Time-Invariant Predictors and Treatment Effects in Longitudinal Models

- Topics:
  - **Introducing time-invariant predictors**
  - Example 1: Between-group treatment effects on weight loss over three occasions in a randomized control trial (RCT)
  - Example 2: Fluid intelligence predicting improvements in response time (RT) across six short-term occasions
  - Example 3: Between-group treatment effects in an RCT with multiple observations within repeated measures conditions
  - Example 4: Within-person treatments in aggregated N-of-1 RCTs
  - Longitudinal designs of individuals in groups

Review: 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**

Now we will add predictors to account for each pile!

(Multivariate) Alternative Covariance Structures

- Models for WP fluctuation typically include only a residual covariance structure, and at most a random intercept (random time slopes won't help in the absence of systematic change)
- Likewise, random time slopes may not be possible given few occasions (especially with piecewise slopes for change)

Between-Person Random Intercept in G + Within-Person Structure in R (to make V)

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

Gets accounted for by time-level predictors

Residual Variance (σ_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=V

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

Total Variance (σ_T^2)

Modeling Time-Invariant Predictors

- Which independent variables can be time-invariant predictors?
  - Aka, "**person-level**" or "**level-2**" or predictors ($\boldsymbol{x_i}$) in two-level models
  - 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
  - **Quantitative** 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 as "factor" variables (e.g., SAS CLASS, SPSS BY, STATA i.), but not in Mplus; I do not like ± 1 coding for group differences (because then 0 = ???)
    - I find indicator or sequential dummy-coding variants most useful

Beware of Missing Predictors

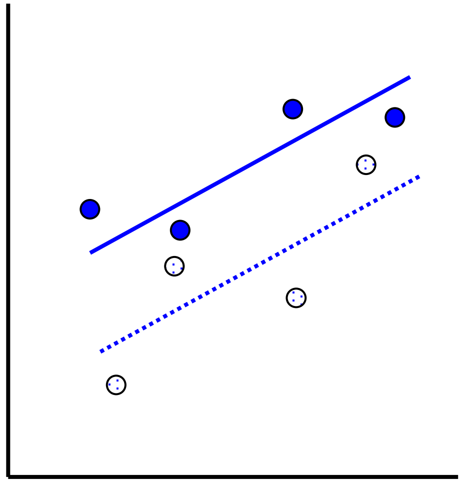
- Any cases missing model predictors that are not part of the joint likelihood* will not be used in that model
  - Not great for time or time-varying predictors (Missing At Random-ish)
  - Really bad for time-invariant predictors (listwise deletion, MCAR)

- Better options for missing predictors:
  - *Bring the predictor into the joint likelihood (only possible in software for truly 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)

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

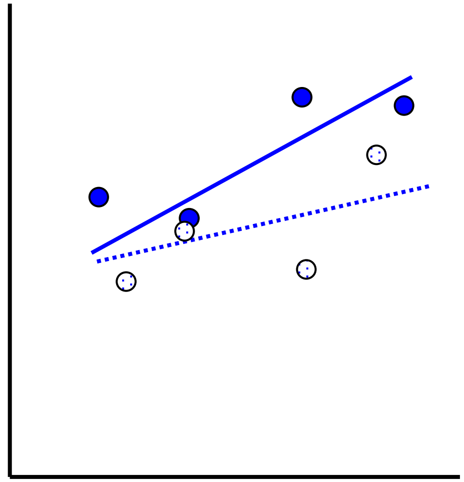
- **In Within-Person Change Models** → Adjust growth curve

Main effect of x_i, no interaction with time



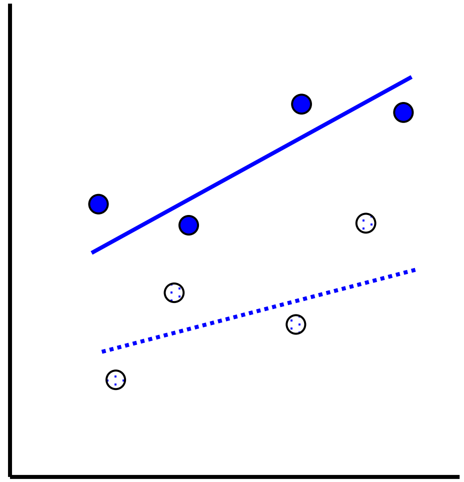
← Time →

Interaction with time, main effect of x_i?



← Time →

Main effect of x_i, and interaction with time

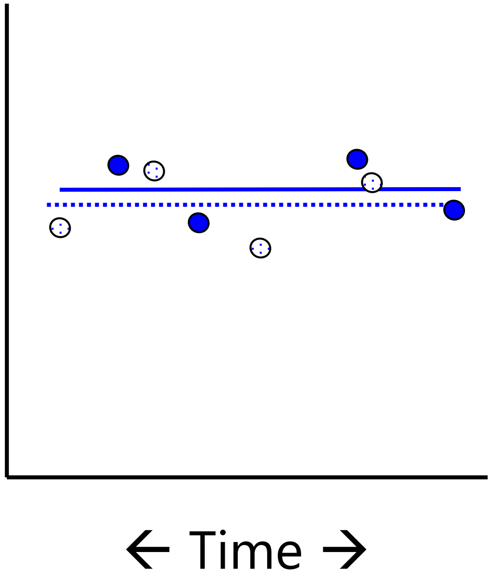


← Time →

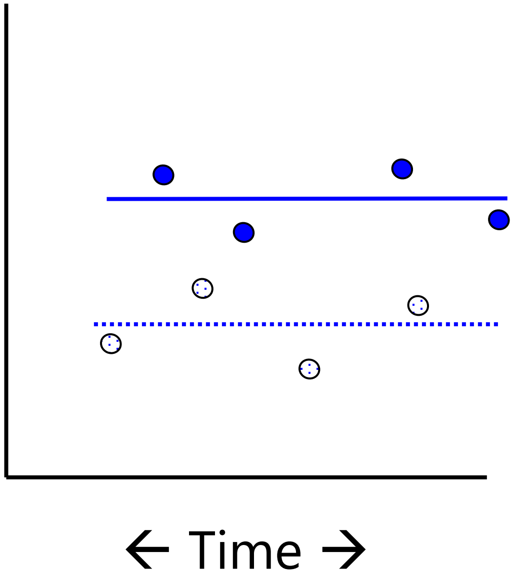
The Role of Time-Invariant Predictors in the **Model for the Means**

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

No main effect of x_i



Main effect of x_i

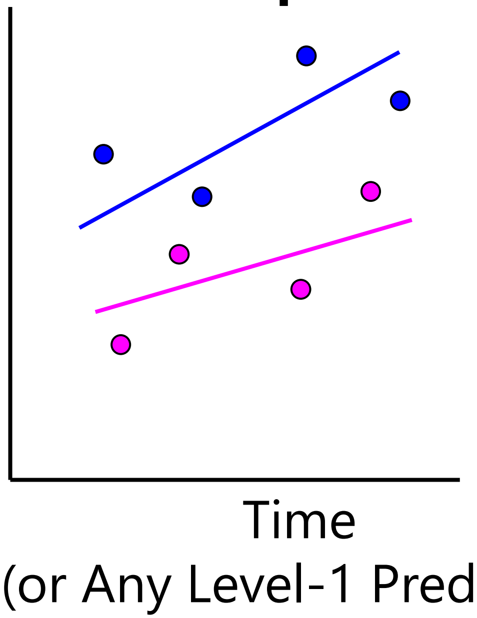


The Role of Time-Invariant Predictors in the **Model for the Variance**

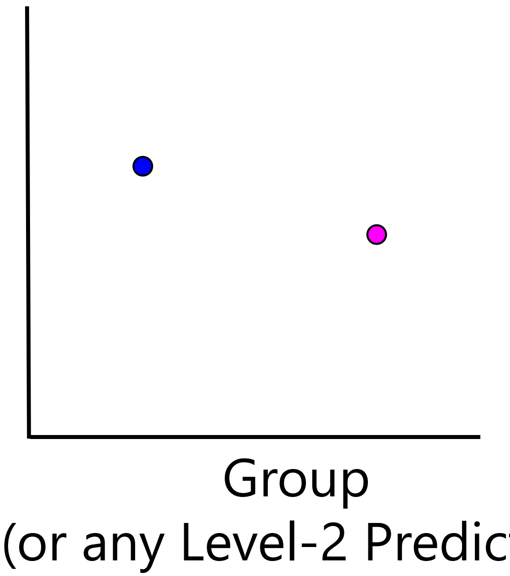
- Beyond fixed effects in the model for the means, time-invariant predictors can be used to allow **heterogeneity of variance** at their level or below in "**location–scale models**"
- e.g., Group as a predictor of heterogeneity of variance:
  - **At level 2**: *Amount* of individual differences in intercepts and/or slopes differs between control and treatment (assumed constant by default!)
  - **At level 1**: *Amount* of within-person residual variation differs between control and treatment (assumed constant by default!)
    - In within-person **fluctuation** model: differential volatility over time
    - In within-person **change** model: differential volatility/inconsistency remaining after controlling for fixed and random effects of time
- These models are harder to estimate and may require custom algorithms (e.g., SAS NLMIXED; in Mplus v 8+ using "logV")

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

Random Slopes for Time



Random Slopes for Group?



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

* *Level-2 predictors can be included as predictors of heterogeneity of variance, which technically is a random slope of sorts (but interpretation is different)*

Sources of Explained Variance by Person-Level-2 Time-Invariant Predictors

- **Fixed effects of level-2 predictors *by themselves*:**
  - Level-2 (BP) main effects reduce level-2 random intercept variance
  - Level-2 (BP) interactions also reduce level-2 random intercept variance
- **Fixed effects of *cross-level interactions* (level-1* level-2):**
  - If a 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 *time* is random, then pred1**time*, pred2**time*, and pred1*pred2**time* can each reduce the level-2 random linear time slope variance
  - If the level-1 predictor <u>not random</u>, any cross-level interaction with it will reduce the level-1 WP **residual variance** instead
    - e.g., if $time^2$ does not have a level-2 random slope, then pred1*$time^2$, pred2*$time^2$, and pred1*pred2*$time^2$ will reduce the level-1 residual variance → Different quadratic slopes by pred1 and pred2 create better level-1 trajectories, thus reducing level-1 residual variance around the trajectories

Variance Explained… Continued

- **Pseudo-R^2** is named that way for a reason... piles of variance can shift around, such that **it can actually become negative**
  - Sometimes is a sign of model mis-specification (but not always)
  - See Rights & Sterba (2019, 2020) for alternative marginal versions of R^2
    - Ensure positive R^2 values, but they don't quantify R^2 for slope variances (boo)

- **A simple alternative: Total R^2** (Singer & Willett, 2003)
  - Generate model-predicted $\hat{y}_{ti}$ from fixed effects only (NOT including random effects, so no cheating) and correlate it with observed y_{ti}
  - Then square that correlation → total R^2 (same as in GLM regression)
  - Total R^2 = total reduction in overall outcome variance across levels
  - Can be "unfair" in models with large unexplained sources of variance (i.e., for sampling dimensions you didn't have predictors for)

