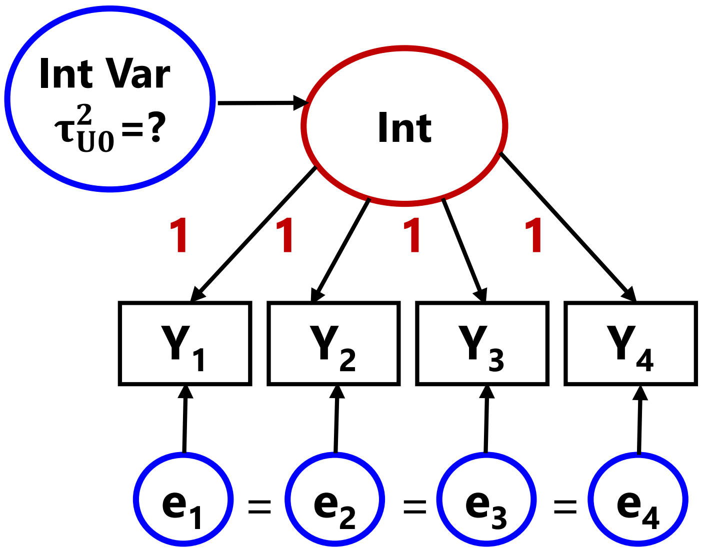
Variance of intercept factor = 0 so far

Residual variance (e) is assumed to be equal across occasions

  - After controlling for the *fixed* intercept, residuals are assumed uncorrelated: **this is a single-level model**

Random Effects as Latent Variables

- **+WP model: U_{0i} + e_{ti} model for the variance**
  - **$y_{ti} = \gamma_{00} + U_{0i} + e_{ti}$**



Mean of the intercept factor = fixed intercept γ_{00}

Loadings of intercept factor= 1 (all occasions contribute equally)

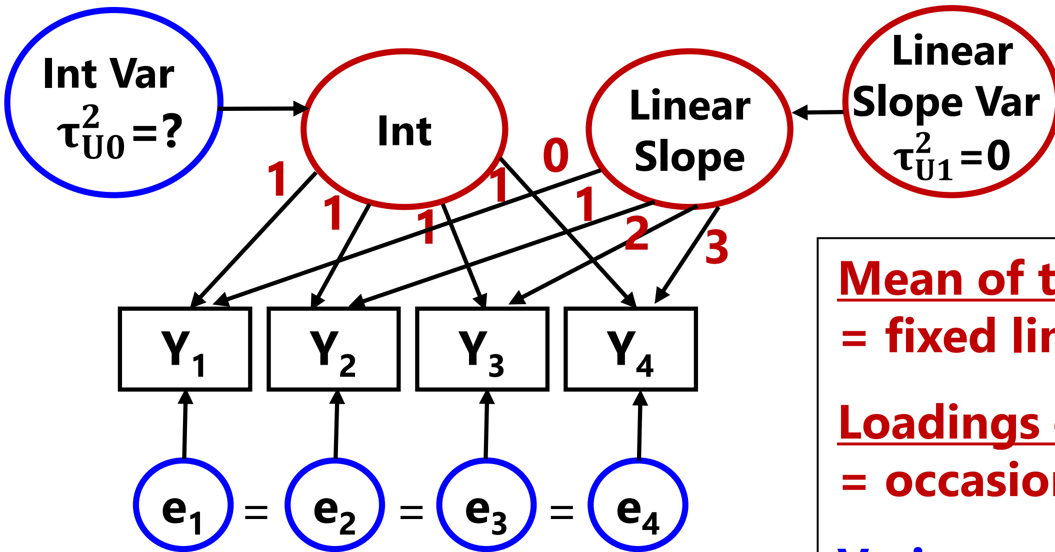
Variance of intercept factor = random intercept variance

Residual variance (e) is assumed to be equal across occasions

  - After controlling for the *random* intercept, residuals are assumed uncorrelated: **now two piles of variance** (aka, an **"empty means, random intercept" model**)

