

# Time-Invariant Predictors in Longitudinal Models

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

# Summary of Steps in Unconditional Longitudinal Modeling

## **For all outcomes:**

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

## **If your outcome shows systematic change:**

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

## **If your outcome does NOT show ANY systematic change:**

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

# I. Empty Means, Random Intercept Model

- Not really predictive, but is a useful statistical baseline model
  - Baseline model fit
  - Partitions variance into between- and within-person variance
- Calculate **ICC** = between / (between + within variance)
  - = Average correlation between occasions
  - = Proportion of variance that is between persons
  - Effect size for amount of person dependency due to mean differences
- Tells you where the action will be:
  - If most of the variance is **between-persons in the random intercept (at level 2)**, you will need **person-level** predictors to reduce that variance (i.e., to account for inter-individual differences)
  - If most of the variance is **within-persons in the residual(at level 1)**, you will need **time-level** predictors to reduce that variance (i.e., to account for intra-individual differences)

## 2. Decide on the Metric of Time

- "Occasion of Study" as Time:
  - Can be used generically for many purposes
  - Include age, time to event as predictors of change
- "Age" as Time:
  - Is equivalent to time-in-study if same age at beginning of study
  - Implies age convergence → that people only differ in age regardless of when they came into the study (BP effects = WP effects)
- "Distance to/from an Event" as Time:
  - Is appropriate if a distinct process is responsible for changes
  - Also implies convergence (BP effects = WP effects)
  - Only includes people that have experienced the event
- Make sure to use exact time regardless of which "time" used

# 3. Decide on a Centering Point

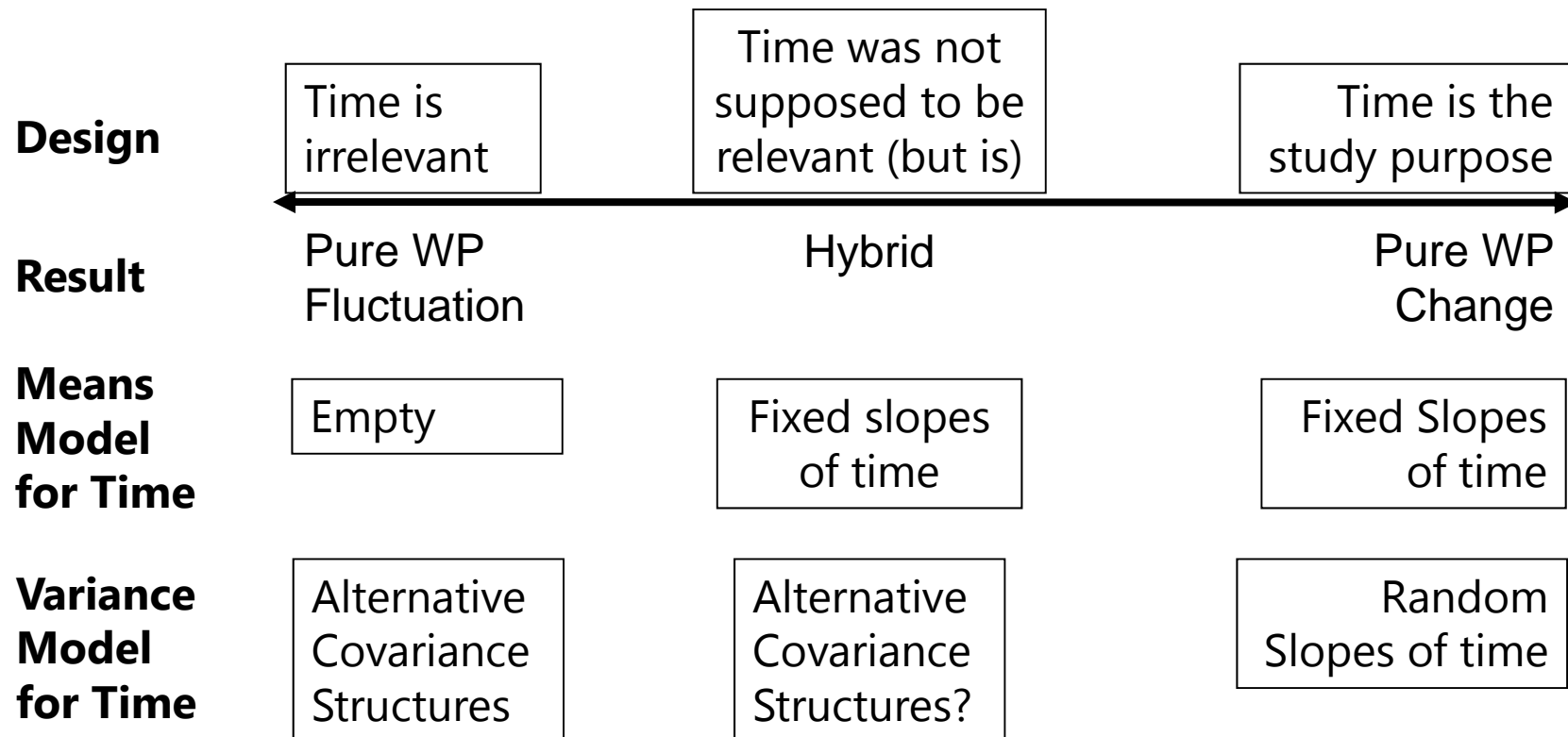
- How to choose: At what occasion would you like a snap-shot of inter-individual differences?
  - Intercept variance represents inter-individual differences at that particular time point (that you can later predict!)
- Where do you want your intercept?
  - Re-code time such that the centering point = 0
  - Multiple variants could be used (e.g., moving snapshots)
- Different versions of time = 0 will produce statistically equivalent models with re-arranged parameters
  - i.e., conditional level and rate of change at time 0