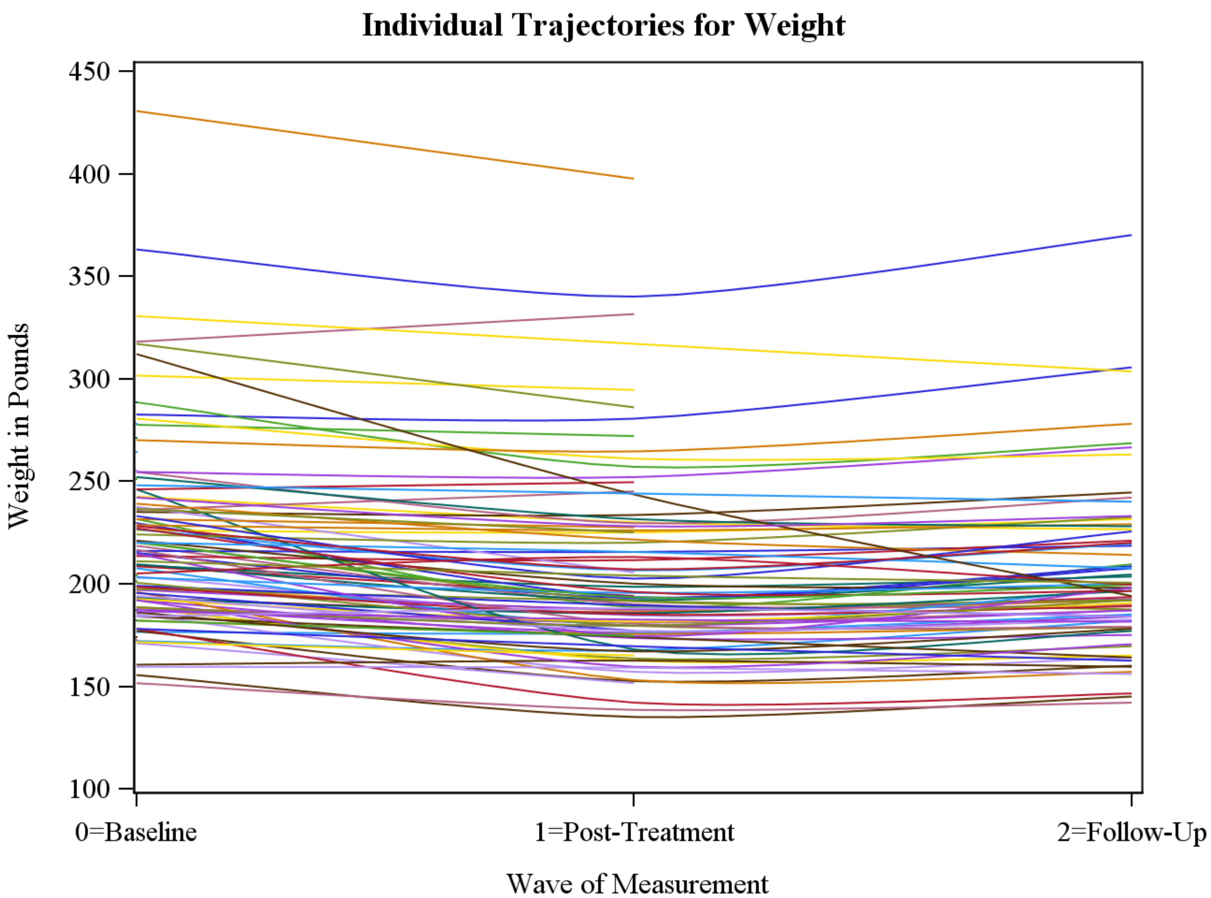
- MORAL OF THE STORY: Specify EXACTLY which kind(s) of R^2 you used—give the formula and a reference!!

Time-Invariant Predictors and Treatment Effects in Longitudinal Models

- Topics:
  - Introducing time-invariant predictors
  - **Example 1: Between-group treatment effects on weight loss over three occasions in a randomized control trial (RCT)**
  - Example 2: Fluid intelligence predicting improvements in response speed across six short-term occasions
  - Example 3: Between-group treatment effects in an RCT with multiple observations within repeated measures conditions
  - Example 4: Within-person treatments in aggregated N-of-1 RCTs
  - Longitudinal designs of individuals in groups

Example 1: Differences in Weight Loss

- Randomized Control Trial for differences in weight loss (and then maintenance of weight loss) between control and treatment groups
  - 105 persons measured up to 3 occasions → balanced but incomplete data
  - Univariate MLM on stacked data (REML estimation, Kenward-Roger DDF)



Example 1: Piecewise Model for the Means

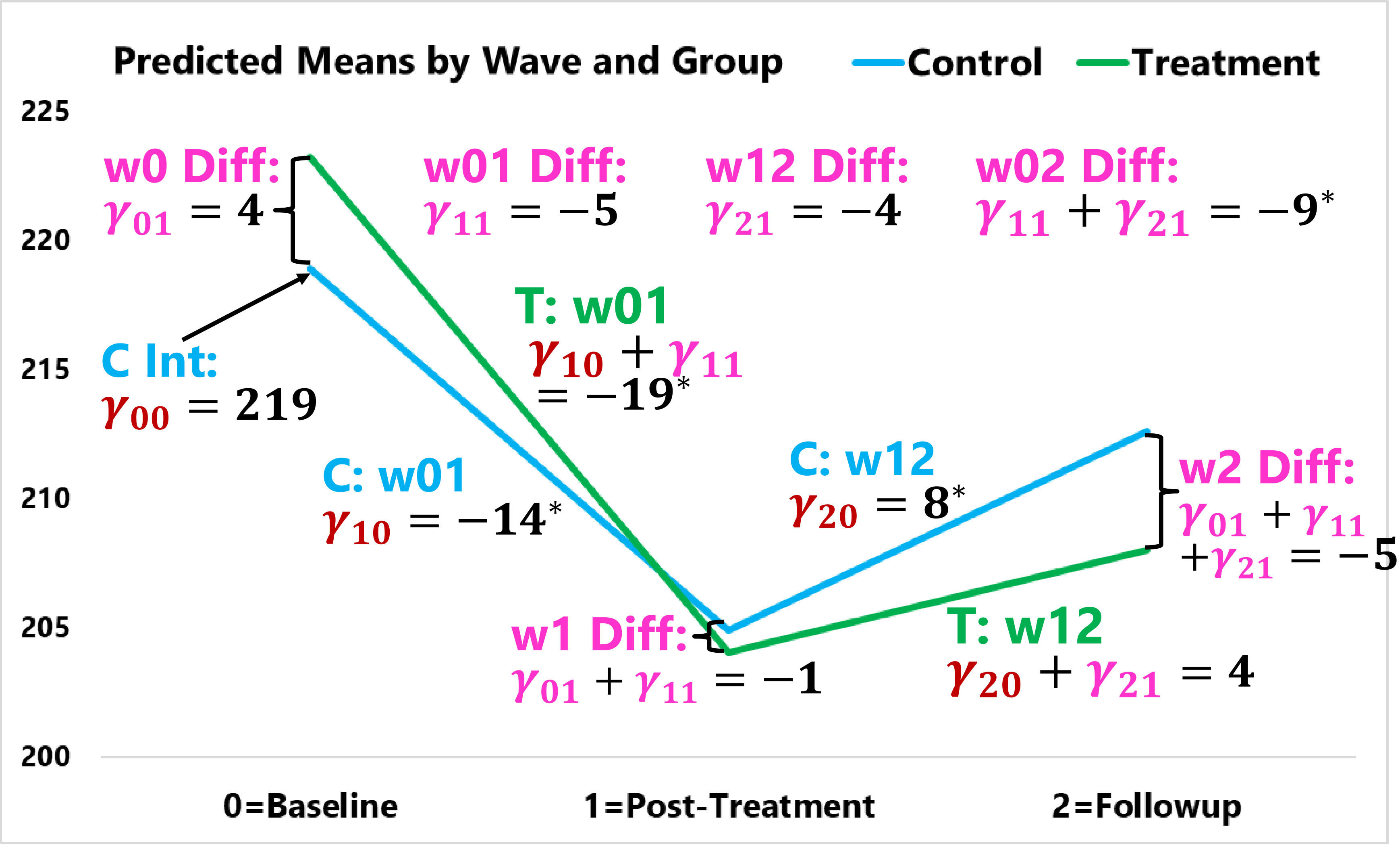
$$\widehat{weight}_{ti} = \gamma_{00} + \gamma_{10}(w01_{ti}) + \gamma_{20}(w12_{ti}) + \gamma_{01}(Tx_i) + \gamma_{11}(w01_{ti})(Tx_i) + \gamma_{21}(w12_{ti})(Tx_i)$$

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

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

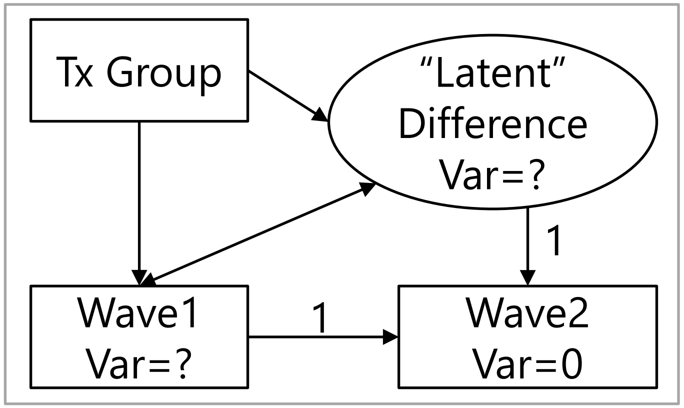
Example 1: Piecewise Model for the Means

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



Example 1: Model for the Variance

- Random effects were not used to model dependency
  - Why? Given 2 slopes across 3 occasions, only 1 of the slopes could have a random effect → assumes parallel change for other time period!
  - The only way to get 2 random slopes from 3 occasions is to remove the WP residual variance, which assumes BP differences in change are measured perfectly!
    - Btw, this is the basis of two-occasion "latent" difference score models (but latent variables are not magic)



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

Example 1: Syntax by Univariate MLM Program (Stacked Data)

SAS:

```
PROC MIXED DATA=work.Example1 COVTEST METHOD=REML;
     CLASS ID wave GroupTx;
     MODEL weight = w01 w12 Tx w01*Tx w12*Tx / SOLUTION DDFM=KR;
     REPEATED wave / RCORR TYPE=UN SUBJECT=ID GROUP=GroupTx;
RUN;
```

R (gls from nlme package)—not sure how to get a different unstructured R matrix by group, described here, and denominator DF will not be correct without follow-up commands:

```
gls(data=Example1, method="REML", model=weight~1+w01+w12+Tx+w01:Tx+w12:Tx,
    correlation=corSymm(form=~as.numeric(wave)|ID), weights=varIdent(form=~1|wave))
```

STATA:

```
mixed weight w01 w12 Tx w01*Tx w12*Tx, || ID: , noconstant variance reml ///
      residuals(unstructured,t(wave) by(GroupTx)) dfmethod(kroger) dftable(pvalue)
```