Random Effects as Latent Variables

- **Fixed linear time, random intercept model:**
  - $y_{ti} = \gamma_{00} + (\gamma_{10}\text{Time}_{ti}) + U_{0i} + e_{ti}$



Mean of the linear slope factor
= fixed linear slope γ_{10}

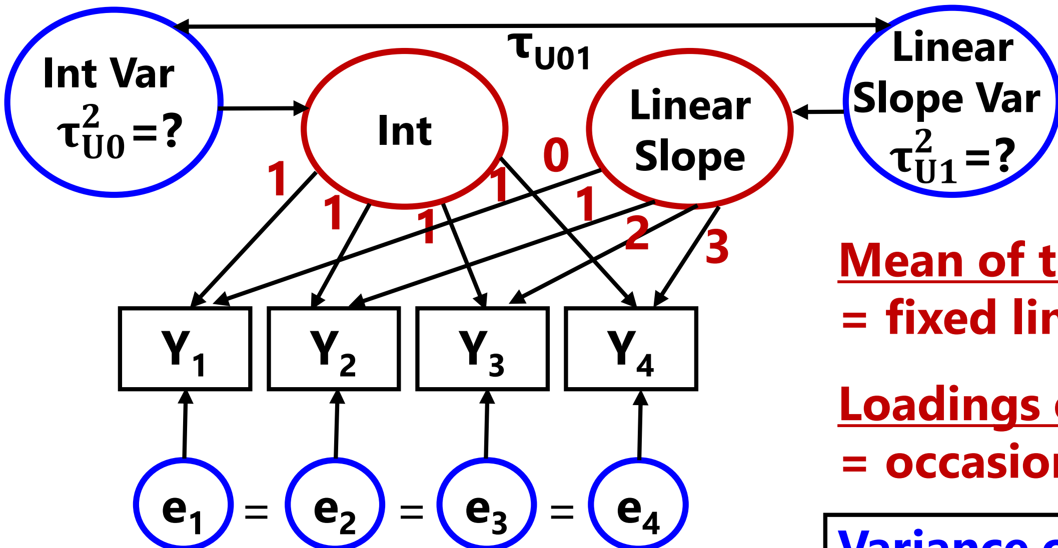
Loadings of linear slope factor
= occasions (keep real time)

Variance of linear slope factor
= 0

  - After controlling for the *fixed linear slope and random intercept*, residuals are assumed uncorrelated

Random Effects as Latent Variables

- **Random linear model:**
  - $y_{ti} = \gamma_{00} + (\gamma_{10}\text{Time}_{ti}) + U_{0i} + (U_{1i}\text{Time}_{ti}) + e_{ti}$