# 4. Plot Saturated Means and Individuals

- If time is balanced across persons:
  - Estimate a saturated means model to generate means
- If time is NOT balanced across persons:
  - Create a rounded time variable to estimate means model ONLY
  - Still use exact time/age variable for analysis!
- Plot the means – what kind of trajectory do you see?
- Please note: ML/REML estimated means per occasion may NOT be the same as the observed means (i.e., as given by PROC MEANS). The estimated means are what would have been obtained *had your data been complete* (assuming MAR), whereas observed means are not adjusted to reflect any missing data (MCAR). Report the ML/REML estimated means.

# What if I have no change?

- Longitudinal studies are not always designed to examine systematic change (e.g., daily diary studies)
- In reality, there is a continuum of fluctuation to change:



## 5. and 6. for **Systematic Change**: Evaluate Fixed and Random Effects of Time

### **Model for the Means:**

- What kind of fixed effects of time are needed to parsimoniously represent the observed means across time points?
  - Linear or nonlinear? Continuous or discontinuous?
  - Polynomials? Pieces? Nonlinear curves?
  - How many parameters do you need to Name That Trajectory?
  - Use obtained  $p$ -values to test significance of fixed effects

### **Model for the Variance (focus primarily on G):**

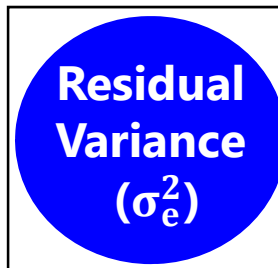
- What kind of random effects of time are needed:
  - To account for individual differences in aspects of change?
  - To describe the variances and covariances over time?
  - Do the residuals show any pattern after accounting for random effects?
  - Use REML  $-2\Delta LL$  tests to test significance of new effects (or ML if big  $N$ )



# Random Effects Models for the Variance

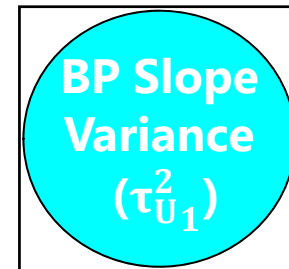
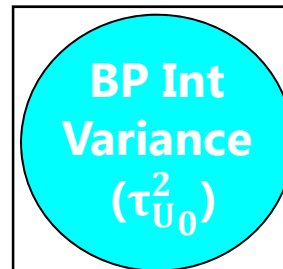
- Each source of correlation or dependency goes into a new variance component (or pile of variance) until each source meets the usual assumptions of GLM: normality, independence, constant variance
- Example 2-level longitudinal model:**