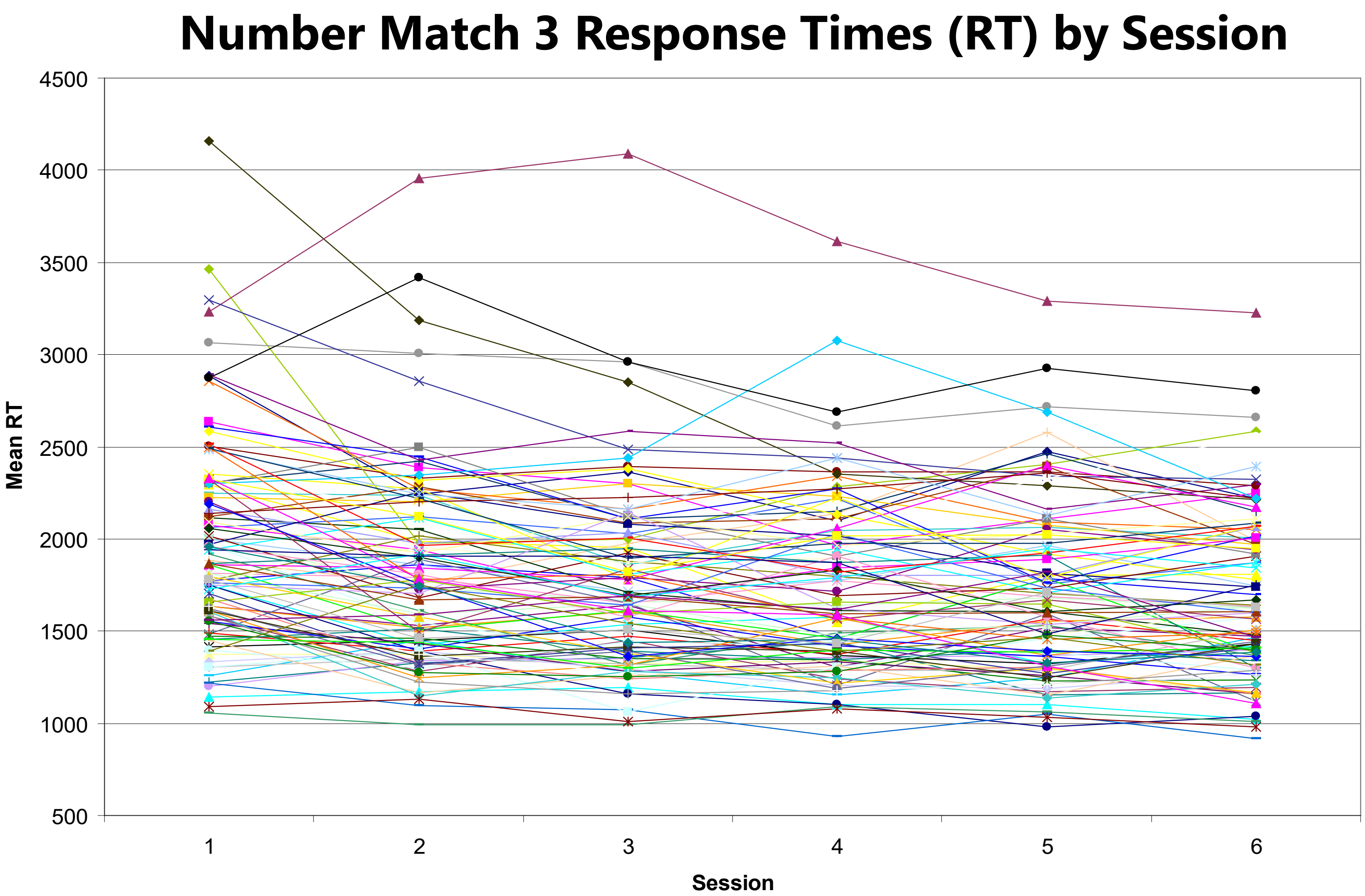
SPSS—I don't think you can get fully different unstructured R matrices by group, but I could be wrong, and only Satterthwaite DF are available:

```
MIXED weight BY ID wave WITH w01 w12 Tx
      /METHOD = REML
      /PRINT = SOLUTION TESTCOV R
      /FIXED = w01 w12 Tx w01*Tx w12*Tx
      /REPEATED = wave | COVTYPE(UN) SUBJECT(ID).
```

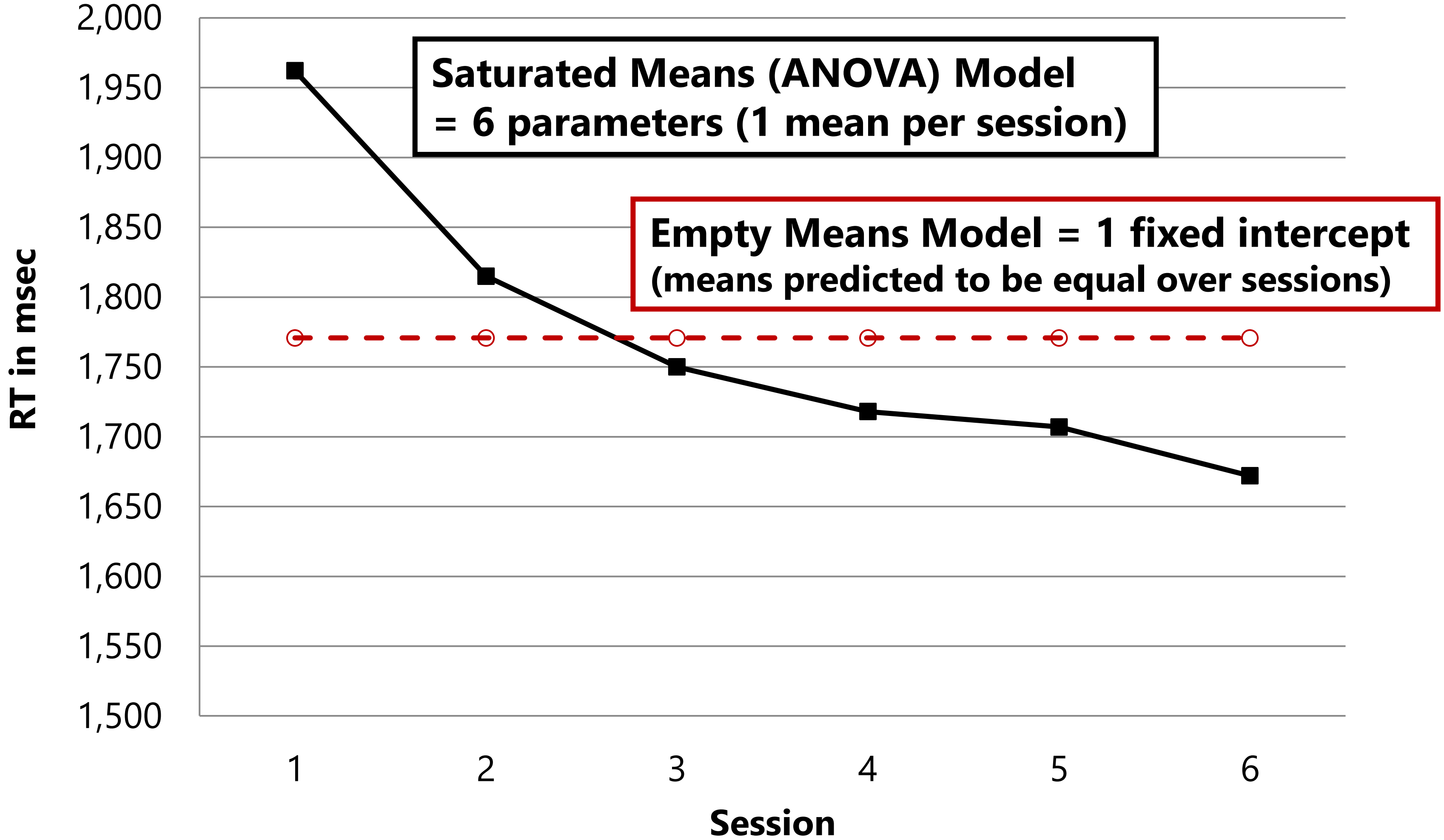
Time-Invariant Predictors and Treatment Effects in Longitudinal Models

- Topics:
  - Introducing time-invariant predictors
  - Example 1: Between-group treatment effects on weight loss over three occasions in a randomized control trial (RCT)
  - **Example 2: Fluid intelligence predicting improvements in response speed across six short-term occasions**
  - Example 3: Between-group treatment effects in an RCT with multiple observations within repeated measures conditions
  - Example 4: Within-person treatments in aggregated N-of-1 RCTs
  - Longitudinal designs of individuals in groups

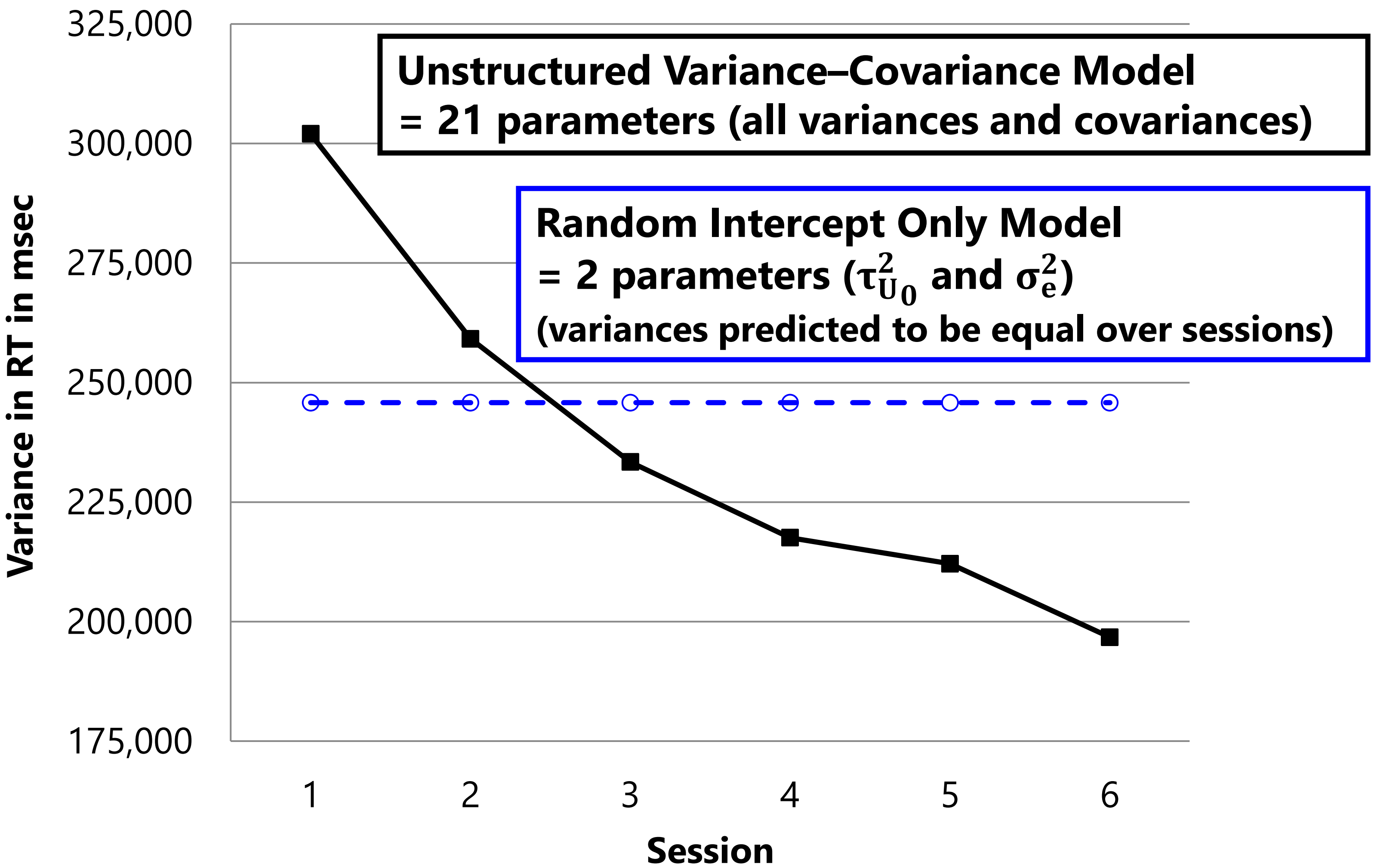
Example 2: Individual Trajectories
101 older adults, 6 occasions within 2 weeks