Mean of the linear slope factor = fixed linear slope γ_{10}

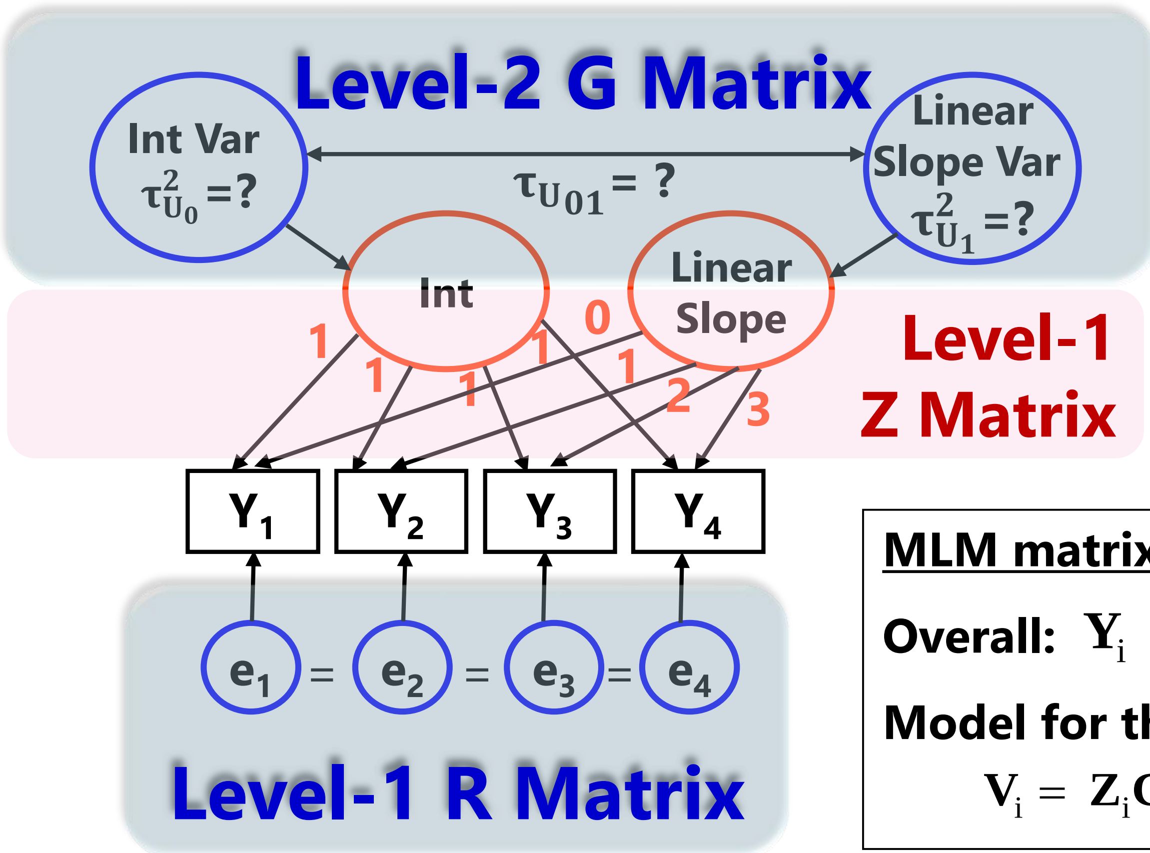
Loadings of linear slope factor = occasions (keep real time)

Variance of linear slope factor = random slope variance (and covariance with random intercept

  - After controlling for the *random linear slope and random intercept*, residuals are assumed uncorrelated: **now three piles of variance** to be predicted (BP int, BP slope, WP res)

Summary: Random Linear Time Model as Latent Variables in SEM

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



MLM matrix version of model

Overall: $\mathbf{Y}_i = \mathbf{X}_i\gamma + \mathbf{Z}_i\mathbf{U}_i + \mathbf{E}_i$

Model for the Variance:

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

Review of Concepts and Terminology in Longitudinal Modeling

- Topics:
  - Concepts and terminology in longitudinal models (and their estimation in current software)
  - Fixed and random effects of time (and their analog through latent variables)
  - **Significance testing and effect size for fixed and random effects (in MLM or SEM)**
  - Modeling time-invariant predictors

Assessing the "Goodness" of the Model

- **Model for the Means** → which **fixed effects** of predictors should be included in the model (e.g., main effects, interactions)
  - **Significance tests** do not require assessment of relative model fit using LL or −2LL (can always use univariate or multivariate Wald tests)
  - **Effect sizes** can come from the significance tests (e.g., F → Cohen's d), or from reductions in variance (pseudo-R^2 or total-R^2)

- **Model for the Variance** → what pattern(s) of variance and covariance the residuals from the same unit have; what **random effects** are needed to describe these pattern(s)
  - **Significance tests** DO require assessing relative model fit via $-2\Delta LL$
    - Cannot use the Wald test *p*-values for variances on the output because those *p*-values use a two-sided sampling distribution for what the variance could be (but variances cannot be negative, so those *p*-values are not valid)
  - **Effect sizes** (less commonly provided) can come from random effects confidence intervals (CI) or random effects reliability measures
    - $\text{Random Effect 95\% CI} = \text{fixed effect} \pm \left(1.96 * \sqrt{\text{Random Variance}}\right)$

Statistical Significance of Fixed Effects: What letter will I get?