**Level 1 (one source of)**  
**Within-Person Variation:**  
gets accounted for by  
time-level predictors



**FIXED** effects make variance go away (explain variance).  
**RANDOM** effects just make a new pile of variance.

**Level 2 (two sources of)**  
**Between-Person Variation:**  
gets accounted for by  
person-level predictors



↑  $\tau_{U_{01}}$  covariance ↑

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

## 5. for **NO Systematic Change**: Evaluate Alternative Covariance Structures

### Model for the Means:

- Be sure you don't need any terms for systematic effects of time
- If not, keep a fixed intercept only

### Model for the Variance (focus primarily on **R**):

- How many parameters do you need to Name... that... Structure?
- I recommend the hybrid: Random Intercept in **G** + Structure in **R**
  - Separates BP and WP variance
  - Likely more parsimonious than just **R**-only model
- Compare alternative models with the same fixed effects
  - Nested? REML  $-2\Delta LL$  test for significance
  - Non-nested? REML AIC and BIC for "supporting evidence"

# Alternative Covariance Structure Models

- Models for fluctuation typically include only a covariance structure, and at most a random intercept (random slopes for time won't help in the absence of systematic change)

**Between-Person Random Intercept in G + Within-Person Structure in R**

**Level 1 (one source of) Within-Person Variation:**

**Gets accounted for by time-level predictors**

**Residual  
Variance  
( $\sigma_e^2$ )**

**Level 2 (one sources of) Between-Person Variation:**

**Gets accounted for by person-level predictors**

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

**TOTAL Structure in R**

**All sources of variation and covariation are held in one matrix, but if dependency is predicted accurately then it's ok.**

**Total  
Variance  
( $\sigma_T^2$ )**

# Why spend so much effort on unconditional models of time? Here is the reasoning...

- The fixed effects of time are what the random effects of time are varying around...
- The random effects of time form the variances that the person-level predictors will account for...
- The effects of person-level predictors are specified as a function of the time effect already in the model...
- The effects of time-varying predictors are supposed to account for variance not accounted for by the model for time...
- What fixed and random time effects of time you include in the model dictate what is to be predicted. **Get time right first!**

# Time-Invariant Predictors in Longitudinal Models

- Topics:
  - Summary of building unconditional models for time
  - **Missing predictors in MLM**
  - **Centering predictors and interpreting interactions**
  - Effects of time-invariant predictors
  - Fixed, systematically varying, and random level-1 effects
  - Model building strategies and assessing significance

# Missing Data in MLM Software

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

# Be Careful of Missing Predictors!

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

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

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

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

Model with Time → DV: 1-6, 8

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

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

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

# So what does this mean for missing data in MLM?

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



# Centering Predictors

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

# Main Effects of Predictors within Interactions

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

# Interactions: Why 0 Matters

- Y = Student achievement (GPA as percentage out of 100)  
X = Parent attitudes about education (measured on 1-5 scale)  
Z = Father's education level (measured in years of education)
- $$\text{GPA}_i = \beta_0 + (\beta_1 * \text{Att}_i) + (\beta_2 * \text{Ed}_i) + (\beta_3 * \text{Att}_i * \text{Ed}_i) + e_i$$
$$\text{GPA}_i = 30 + (1 * \text{Att}_i) + (2 * \text{Ed}_i) + (0.5 * \text{Att}_i * \text{Ed}_i) + e_i$$
- Interpret  $\beta_0$ :
- Interpret  $\beta_1$ :
- Interpret  $\beta_2$ :
- Interpret  $\beta_3$ : **Attitude** as Moderator:  
**Education** as Moderator:
- **Predicted GPA** for **attitude of 3** and **Ed of 12**?  
$$75 = 30 + 1*(3) + 2*(12) + 0.5*(3)*(12)$$