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



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



Random Quadratic Time Unconditional Model

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

Level 2 Equations (one per β):

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

- 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 Time Slope** for person *i*
- γ_{10}: Fixed (mean) **Linear Slope**
- U_{1i}: Random (Deviation) **Linear Slope**

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

- β_{2i}: **Quadratic Time Slope** for person *i*
- γ_{20}: Fixed (mean) **Quad Slope**
- U_{2i}: Random (Deviation) **Quad Slope**

Time = session − 1
REML estimation using stacked data (univ MLM)
U_i covariances also estimated

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

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

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

Need n = 4 occasions to fit random quadratic time model

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

<u>Level 1</u>: $\mathbf{RT_{ti} = \beta_{0i} + \beta_{1i}Time_{ti} + \beta_{2i}Time_{ti}^2 + e_{ti}}$

<u>Level 2 Equations (one per β):</u>

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

Intercept for person *i*

Fixed Intercept when Time=0 and Reas=22

Δ in Intercept per unit Δ in Reas

Random (Deviation) Intercept after controlling for Reas

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

Linear Slope for person *i*

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

Δ in Linear Time Slope per unit Δ in Reas (=Reas*time)

Random (Deviation) Linear Time Slope after controlling for Reas

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

Quad Slope for person *i*

Fixed Quad Time Slope when Reas=22

Δ in Quad Time Slope per unit Δ in Reas (=Reas*time^2)

Random (Deviation) Quad Time Slope after controlling for Reas

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

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

Level 2 Equations (one per β):

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

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

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

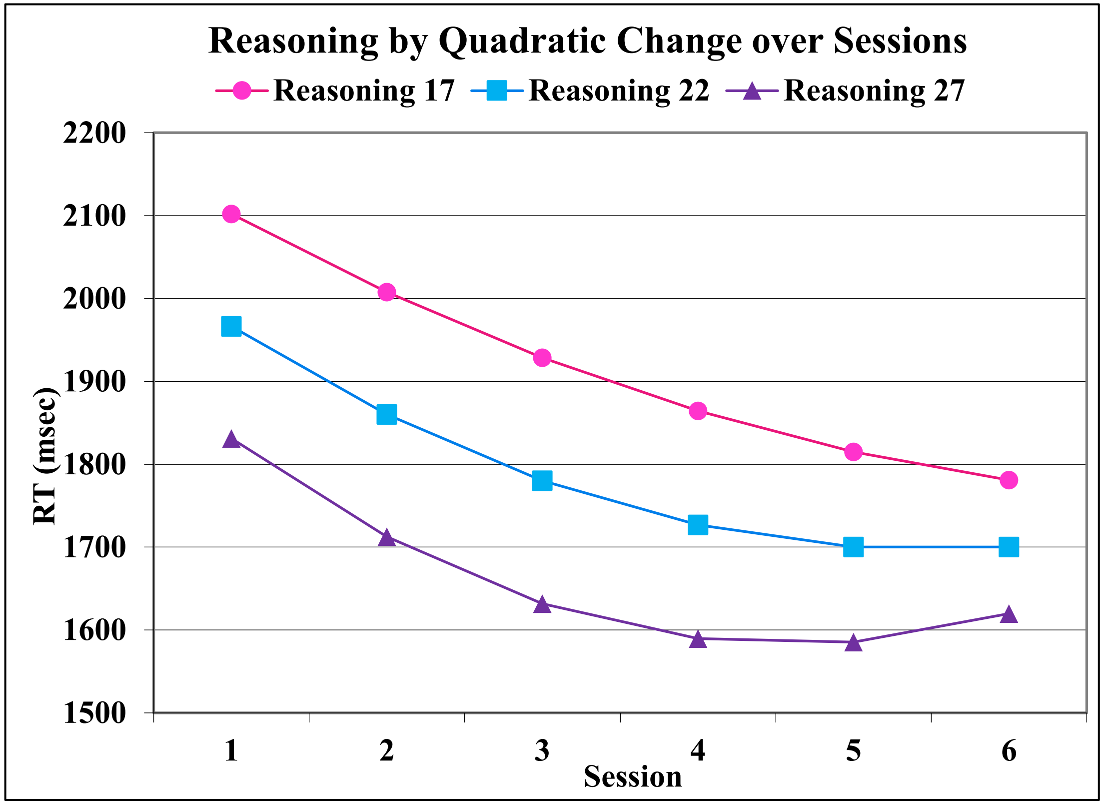
- Composite equation:
- $$y_{ti} = (\gamma_{00} + \gamma_{01}Reas_i + U_{0i}) + (\gamma_{10} + \gamma_{11}Reas_i + U_{1i})Time_{ti} + (\gamma_{20} + \gamma_{21}Reas_i + U_{2i})Time_{ti}^2 + e_{ti}$$

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

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

BP Pseudo-R^2 Values:

Intercept U_{0i} = .049
Linear Time U_{1i} = −.006
Quadratic Time U_{2i} = .024
WP Residual e_{ti} = **0**



People with better reasoning:

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

Example 2: Syntax by Univariate MLM Program (Stacked Data)

SAS:

```
PROC MIXED DATA=work.Example2 COVTEST METHOD=REML;
     CLASS ID;
     MODEL RT = time timesq reas time*reas timesq*reas / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT time timesq / GCORR TYPE=UN SUBJECT=ID;
RUN;
```

R (lmer from lme4 package)—using lmertest package, which does provide correct denominator DF:

```
model2 = lmer(data=Example2, REML=TRUE,
         formula=RT~1+time+timesq+reas
                     +time:reas+timesq:reas+(1+time+timesq|ID))
summary(model2, ddf="Satterthwaite")
```

STATA:

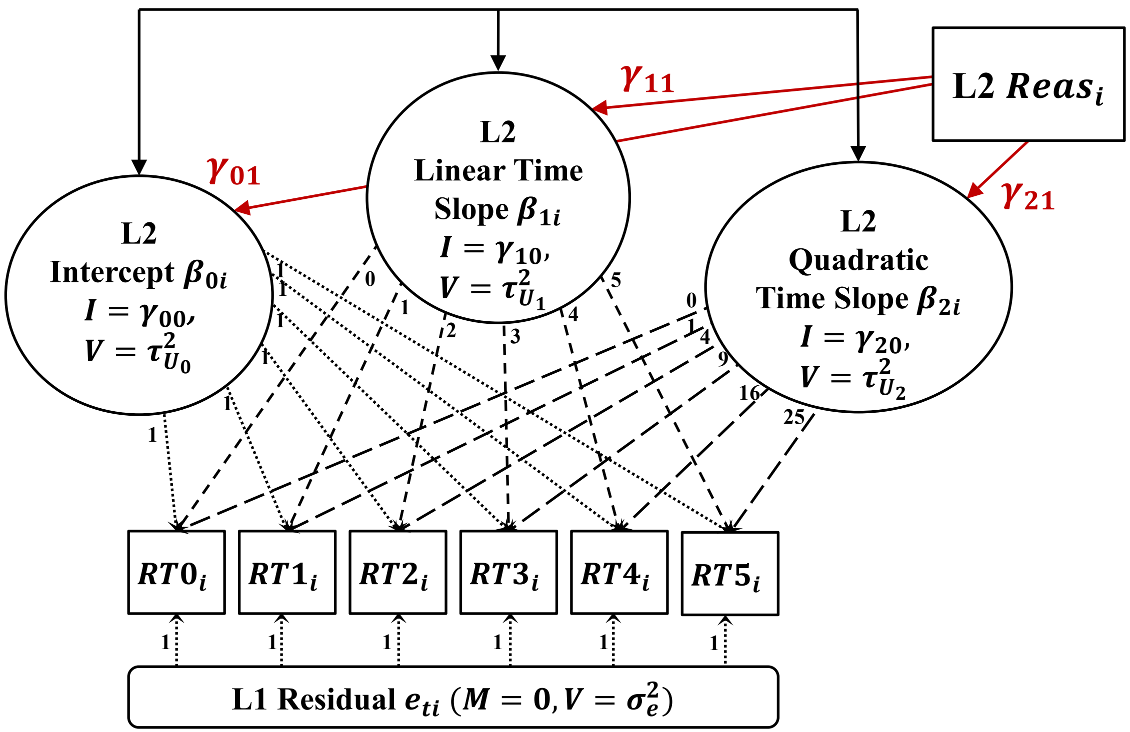
```
mixed RT time timesq reas time#reas timesq#reas, || ID: time timesq, ///
      variance reml covariance(un) dfmethod(satterthwaite) dftable(pvalue)
```

SPSS:

```
MIXED RT BY ID WITH time timesq reas
      /METHOD = REML
      /PRINT = SOLUTION TESTCOV
      /FIXED = time timesq reas time*reas timesq*reas
      /RANDOM = INTERCEPT time timesq | COVTYPE(UN) SUBJECT(ID).
```

Should I have used a "latent" growth curve model (on wide data in SEM) instead?

$$\mathbf{RT_{ti}} = (\gamma_{00} + \gamma_{01}\mathrm{Reas}_i + U_{0i}) + (\gamma_{10} + \gamma_{11}\mathrm{Reas}_i + U_{1i})\mathrm{Time}_{ti} + (\gamma_{20} + \gamma_{21}\mathrm{Reas}_i + U_{2i})\mathrm{Time}_{ti}^2 + e_{ti}$$



Cons:

- No REML, no DDF → Type I error for small *N*
- Requires balanced time *(or definition variables for individual time loadings)*

Pros:

- Latent basis nonlinear change *(fix 1st loading to 0, last to 1, estimate other loadings for % change)*
- More flexibility in WP residual heterogeneity of variance and covariance
- Change in latent variables instead of observed