In ML or REML, fixed effects can be tested via **Wald** tests: the ratio of its estimate/SE forms a statistic we compare to a distribution

	Denominator DF is assumed infinite	Denominator DF is estimated → "Modified"
Numerator DF = 1 *(test one fixed effect)* is **Univariate Wald Test**	use ***z*** distribution (e.g., Mplus, STATA)	use ***t*** distribution (e.g., SAS, SPSS)
Numerator DF > 1 *(test 2+ fixed effects)* is **Multivariate Wald Test**	use $\boldsymbol{\chi^2}$ distribution (e.g., Mplus, STATA)	use ***F*** distribution (e.g., SAS, SPSS)
Denominator DF options (important in small level-2 samples)	not applicable, so DDF is not given	SAS, SPSS, and STATA 14: **Satterthwaite** SAS and Stata 14: Between-Within, Kenward-Roger (best)

Pseudo-R^2 Effect Size of Fixed Effects

- Pseudo-R^2 = proportion of variance accounted for by fixed effects of predictors **in each pile of variance → multiple pseudo-R^2 values**
- For example, a fixed linear effect of WP time will reduce level-1 residual variance σ_e^2 in **R** by this much:

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

$$\text{More generally, Pseudo } R^2 = \frac{\text{was} - \text{is}}{\text{was}}$$

"**fewer**" = "**was**" = from model with fewer parameters
"**more**" = "**is**" = from model with more parameters

- But whenever <u>only</u> level-1 residual variance σ_e^2 is reduced, the level-2 random intercept variance $\tau_{U_0}^2$ will <u>INCREASE</u> as a result. Why?
  - Likelihood-based estimates of "true" $\tau_{U_0}^2$ use (σ_e^2 / level-1 *n*) as correction factor for the amount of between-person difference attributable to chance:
    True $\tau_{U_0}^2$ = Observed $\tau_{U_0}^2$ – (σ_e^2 / level-1 *n*)
  - For example: observed level-2 $\tau_{U_0}^2$=4.65, level-1 σ_e^2=7.06, *n*=4
    - True $\tau_{U_0}^2$= 4.65 –(**7.60**/4) = **2.88** in empty means model
    - Add fixed linear time slope → reduce σ_e^2 from 7.06 to 2.17 (Pseudo-R^2 = .69)
    - But now True $\tau_{U_0}^2$= 4.65 –(**2.17**/4) = **4.10** in fixed linear time model

Variance Accounted for by Fixed Effects For Real: Total-R^2

- **Pseudo-R^2** is named that way for a reason… piles of variance can shift around, such that it can be negative or explained for no reason
  - Sometimes a sign of model mis-specification (but not always)
  - Usually hard to explain when it happens!

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

- Because R^2 does not always mean the same thing, specify which kind of R^2 you used—provide the formula and a reference!

Significance Tests for Choosing Amongst Models for the Variance

- Requires assessment of **relative model fit**: how well does the model fit relative to other possible models?
  - Assessment of *absolute* model fit is only possible for balanced data

- Relative fit is indexed by overall model **log-likelihood** (**LL**):
  - Log of likelihood for each person's outcomes given model parameters
  - Sum log-likelihoods across all independent persons = **model LL**
  - Two flavors: Maximum Likelihood (ML) or Restricted ML (REML)

- What you get for this on your output varies by software…

- Given as −2*log likelihood (−2LL) in SAS or SPSS MIXED:
  −2LL gives BADNESS of fit, so **smaller** value = better model

- Given as just log-likelihood (LL) in STATA MIXED and Mplus:
  LL gives GOODNESS of fit, so **bigger** value = better model

Comparing Models for the Variance

- **Two strategies for choosing a model for the variance:**
  - Does the more complex model fit better (than a simpler model)?
  - Does the simpler model fit worse (than a more complex model)?

- Nested models are compared using a **"likelihood ratio test": −2ΔLL test** (aka, "χ^2 test" in SEM; "deviance difference test" in MLM)

"fewer" = from model with fewer parameters
"more" = from model with more parameters

Results of 1. & 2. must be positive values!