# Model-Implied Simple Main Effects

- **Original:**  $GPA_i = \beta_0 + (\beta_1 * Att_i) + (\beta_2 * Ed_i) + (\beta_3 * Att_i * Ed_i) + e_i$   
 $GPA_i = 30 + (1 * Att_i) + (2 * Ed_i) + (0.5 * Att_i * Ed_i) + e_i$
- Given any values of the predictor variables, the model equation provides predictions for:
  - Value of outcome (model-implied intercept for non-zero predictor values)
  - Any conditional (simple) main effects implied by an interaction term
  - **Simple Main Effect = what it is + what *modifies* it**
- Step 1: **Identify** all terms in model involving the predictor of interest
  - e.g., Effect of Attitudes comes from:  $\beta_1 * Att_i + \beta_3 * Att_i * Ed_i$
- Step 2: **Factor out** common predictor variable
  - Start with  $[\beta_1 * Att_i + \beta_3 * Att_i * Ed_i] \rightarrow [Att_i (\beta_1 + \beta_3 * Ed_i)] \rightarrow Att_i$  (new  $\beta_1$ )
  - Value given by ( ) is then the model-implied coefficient for the predictor
- Step 3: **ESTIMATEs** calculate model-implied simple effect and SE
  - Let's try it for a new reference point of **attitude = 3** and **education = 12**

# Model-Implied Simple Main Effects

- Old Equation using uncentered predictors:

$$\text{GPA}_i = \beta_0 + (\beta_1 * \text{Att}_i) + (\beta_2 * \text{Ed}_i) + (\beta_3 * \text{Att}_i * \text{Ed}_i) + e_i$$

$$\text{GPA}_i = 30 + (1 * \text{Att}_i) + (2 * \text{Ed}_i) + (0.5 * \text{Att}_i * \text{Ed}_i) + e_i$$

- New equation using centered predictors ( $\text{Att}_i - 3$  and  $\text{Ed}_i - 12$ ):

$$\text{GPA}_i = \_ + \_ * (\text{Att}_i - 3) + \_ * (\text{Ed}_i - 12) + \_ * (\text{Att}_i - 3) * (\text{Ed}_i - 12) + e_i$$

- **Intercept: expected value of GPA when  $\text{Att}_i = 3$  and  $\text{Ed}_i = 12$**

$$\beta_0 = 75$$

- **Simple main effect of Att if  $\text{Ed}_i = 12$**

$$\beta_1 * \text{Att}_i + \beta_3 * \text{Att}_i * \text{Ed}_i \rightarrow \text{Att}_i(\beta_1 + \beta_3 * \text{Ed}_i) \rightarrow \text{Att}_i(1 + 0.5 * 12)$$

- **Simple main effect of Ed if  $\text{Att}_i = 3$**

$$\beta_2 * \text{Ed}_i + \beta_3 * \text{Att}_i * \text{Ed}_i \rightarrow \text{Ed}_i(\beta_2 + \beta_3 * \text{Att}_i) \rightarrow \text{Ed}_i(2 + 0.5 * 3)$$

- **Two-way interaction of Att and Ed:**

$$(0.5 * \text{Att}_i * \text{Ed}_i)$$

# Model-Implied Simple Main Effects

- Old Equation using uncentered predictors:  
$$\text{GPA}_i = \beta_0 + (\beta_1 * \text{Att}_i) + (\beta_2 * \text{Ed}_i) + (\beta_3 * \text{Att}_i * \text{Ed}_i) + e_i$$
$$\text{GPA}_i = 30 + (1 * \text{Att}_i) + (2 * \text{Ed}_i) + (0.5 * \text{Att}_i * \text{Ed}_i) + e_i$$
- Intercept: **expected value of GPA** when  $\text{Att}_i=3$  and  $\text{Ed}_i=12$
- Simple main effect of **Att** if  $\text{Ed}_i=12 \rightarrow \text{Att}_i(\beta_1 + \beta_3 * \text{Ed}_i)$
- Simple main effect of **Ed** if  $\text{Att}_i=3 \rightarrow \text{Ed}_i(\beta_2 + \beta_3 * \text{Att}_i)$

```
TITLE "Calculating Model-Implied Parameters";
PROC MIXED DATA=dataname ITDETAILS METHOD=ML;
MODEL y = att ed att*ed / SOLUTION;
ESTIMATE "GPA if Att=3, Ed=12"      intercept 1 att 3 ed 12 att*ed 36;
ESTIMATE "Effect of Att if Ed=12"   att 1 att*ed 12;
ESTIMATE "Effect of Ed if Att=3"    ed 1 att*ed 3;
RUN;
```

In ESTIMATE statements, the variables refer to their betas; the numbers refer to the operations of their betas.