Example 2: Mplus Single-Level SEM Syntax

Just showing **MODEL** part, which would be preceded by **DATA**, **VARIABLE**, and **ANALYSIS** as usual (estimated using **wide** data)

```
!!!!! Random quadratic model of change

! Factor loadings fixed by @
  Int BY RT0@1 RT1@1 RT2@1 RT3@1 RT4@1  RT5@1;
  Lin BY RT0@0 RT1@1 RT2@2 RT3@3 RT4@4  RT5@5;
  Qua BY RT0@0 RT1@1 RT2@4 RT3@9 RT4@16 RT5@25;

! Factor intercepts estimated = fixed effects
  [Int Lin Qua];
! Level-2 factor variances estimated (in G)
  Int Lin Qua;
! Level-2 factor covariances estimated (in G)
  Int Lin Qua WITH Int Lin Qua;

! Per-occasion intercepts fixed to 0
  [RT0@0 RT1@0 RT2@0 RT3@0 RT4@0 RT5@0];

! Level-1 residual variances held equal (in R)
  RT0 RT1 RT2 RT3 RT4 RT5 (ResVar);

! Fixed effects of reasoning → latent factors
  Int Lin Qua ON reas;
```

```
!!!!! Random latent basis model of change

! Factor loadings fixed by @
  Int BY RT0@1 RT1@1 RT2@1 RT3@1 RT4@1 RT5@1;
  Slp BY RT0@0 RT1*  RT2*  RT3*  RT4*  RT5@1;
! Loadings estimated as 0.57, 0.76, 0.90, 0.97

! Factor intercepts estimated = fixed effects
  [Int Slp];
! Level-2 factor variances estimated (in G)
  Int Slp;
! Level-2 factor covariance estimated (in G)
  Int WITH Slp;

! Per-occasion intercepts fixed to 0
  [RT0@0 RT1@0 RT2@0 RT3@0 RT4@0 RT5@0];

! Level-1 residual variances held equal (in R)
  RT0 RT1 RT2 RT3 RT4 RT5 (ResVar);

! Fixed effects of reasoning → latent factors
  Int Slp ON reas;
```

Note: There are Mplus syntax shortcuts for growth models I am not using: (1) to be explicit about what the model contains, (2) to not estimate separate residual variances

Example 2: R Single-Level SEM Syntax

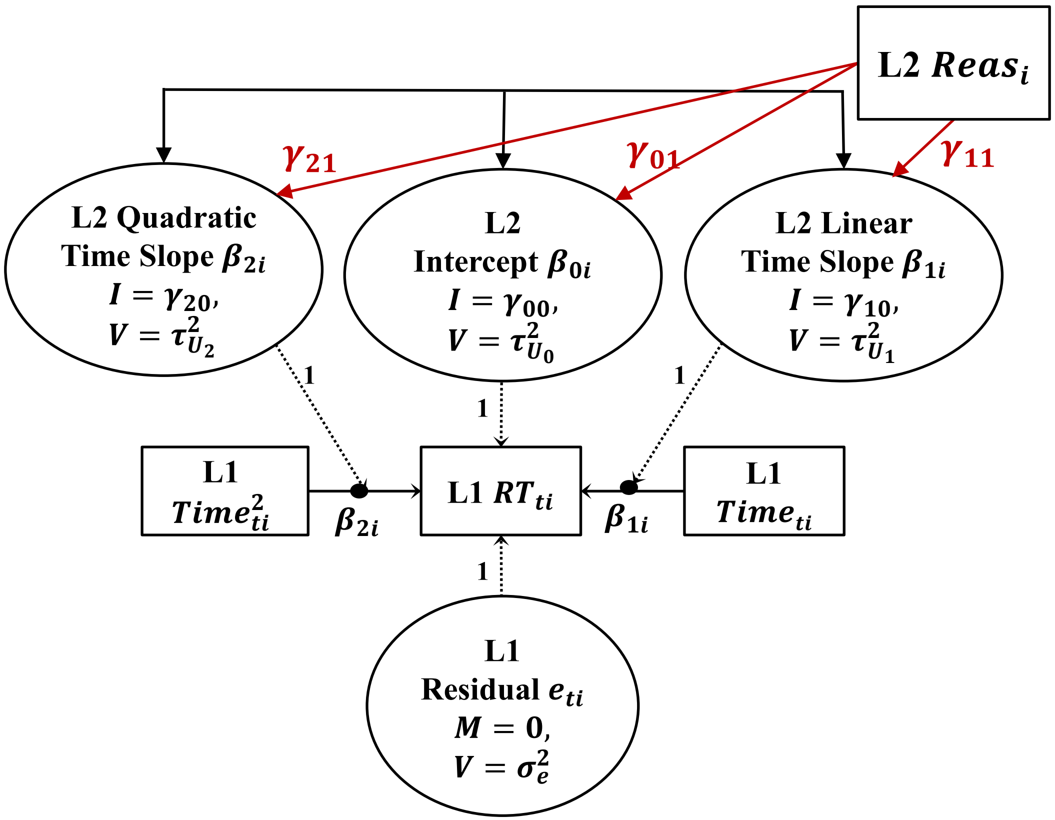
```
RandQuadSyntax = "
# Factor loadings fixed by *
Int =~ 1*RT0 + 1*RT1 + 1*RT2 + 1*RT3 +  1*RT4 +  1*RT5
Lin =~ 0*RT0 + 1*RT1 + 2*RT2 + 3*RT3 +  4*RT4 +  5*RT5
Qua =~ 0*RT0 + 1*RT1 + 4*RT2 + 9*RT3 + 16*RT4 + 25*RT5

# Factor intercepts estimated = fixed effects
  Int ~ 1;  Lin ~ 1;  Qua ~ 1
# Level-2 factor variances estimated (in G)
  Int ~~ Int;  Lin ~~ Lin;  Qua ~~ Qua
# Level-2 factor covariances estimated (in G)
  Int ~~ Lin + Qua;  Lin ~~ Qua

# Per-occasion intercepts fixed to 0
  RT0 ~ 0; RT1 ~ 0; RT2 ~ 0
  RT3 ~ 0; RT4 ~ 0; RT5 ~ 0

! Level-1 residual variances held equal (in R)
  RT0 ~~ (ResVar)*RT0;  RT1 ~~ (ResVar)*RT1
  RT2 ~~ (ResVar)*RT2;  RT3 ~~ (ResVar)*RT3
  RT4 ~~ (ResVar)*RT4;  RT5 ~~ (ResVar)*RT5

# Fixed effects of reasoning --> latent factors
  Int + Lin + Qua ~ reas
"
RQModel = lavaan(data=Example2wide,
                 model=RandQuadSyntax,
                 estimator="ML", mimic="mplus")
summary(RQModel, fit.measures=TRUE, rsquare=TRUE,
        standardized=TRUE)
```

```
LatentBasisSyntax = "
# Factor loadings fixed by *
Int =~ 1*RT0 + 1*RT1 + 1*RT2 + 1*RT3 +  1*RT4 +  1*RT5
Slp =~ 0*RT0 +   RT1 +   RT2 +   RT3 +    RT4 +  1*RT5
# Loadings estimated as 0.57, 0.76, 0.90, 0.97

# Factor intercepts estimated = fixed effects
  Int ~ 1;  Slp ~ 1
# Level-2 factor variances estimated (in G)
  Int ~~ Int;  Slp ~~ Slp
# Level-2 factor covariances estimated (in G)
  Int ~~ Slp

# Per-occasion intercepts fixed to 0
  RT0 ~ 0; RT1 ~ 0; RT2 ~ 0
  RT3 ~ 0; RT4 ~ 0; RT5 ~ 0

! Level-1 residual variances held equal (in R)
  RT0 ~~ (ResVar)*RT0;  RT1 ~~ (ResVar)*RT1
  RT2 ~~ (ResVar)*RT2;  RT3 ~~ (ResVar)*RT3
  RT4 ~~ (ResVar)*RT4;  RT5 ~~ (ResVar)*RT5

# Fixed effects of reasoning --> latent factors
  Int + Slp ~ reas
"
LBModel = lavaan(data=Example2wide,
                 model=LatentBasisSyntax,
                 estimator="ML", mimic="mplus")
summary(LBModel, fit.measures=TRUE, rsquare=TRUE,
        standardized=TRUE)
```

Note: There are lavaan syntax shortcuts for growth models I am not using: (1) to be explicit about what the model contains, (2) to not estimate separate residual variances

Should I have used "multilevel SEM" (on long data) instead? Not in this case…

$$\mathbf{RT}_{ti} = (\gamma_{00} + \gamma_{01}\mathbf{Reas}_i + U_{0i}) + (\gamma_{10} + \gamma_{11}\mathbf{Reas}_i + U_{1i})\mathbf{Time}_{ti} + (\gamma_{20} + \gamma_{21}\mathbf{Reas}_i + U_{2i})\mathbf{Time}_{ti}^2 + e_{ti}$$



Cons:

- No REML, no DDF → Type I error for small *N*
- ~~Requires balanced time *(or definition variables for individual time loadings)*~~

Pros:

- ~~Latent basis nonlinear change *(fix 1st loading to 0, last to 1, estimate other loadings for % change)*~~
- ~~More flexibility in WP residual heterogeneity of variance and covariance~~
- Change in latent variables instead of observed

Example 2: Mplus M-SEM Syntax

Just showing **MODEL** part, which would be preceded by **DATA**, **VARIABLE**, and **ANALYSIS** as usual (estimated using **long** data)