1. Calculate **−2ΔLL:** if given −2LL, do $-2\Delta LL = (-2LL_{fewer}) - (-2LL_{more})$
   if given LL, do $-2\Delta LL = -2 * (LL_{fewer} - LL_{more})$
2. Calculate **Δdf** = $(\#\ Parms_{more}) - (\#\ Parms_{fewer})$
3. **Compare −2ΔLL to χ^2 distribution with df = Δdf**
4. Get *p*-value from CHIDIST in Excel or LRTEST option in STATA

Comparing Models for the Variance

- <u>What your *p*-value for the −2ΔLL test means:</u>
  - If you **ADD** parameters, then your model can get **better** (if −2ΔLL test is significant) or **not better** (not significant)
  - If you **REMOVE** parameters, then your model can get **worse** (if −2ΔLL test is significant) or **not worse** (not significant)

- Nested or non-nested models can also be compared by **Information Criteria** that also reflect model parsimony
  - No significance tests or critical values, just "smaller is better"
  - **AIC** = Akaike IC = −2LL + 2 *(#parameters)
  - **BIC** = Bayesian IC = −2LL + log(N)*(#parameters)
  - What "parameters" means depends on flavor (except in STATA):
    - ML = ALL parameters; REML = variance model parameters only

Flavors of Maximum Likelihood

- Remember that likelihood estimation comes in two flavors:
- **"Restricted (or residual) maximum likelihood"**
  - Only available for general linear models or general linear mixed models (that assume normally distributed residuals); not in SEM software
  - Is same as LS given complete outcomes, but it doesn't require them
  - Estimates variances the same way as in LS (accurate) → $\frac{\sum(y_i - y_{pred})^2}{N-k}$
- **"Maximum likelihood" (ML; also called FIML*)**
  - Is more general, is available for the above plus for non-normal outcomes and latent variable models (CFA/SEM/IRT; multilevel SEM)
  - Is NOT the same as LS: it under-estimates variances by not accounting for the # of estimated fixed effects → $\frac{\sum(y_i - y_{pred})^2}{N}$
- **FI = Full information → it uses all original data (they both do)*

ML vs. REML Estimation in a Nutshell

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

"Population"