These estimates would be given directly by the model parameters instead if you re-centered the predictors as:  $\text{Att}-3$ ,  $\text{Ed}-12$ .

# More Generally...

- Can decompose a **2-way interaction** by testing the simple effect of X at different levels of Z (and vice-versa)
  - Use ESTIMATEs to request simple effects at any point of the interacting predictor
  - Re-centering the interacting predictor at those points will also work
- More general rules, given a **3-way interaction**:
  - *Simple (main) effects move the intercept*
    - 1 possible interpretation for each simple main effect
    - Each simple effect is conditional on other two variables = 0
  - *The 2-way interactions (3 of them in a 3-way model) move the simple effects*
    - 2 possible interpretations for each 2-way interaction
    - Each 2-way interaction is conditional on third variable = 0
  - *The 3-way interaction moves each of the 2-way interactions*
    - 3 possible interpretations of the 3-way interaction
    - Is highest-order term in model, so is unconditional (applies always)

# Time-Invariant Predictors in Longitudinal Models

- Topics:
  - Summary of building unconditional models for time
  - Missing predictors in MLM
  - Centering predictors and interpreting interactions
  - **Effects of time-invariant predictors**
  - Fixed, systematically varying, and random level-1 effects
  - Model building strategies and assessing significance



# Modeling Time-Invariant Predictors

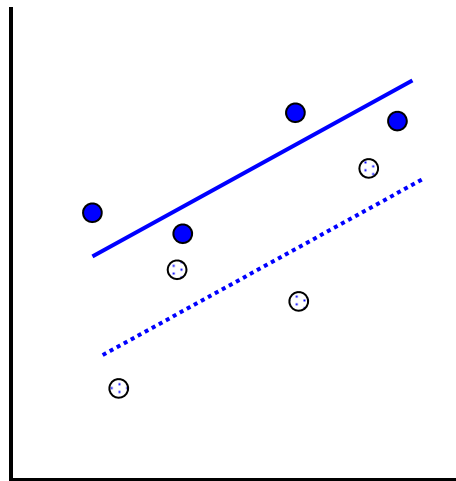
What independent variables can be time-invariant predictors?

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

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

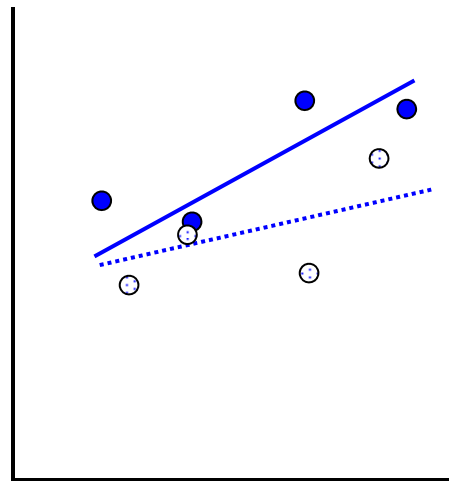
- **In Within-Person Change Models** → Adjust growth curve

Main effect of X, No interaction with time



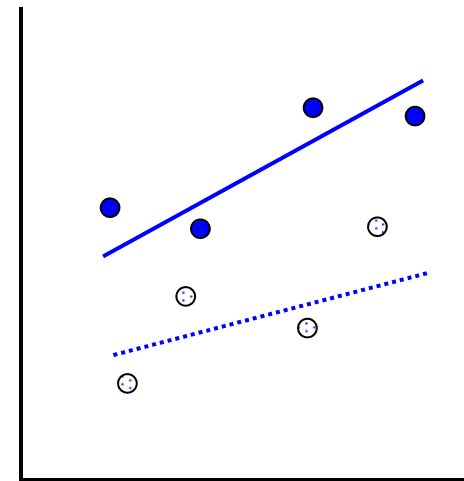
← Time →

Interaction with time, Main effect of X?