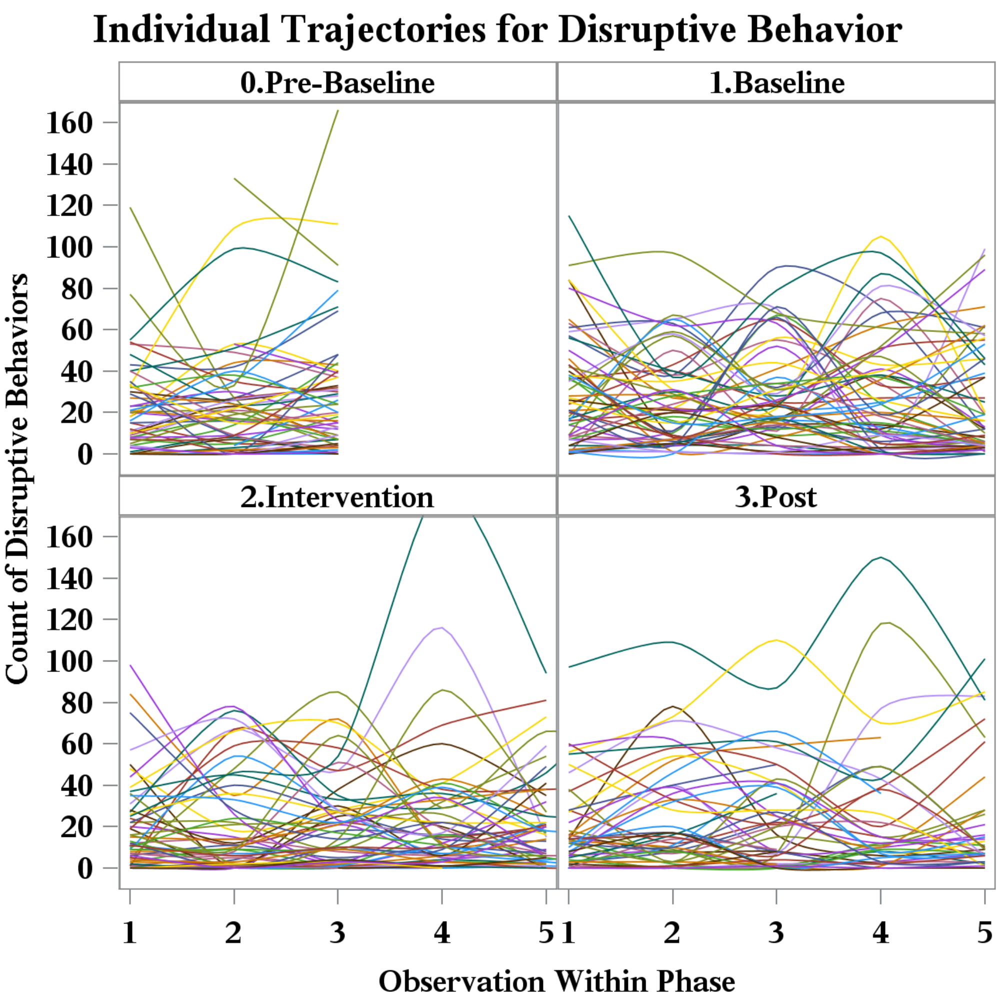
```
%WITHIN%
  RT;                    ! Level-1 residual variance
  Lin | RT ON time;      ! Create beta1i placeholder
  Qua | RT ON timesq;    ! Create beta2i placeholder

%BETWEEN%
  [RT Lin Qua];                   ! Intercepts
   RT Lin Qua;                    ! Level-2 random effect variances
   RT Lin Qua WITH RT Lin Qua;    ! Level-2 random effect covariances
   RT Lin Qua ON reas;            ! Fixed effects of reasoning
```

- Note: R's lavaan package does have M-SEM, but it is much more limited than M-SEM in Mplus:
  - Listwise deletion for any rows (occasions) with missing values
  - No random slopes!

Time-Invariant Predictors and Treatment Effects in Longitudinal Models

- Topics:
  - Introducing time-invariant predictors
  - Example 1: Between-group treatment effects on weight loss over three occasions in a randomized control trial (RCT)
  - Example 2: Fluid intelligence predicting improvements in response speed across six short-term occasions
  - **Example 3: Between-group treatment effects in an RCT with multiple observations within repeated measures conditions**
  - Example 4: Within-person treatments in aggregated N-of-1 RCTs
  - Longitudinal designs of individuals in groups

Example 3: Multiple Condition RCT



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

Example 3: Multivariate MLM

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

Level 2:

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

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

- **Model for the Variance—Multivariate* Two-Level Model:**
  - Unstructured level-2 G matrix for 4 random intercepts → separate variances and covariances by condition (*which makes it a multivariate MLM, reviewer 2)
  - Heterogeneity of level-1 residual variance examined by condition (not found for DB)
  - Could have also examined individual differences *in level-1 variance* in “scale” model!

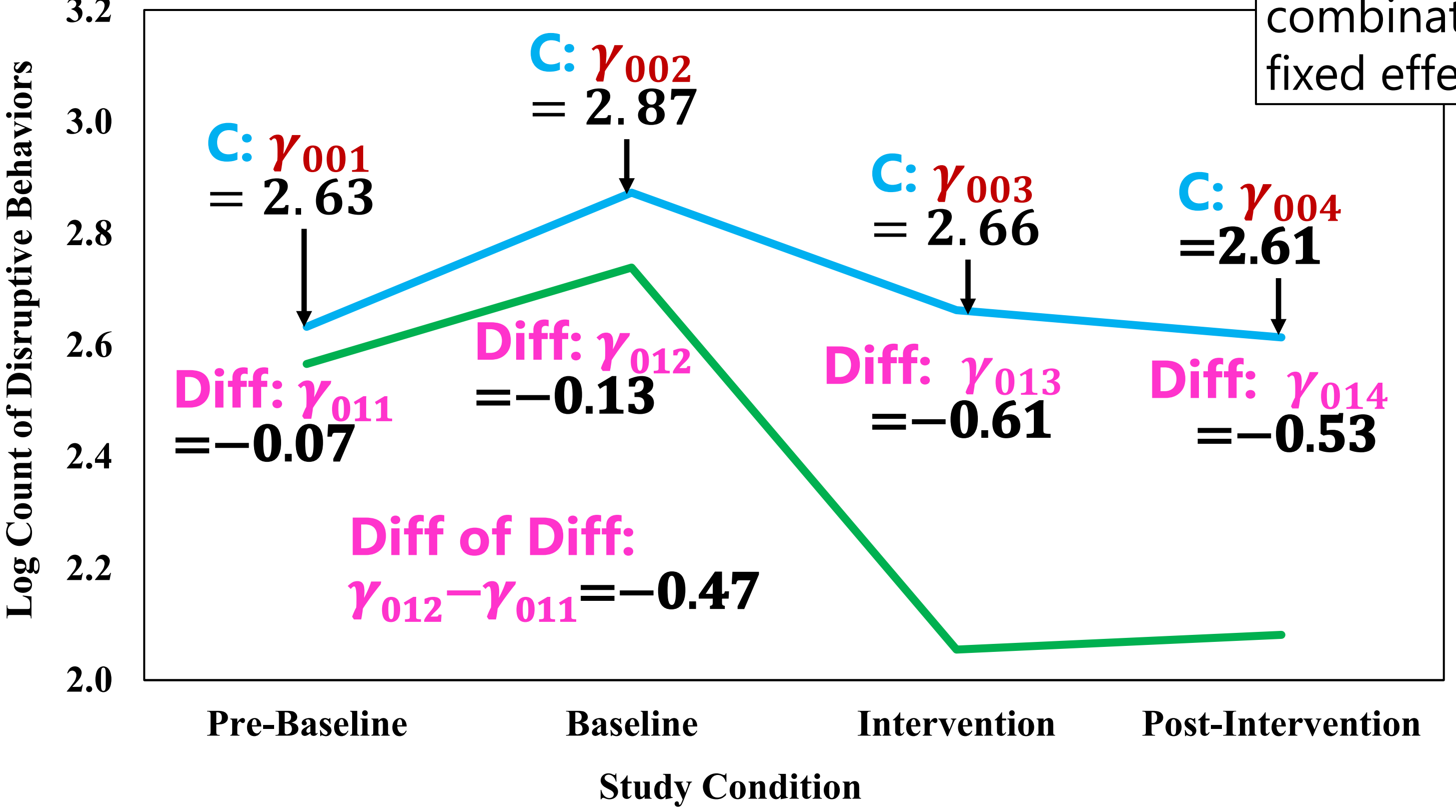
Example 3: Multivariate MLM

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

Level 2:

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

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



Example 3: Syntax by Univariate MLM Program (Stacked Data)

SAS:

```
PROC MIXED DATA=work.Example3 COVTEST METHOD=REML;
     CLASS ID;
     MODEL logDB = Pre Base Inter Post
                   Pre*Tx Base*Tx Inter*Tx Post*Tx / NOINT SOLUTION DDFM=KR;
     RANDOM Pre Base Inter Post / GCORR TYPE=UN SUBJECT=ID;
RUN;
```

R (lmer from lme4 package)—using lmertest package, which does provide correct denominator DF:

```
model3 = lmer(data=Example3, REML=TRUE, formula=logDB~0+Pre+Base+Inter+Post
              +Pre:Tx+Base:Tx+Inter:Tx+Post:Tx+(0+Pre+Base+Inter+Post|ID))
summary(model3, ddf="Kenward-Roger")
```

STATA:

```
mixed logDB Pre Base Inter Post Pre#Tx Base#Tx Inter#Tx Post#Tx , ///
      || ID: Pre Base Inter Post , noconstant variance reml covariance(un) ///
      dfmethod(kroger) dftable(pvalue)
```

SPSS—only Satterthwaite DF are available:

```
MIXED logDB BY ID WITH Pre Base Inter Post Tx
      /METHOD = REML
      /PRINT = SOLUTION TESTCOV G
      /FIXED = Pre Base Inter Post Pre*Tx Base*Tx Inter*Tx Post*Tx
      /RANDOM = Pre Base Inter Post | COVTYPE(UN) SUBJECT(ID).
```

Time-Invariant Predictors and Treatment Effects in Longitudinal Models

- Topics:
  - Introducing time-invariant predictors
  - Example 1: Between-group treatment effects on weight loss over three occasions in a randomized control trial (RCT)
  - Example 2: Fluid intelligence predicting improvements in response speed across six short-term occasions
  - Example 3: Between-group treatment effects in an RCT with multiple observations within repeated measures conditions
  - **Example 4: Within-person treatments in aggregated N-of-1 RCTs**
  - Longitudinal designs of individuals in groups

Example 4: Fixed vs Random 8 Persons

- Data simulated to mimic **aggregated N-of-1 RCT** to improve daily minutes of vigorous physical activity (MVPA) in kids
  - Cushing, Walters, and Hoffman 2014

- 8 kids measured for 30 days (15 treatment, 15 control)
  - Randomly assigned to receive motivating texts on half of the days → **treatment is a time-varying predictor** *with no between-person variance (so it can only have a* ***within-person effect****, stay tuned)*
  - Days are treated as interchangeable here (but in practice one may need to control for day of week, schedule differences, etc)

- Two alternative analytic approaches demonstrated:
  - Persons as **fixed effects** (single-level model for the variance): ***For which*** *kids did the intervention work?*
  - Persons as **random effects** (two-level model for the variance): ***Why*** *does the intervention work for some kids more than others?*

Example 4: Fixed vs Random 8 Persons

- **Persons as fixed effects (single-level model):**

$$\boldsymbol{MVPA_{ti} = \gamma_{00} + \gamma_{10}(Text_{ti}) + \gamma_{01}(ID1_i) + \cdots \gamma_{07}(ID7_i)}$$
$$\boldsymbol{+ \gamma_{11}(Text_{ti})(ID1_i) + \cdots \gamma_{17}(Text_{ti})(ID7_i) + e_{ti}}$$

- **Model for the Means:**
  - $ID1_i$ to $ID7_i$: kid-specific dummy codes (0=not them, 1=them)
  - Reference kid ($ID_i = 8$): intercept (activity on control days) = $\boldsymbol{\gamma_{01}}$ and treatment difference (activity change on texting days) = $\boldsymbol{\gamma_{10}}$
  - Per-kid intercept difference relative to reference kid ($ID_i = 8$): $\boldsymbol{\gamma_{01}}$… $\boldsymbol{\gamma_{07}}$
  - Per-kid treatment difference relative to reference kid ($ID_i = 8$): $\boldsymbol{\gamma_{10}}$… $\boldsymbol{\gamma_{17}}$
  - All other differences are given as linear combinations of fixed effects
- **Model for the Variance—Single-Level Model:**
  - All between-person variances are already covered by the fixed effects!
  - Within-person residual $\boldsymbol{e_{ti}}$ variance = remaining variation among days of same kind (could be estimated separately per kid, too)