$$\frac{\Sigma(y_i - y_{pred})^2}{N}$$

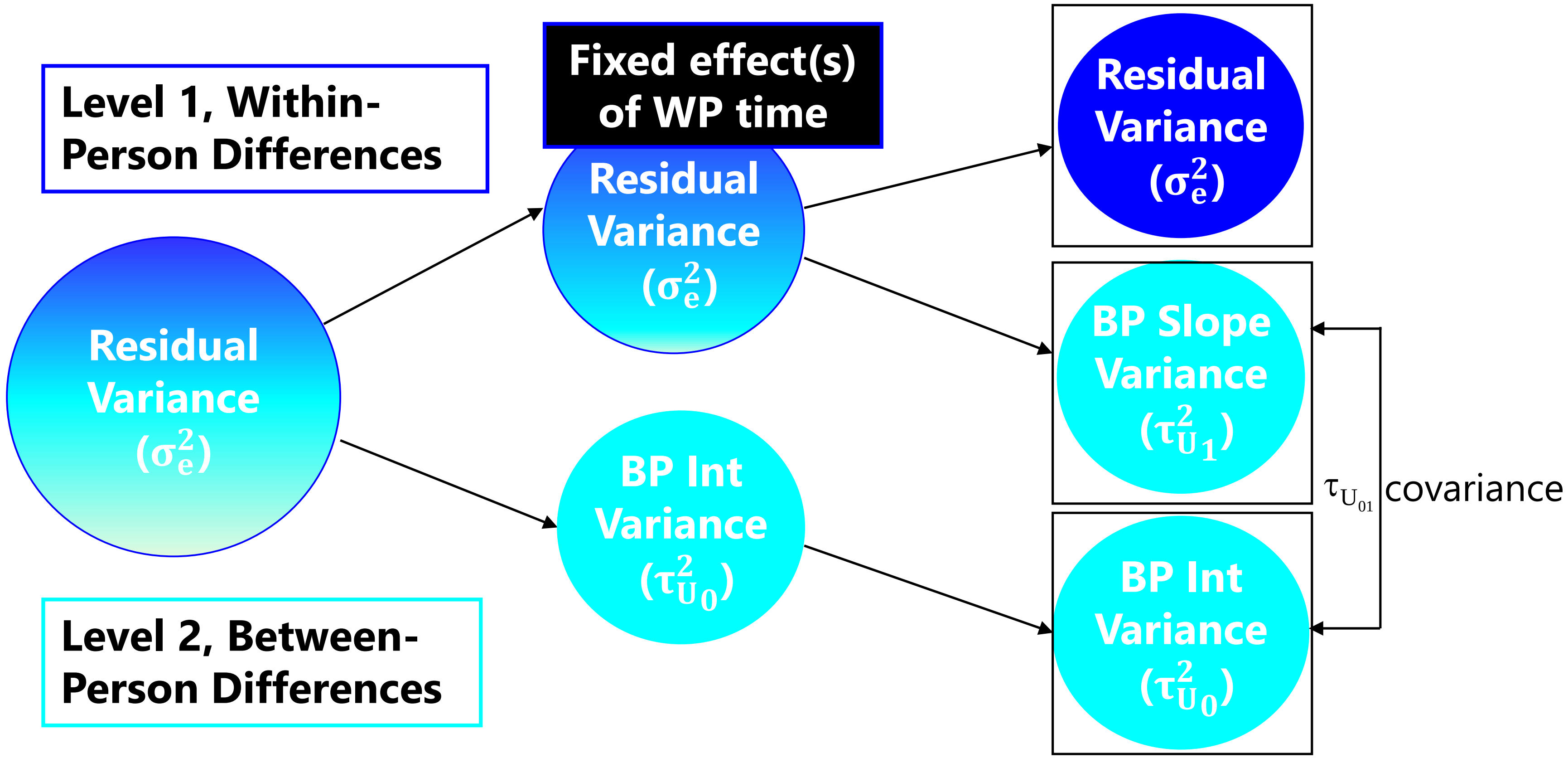
"Sample"

$$\frac{\Sigma(y_i - y_{pred})^2}{N - k}$$

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

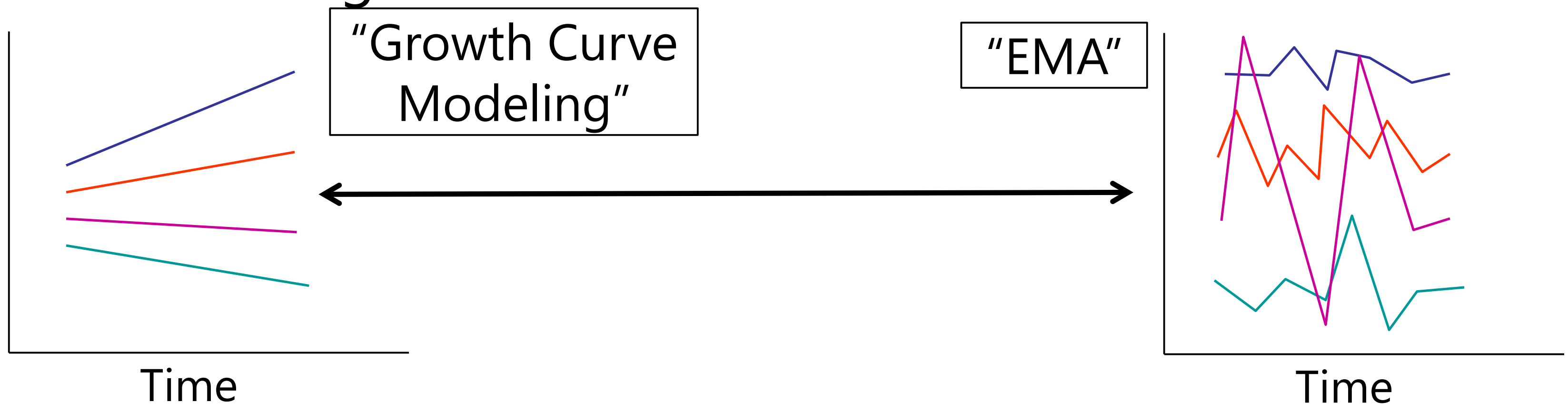
Summary: Unconditional Models for Time

- The process of fitting "unconditional models for time" (fixed and random effects) can be depicted as follows:



Summary: Unconditional Models for Time

Pure WP Change Pure WP Fluctuation



<u>Role of "Time" in the Model for the Means:</u>

- WP Change → describe pattern of ***average*** change (e.g., growth curves)
- WP Fluctuation → describe ***average*** time-specific trends that may not have been expected (e.g., reactivity, day of the week, circadian/schedule effects)

<u>Role of "Time" in the Model for the Variance:</u>

- WP Change → describe ***individual differences*** in change (random effects)
  → this allows variances and covariances to differ over time
- WP Fluctuation → mostly describe pattern(s) of covariance over time (may need random effects of time for differing variances)

Summary: Unconditional Models for Time

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

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 ($\tau_{U_1}^2$)

Multiple BP time slope variances are possible…

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

Next we will add predictors to account for each pile!

Review of Concepts and Terminology in Longitudinal Modeling

- Topics:
  - Concepts and terminology in longitudinal models (and their estimation in current software)
  - Fixed and random effects of time (and their analog through latent variables)
  - Significance testing and effect size for fixed and random effects (in MLM or SEM)
  - **Modeling time-invariant predictors**

Beware of Missing Predictors

- Any cases missing model predictors (that are not part of the likelihood*) <u>will not be used in that model</u>
  - Less than ideal for time or time-varying predictors (MARish)
  - Really bad for time-invariant predictors (listwise deletion, MCAR)

- Other options for missing predictors:
  - *Bring the predictor into the likelihood (only possible in software for multivariate MLMs, such as Mplus, or in SEM programs)
    - Its mean, variance, and covariances "get found" as model parameters
    - Predictor then has distributional assumptions (default is multivariate normal), which may not be plausible for all predictors
    - Mplus v. 8 still will not do this for non-normal "predictors" in multivariate MLM
  - Multiple imputation (and analysis of *each* imputed dataset)
    - Imputation also requires distributional assumptions for imputed variables!
    - Also requires all parameters of interest for the analysis model to be in the imputation model, too (which is problematic for interactions or random effects)

Modeling Time-Invariant Predictors

- Which independent variables can be time-invariant predictors?
  - Also known as "**person-level**" or "**level-2**" or "**macro**" predictors
  - Includes substantive predictors, controls, and predictors of missingness
  - Includes anything that either **does not change across time**, or that might change across time but that **you've only measured once** (you may need to argue why this is conceptually ok or limit conclusions accordingly)
  - Also includes **BP variance in time or time-varying predictors** (stay tuned)

- All predictors should be **centered** so that 0 values are meaningful:
  - This is needed to create a meaningful fixed/random intercept, and/or meaningful fixed main effects of predictors also included in interactions
    - e.g., if fixed effects of X, Z, and X*Z, the main effect of X is specifically for Z=0
  - **Continuous** predictors can be **centered at any constant**, such as the sample mean (common and useful if it has an unfamiliar scale) or any meaningful reference (better for translating across studies)
  - **Categorical** predictors can have their **dummy-code contrasts** created for you in most programs (e.g., SAS CLASS, SPSS BY, STATA i.), but not in Mplus; I do not like ± 1 coding for group differences (because then 0 = ???)