← Time →

Main effect of X, and Interaction with time

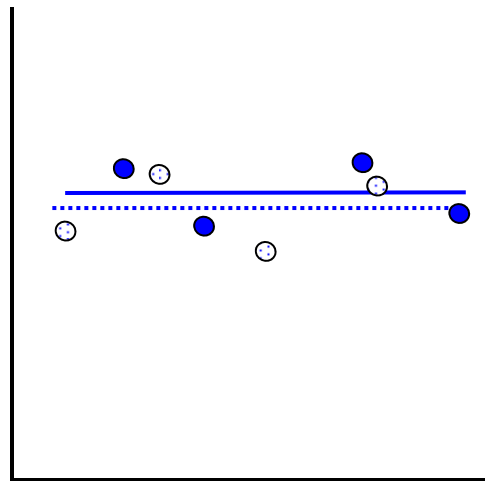


← Time →

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

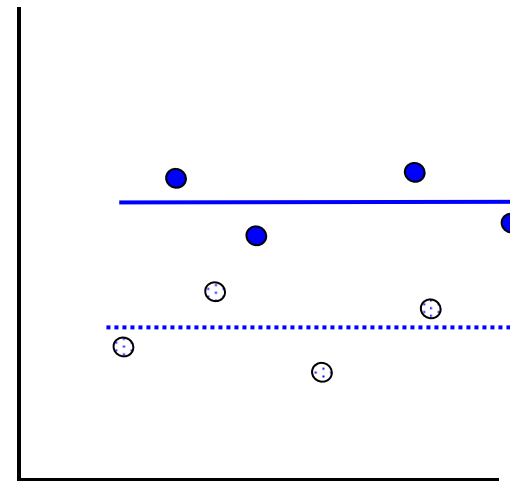
- In **Within-Person Fluctuation Models** → Adjust mean level

No main effect of X



← Time →

Main effect of X



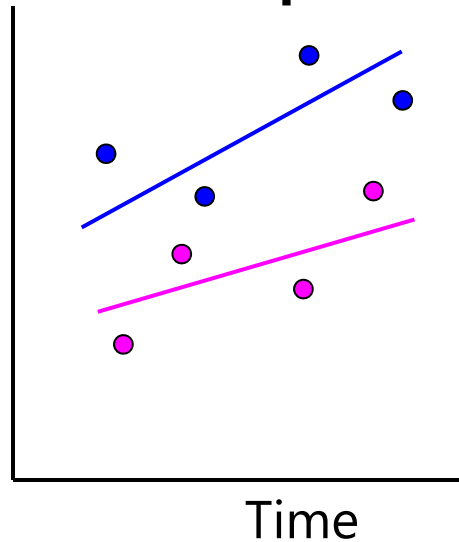
← Time →

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

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

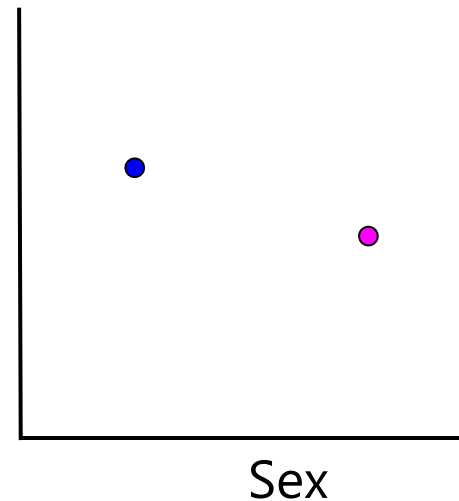
# Why Level-2 Predictors Cannot Have Random Effects in 2-Level Models

Random Slopes for Time



(or Any Level-1 Predictor)

Random Slopes for Sex?



(or any Level-2 Predictor)

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

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

- Main Effect of Education = Education\*Intercept Interaction
  - Moderates the intercept → Increase or decrease in expected outcome at time 0 for every year of education
- Effect of Education on Linear Time = Education\*Time Interaction
  - Moderates the linear time slope → Increase or decrease in expected rate of change at time 0 for every year of education
- Effect of Education on Quadratic Time = Education\*Time<sup>2</sup> Interaction
  - Moderates the quadratic time slope → Increase or decrease in half of expected acceleration/deceleration of linear rate of change for every year of education

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

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

Level 2 Equations (one per  $\beta$ ):