Example 4: Fixed Effects Model Syntax by GLM Program (Stacked Data)

SAS:

```
PROC GLM DATA=work.Example4
     CLASS ID;
     MODEL MVPA = ID Text ID*Text / SOLUTION;
RUN;
```

R (lm from base R):

```
model4F = lm(data=Example4, formula=MVPA~1+factor(ID)+Text+factor(ID):Text)
summary(model4F)
```

STATA:

```
regress MVPA ib(last).ID Text ib(last).ID#Text
```

SPSS:

```
GLM MVPA BY ID WITH Text
    /DESIGN = Intercept, ID, Text, ID BY Text.
```

Example 4: Fixed vs Random 8 Persons

Persons as random effects (two-level model):

- **Level 1:** $MVPA_{ti} = \beta_{0i} + \beta_{1i}(Text_{ti}) + e_{ti}$
- **Level 2 Intercept:** $\beta_{0i} = \gamma_{00} + \gamma_{01}(Eff_i) + U_{0i}$
  Level 2 Texting Effect: $\beta_{0i} = \gamma_{10} + \gamma_{11}(Eff_i) + U_{1i}$

- **Model for the Means (Eff_i = self-efficacy):**
  - Mean intercept (activity on control days) for $Eff_i = 0$: γ_{00}
  - Mean texting effect (activity change) for $Eff_i = 0$: γ_{10}
  - Difference in mean intercept per unit Eff_i : γ_{01}
  - Difference in mean texting effect per unit Eff_i : γ_{11}
- **Model for the Variance—Two-Level Model:**
  - Remaining BP differences on control days: U_{0i} variance
  - Remaining BP differences in texting effect: U_{1i} variance
  - Remaining WP variation among days of same kind: residual e_{ti} variance (could have a random intercept variance in a "scale model", too!)

Example 4: Random Effects Syntax by Univariate MLM Program (Stacked Data)

SAS:

```
PROC MIXED DATA=work.Example4 COVTEST METHOD=REML;
     CLASS ID session;
     MODEL MVPA = Text Eff Text*Eff / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT Text / GCORR TYPE=UN SUBJECT=ID;
RUN;
```

R (lmer from lme4 package)—using lmertest package, which does provide correct denominator DF:

```
model4R = lmer(data=Example4, REML=TRUE,
          formula=MVPA~1+Text+Eff+Text:Eff+(1+Text|ID))
summary(model4R, ddf="Satterthwaite")
```

STATA:

```
mixed MPVA Text Eff Text#Eff, || ID: Text, ///
      variance reml covariance(un) dfmethod(satterthwaite) dftable(pvalue)
```

SPSS:

```
MIXED MVPA BY ID WITH Text Eff
      /METHOD = REML
      /PRINT = SOLUTION TESTCOV
      /FIXED = Text Eff Text*Eff
      /RANDOM = INTERCEPT Text | COVTYPE(UN) SUBJECT(ID).
```

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 (i.e., the predictors are the reasons WHY people need their own intercepts and slopes)
  - If a level-1 effect is not random, it can still be moderated by a cross-level interaction with a time-invariant predictor…
    - … but then it will predict L1 residual variance instead

Time-Invariant Predictors and Treatment Effects in Longitudinal Models

- Topics:
  - Introducing time-invariant predictors
  - Example 1: Between-group treatment effects on weight loss over three occasions in a randomized control trial (RCT)
  - Example 2: Fluid intelligence predicting improvements in response speed across six short-term occasions
  - Example 3: Between-group treatment effects in an RCT with multiple observations within repeated measures conditions
  - Example 4: Within-person treatments in aggregated N-of-1 RCTs
  - **Longitudinal designs of individuals in groups**

2 Options for Differences Across Groups

Represent Group Dependency as Fixed Effects

- Include (#groups-1) contrasts for group membership in the **model for the means fixed effects**→ so group is NOT another "level"
- Permits inference about differences between specific groups, but you cannot include between-group predictors (group is saturated)
- Recommended if #groups < 15ish (b/c variances are unreliable)

Represent Group Differences as a Random Effect

- Include a **random intercept variance in the model for the variance**, such that group differences become another "level"
- Permits inference about differences across groups more generally, for which you can test effects of between-group predictors
- Better if #groups > 15ish and you want to **predict** group differences

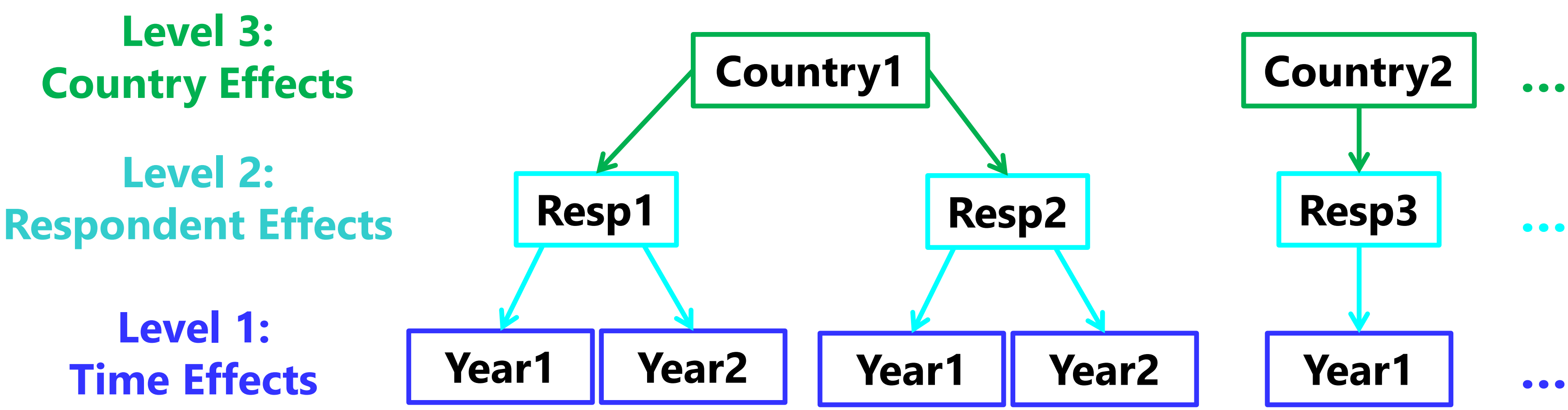
What determines the number of levels?

- **Answer: the model for the outcome variance ONLY**
- How many **dimensions of sampling** in the **outcome**?
  - Longitudinal, one kid per teacher, one school? → 2-level model
  - Longitudinal, 2+ kids per teacher, one school? → 3-level model
  - Longitudinal, 2+ kids per teacher, many schools? → 4-level model
  - Sampling dimensions may also be crossed instead of nested, or may be modeled with fixed effects if the # units is small

- Need at least one pile of variance per dimension (for 3 levels, that's 2 sets of random effects and a residual)
  - Include whatever predictors you want per level, but keep in mind that the usefulness of your predictors will be constrained by the amount of outcome variance in its relevant sampling dimension

Types of 3-Level Designs: Clustered Longitudinal

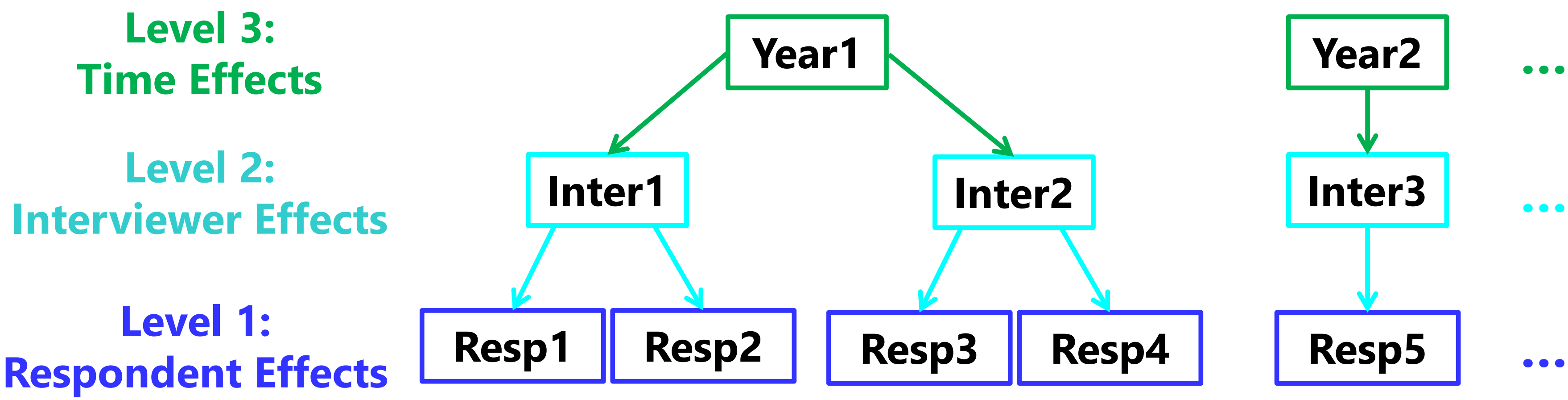
- First scenario: Predicting **time-specific respondent outcomes** for people nested in countries, collected over several years (all same people and same countries are measured over time)



- Country predictors can be included at level 3 only (no random effects)
- Person predictors should be included at levels 2 and 3 (+random over 3)
- What about effects of time-varying predictors *(stay tuned)*?
  - For People: effects should be included at all 3 levels (+random over 2 and 3)
  - For Countries: effects are only possible at levels 1 and 3 (+random over 3)

Other Types of 3-Level Designs

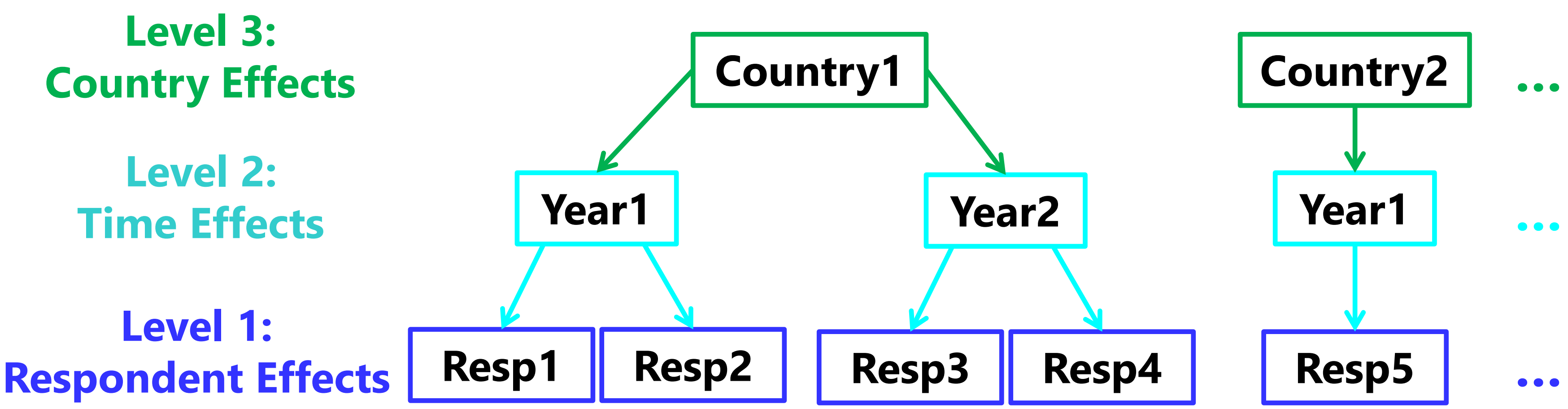
- The sampling design for the outcome (not the predictors) dictates what your levels will be, **so time may not always be level 1**
- Example: Predicting answer compliance in respondents nested in interviewers, collected over several years (all different people)