Variance Accounted For By Level-2, Person-Level, Time-Invariant Predictors

- **Fixed effects of level-2 predictors *by themselves*:**
  - Level-2 (BP) main effects reduce L2 (BP) random intercept variance
  - Level-2 (BP) interactions also reduce L2 (BP) random intercept variance
- **Fixed effects of *cross-level interactions* (level 1* level 2):**
  - **Always test the level-2 random slope for the corresponding level-1 predictor BEFORE fitting any cross-level interactions for it!**
  - If the interacting level-1 predictor is <u>random</u>, any cross-level interaction with it will reduce its corresponding level-2 BP **random slope variance**
    - e.g., if L1 *time* is random, then group**time,* ed**time*, and group*ed**time* can each reduce the L2 random linear time slope variance
  - If the interacting level-1 predictor <u>not random</u>, cross-level interactions with it will reduce the level-1 WP **residual variance** instead
    - The effect of time is then called ***systematically varying***: it's not *fixed* (the same for everybody) or *random* (different for everybody), but a middle ground—each person gets their own slope as a function of the L2 predictor(s)

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

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

Level 2 Equations (one per β):

Composite Equation:

$$y_{ti} = (\gamma_{00} + \gamma_{01}Ed_i + U_{0i}) + (\gamma_{10} + \gamma_{11}Ed_i + U_{1i})Time_{ti} + e_{ti}$$

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

- β_{0i}: **Intercept** for person *i*
- γ_{00}: **Fixed Intercept when Time=0 and Ed=12**
- γ_{01}: **Δ in Intercept per unit Δ in Ed**
- U_{0i}: **Random (Deviation) Intercept after controlling for Ed**

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

- β_{1i}: **Linear Slope** for person *i*
- γ_{10}: **Fixed Linear Time Slope when Ed=12**
- γ_{11}: **Δ in Linear Time Slope per unit Δ in Ed (=Ed*time)**
- U_{1i}: **Random (Deviation) Linear Time Slope after controlling for Ed**

We would calculate **pseudo-R^2 values** as follows:

- For the L2 Ed main effect **(γ_{01})**, using the L2 random intercept **U_{0i} variance**
- For the L2 Ed by L1 Time cross-level interaction **(γ_{11})**:
  - Using the L2 random time slope **U_{1i} variance** if present in the model (as it is here)
  - Otherwise using the L1 residual **e_{ti} variance** instead (so time → "systematically varying")

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

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

- β_{0i}: **Intercept** for person *i*
- γ_{00}: **Fixed Intercept when Time=0 and Ed=12**
- γ_{01}: **Δ in Intercept per unit Δ in Ed**
- U_{0i}: **Random (Deviation) Intercept after controlling for Ed**

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

- β_{1i}: **Linear Slope** for person *i*
- γ_{10}: **Fixed Linear Time Slope when Time=0 and Ed=12**
- γ_{11}: **Δ in Linear Time Slope per unit Δ in Ed (=Ed*time)**
- U_{1i}: **Random (Deviation) Linear Time Slope after controlling for Ed**

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

- β_{2i}: **Quad Slope** for person *i*
- γ_{20}: **Fixed Quad Time Slope when Ed = 12**
- γ_{21}: **Δ in Quad Time Slope per unit Δ in Ed (=Ed*time²)**
- U_{2i}: **Random (Deviation) Quad Time Slope after controlling for Ed**

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

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

Level 2 Equations (one per β):

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

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

Each fixed effect of education will predict the U in its equation if present, or residual variance otherwise.

- Composite equation:
- $\mathbf{y_{ti} = (\gamma_{00} + \gamma_{01}Ed_i + U_{0i}) + (\gamma_{10} + \gamma_{11}Ed_i + U_{1i})Time_{ti} + (\gamma_{20} + \gamma_{21}Ed_i + U_{2i})Time_{ti}^2 + e_{ti}}$

Summary: Time-Invariant Predictors

- Univariate MLMs use ONLY COMPLETE rows (occasions) of data—both predictors and outcomes must be there!
  - Using whatever data you do have for each person will likely lead to better inferences and more statistical power than using only complete persons (so avoid listwise deletion if you can)
  - Missingness on predictors is possible if they are "outcomes" in multivariate software, which implies distributional assumptions

- Time-invariant predictors modify the level-1 created time model → predict individual intercepts and slopes
  - They account for random effect variances (the predictors are the reasons WHY people need their own intercepts and slopes)
  - If a level-1 effect is not random, it can still be moderated by a cross-level interaction with a time-invariant predictor…
    - … but then it will predict L1 residual variance instead