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

Intercept for person  $i$

Fixed Intercept when Time=0 and Ed=12

$\Delta$  in Intercept per unit  $\Delta$  in Ed

Random (Deviation) Intercept after controlling for Ed

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

Linear Slope for person  $i$

Fixed Linear Time Slope when Time=0 and Ed=12

$\Delta$  in Linear Time Slope per unit  $\Delta$  in Ed (=Ed\*time)

Random (Deviation) Linear Time Slope after controlling for Ed

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

Quad Slope for person  $i$

Fixed Quad Time Slope when Ed = 12

$\Delta$  in Quad Time Slope per unit  $\Delta$  in Ed (=Ed\*time<sup>2</sup>)

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

Level 2 Equations (one per  $\beta$ ):

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

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

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

- Composite equation:

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

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



# Time-Invariant Predictors in Longitudinal Models

- Topics:
  - Summary of building unconditional models for time
  - Missing predictors in MLM
  - Centering predictors and interpreting interactions
  - Effects of time-invariant predictors
  - **Fixed, systematically varying, and random level-1 effects**
  - Model building strategies and assessing significance

# Fixed Effects of Time-Invariant Predictors

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

$$\begin{array}{lcl} \beta_{0i} = \gamma_{00} + U_{0i} & \longrightarrow & \beta_{0i} = \gamma_{00} + \gamma_{01}Ed_i + U_{0i} \\ \beta_{1i} = \gamma_{10} + U_{1i} & & \beta_{1i} = \gamma_{10} + \gamma_{11}Ed_i + U_{1i} \\ \beta_{2i} = \gamma_{20} + U_{2i} & & \beta_{2i} = \gamma_{20} + \gamma_{21}Ed_i + U_{2i} \end{array}$$

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

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

# Fixed Effects of Time-Invariant Predictors

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

**This should be ok to do...**

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

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

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

**Is this still ok to do?**

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

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

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

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

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

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

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

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

# Variance Accounted For By Level-2 Time-Invariant Predictors

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

# Variance Accounted for... For Real

- **Pseudo- $R^2$**  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!
- **One last simple 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  $\rightarrow$  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
- MORAL OF THE STORY: Specify EXACTLY which kind of pseudo- $R^2$  you used—give the formula and the reference!!

# Time-Invariant Predictors in Longitudinal Models

- Topics:
  - Summary of building unconditional models for time
  - Missing predictors in MLM
  - Centering predictors and interpreting interactions
  - Effects of time-invariant predictors
  - Fixed, systematically varying, and random level-1 effects
  - **Model building strategies and assessing significance**



# Model-Building Strategies

- It may be helpful to examine predictor effects in separate models at first, including interactions with all growth terms to see the total pattern of effects for a single predictor
  - Question: Does age matter at all in predicting change over time?
  - e.g., random quadratic model + age, age\*time, age\*time<sup>2</sup>
- Then predictor effects can be combined in layers in order to examine unique contributions (and interactions) of each
  - Question: Does age *still* matter after considering reasoning?
  - random quadratic + age, age\*time, age\*time<sup>2</sup>,  
+ reason, reason\*time, reason\*time<sup>2</sup>
  - Potentially also + age\*reason, age\*reason\*time, age\*reason\*time<sup>2</sup>
- Sequence of predictors should be guided by theory and research questions—there may not be a single “best model”
  - One person’s “control” is another person’s “question”, so may not end up in the same place given different orders of predictor inclusion



# Evaluating Statistical Significance of Multiple New Fixed Effects at Once

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

# Evaluating Statistical Significance of New Individual Fixed Effects

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

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

# Denominator DF (DDF) Methods

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

# Denominator DF (DDF) Methods

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

# Wrapping Up...

- MLM uses ONLY rows of data that are COMPLETE: both predictors AND outcomes must be there!
  - Using whatever data you do have for each person will likely lead to better inferences and more statistical power than using only complete persons (listwise deletion)
  - All predictors need to have a meaningful 0 value
- Time-invariant predictors modify the level-1 created growth curve → predict individual intercepts and slopes
  - They account for random effect variances (the predictors are the reasons WHY people need their own intercepts and slopes)
  - If a level-1 effect is not random, it can still be moderated by a cross-level interaction with a time-invariant predictor...
    - ... but then it will predict L1 residual variance instead