- Based on the sampling of time, time may be modeled…
  - As fixed effects in the model for the means → 2-level model instead
    - Best to use dummy codes for time if few occasions OR no time-level predictors of interest
  - As a random effect in the model for the variance → 3-level model
    - Then differences in compliance rates over time can be predicted by time-level predictors

Other Types of 3-Level Designs

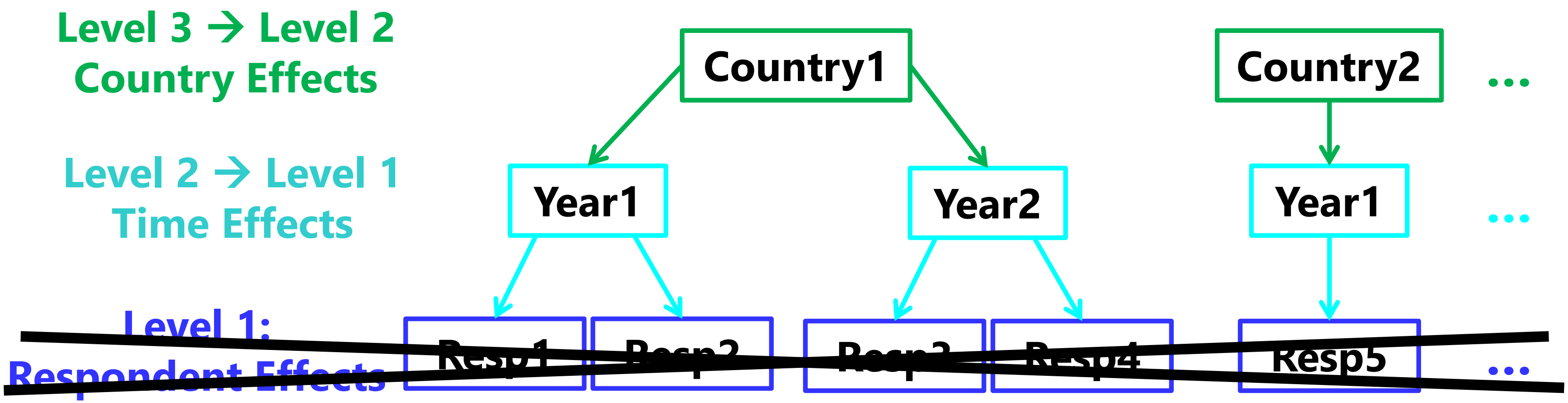
- Another scenario: Predicting **time-specific respondent outcomes** for people nested in countries, collected over several years (<u>all different people</u>, but the <u>same countries</u> measured over time)



- Before including any fixed effects of time, the dimensions of country and time are actually crossed, not nested as shown here
  - Are nested *after* controlling for which occasion is which via fixed effects (using dummy codes per mean or a time trend that describes the means)
  - Time is still a level because not all countries change the same way

3-Level Designs: Predictors vs. Outcomes

- Same scenario: What if, instead of respondent outcomes, we wanted to predict **time-varying country outcomes**?



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

- Respondents are no longer a level at all (no outcomes for them)
- So there is nothing for respondent predictors to do, except at higher levels
  - **Time-specific averages** of respondent predictors → time-level outcome variation
  - **Across time, country averages** of respondent predictors → country-level outcome variation

Empty Means, 3-Level Random Intercept Model: Example for Clustered Longitudinal Data

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

Level 1: $y_{tij} = \beta_{0ij} + e_{tij}$

Residual = time-specific deviation from **person's** predicted outcome

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

Person Random Intercept = person-specific deviation from **group's** predicted outcome

Level 3: $\delta_{00j} = \gamma_{000} + V_{00j}$

Fixed Intercept =grand mean of group means

Group Random Intercept = group-specific deviation from fixed intercept

3 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

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

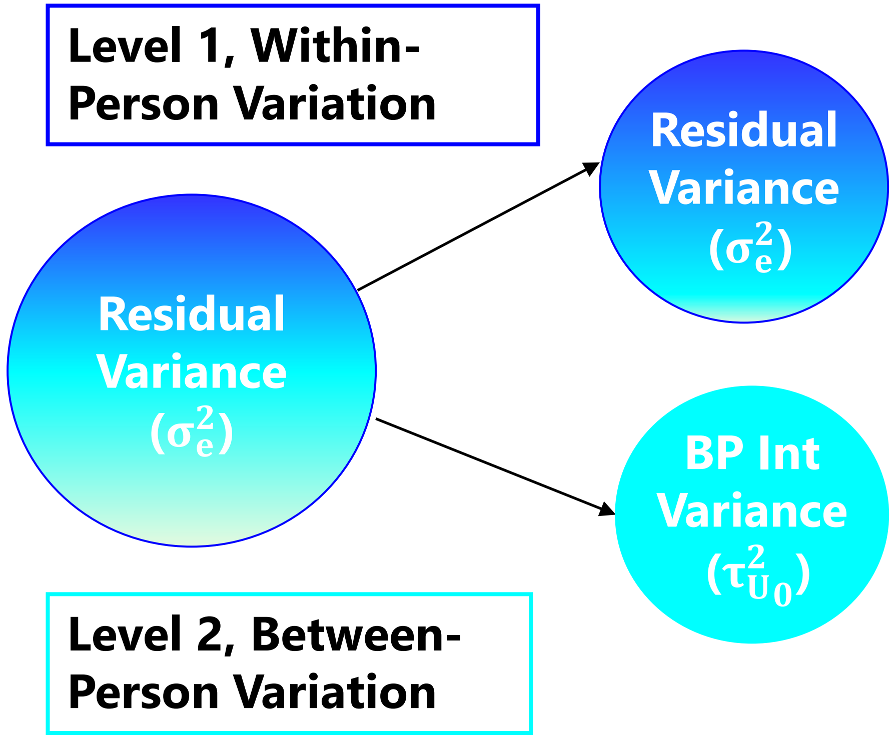
Composite equation:

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

Btw: My bad for reusing "V"

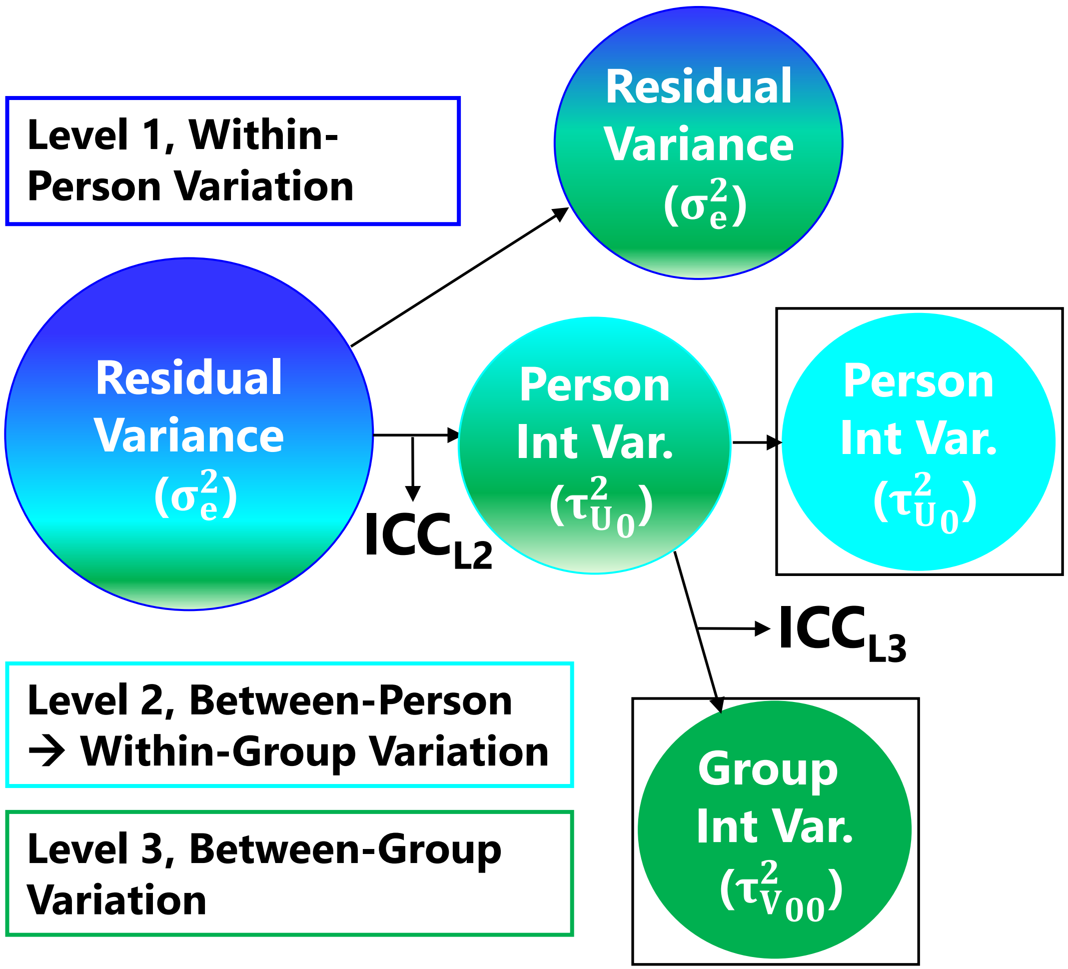
2-Level Random Intercept Model

- Where does each kind of person dependency go? Into a new random effects variance component (or "pile" of variance):
- Let's start with an empty means, random intercept 2-level model for time within person:



3-Level Random Intercept Model

- Now let's see what happens in an empty means, random intercept 3-level model of time within person within groups:



ICCs in a 3-Level Random Intercept Model Example: Time within Person within Group

- **<u>ICC for level 2 (and level 3) relative to level 1:</u>**

- $\text{ICC}_{\text{L2}} = \frac{\text{Between}-\text{Person}}{\text{Total}} = \frac{\text{L3+L2}}{\text{L3+L2+L1}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}}}{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}} + \boldsymbol{\sigma}^2_{\mathbf{e}}}$

  → This ICC expresses similarity of occasions from same person (and by definition, from the same group) → of the ***total variation in Y***, how much of it is ***between persons, or cross-sectional (not due to time)***?

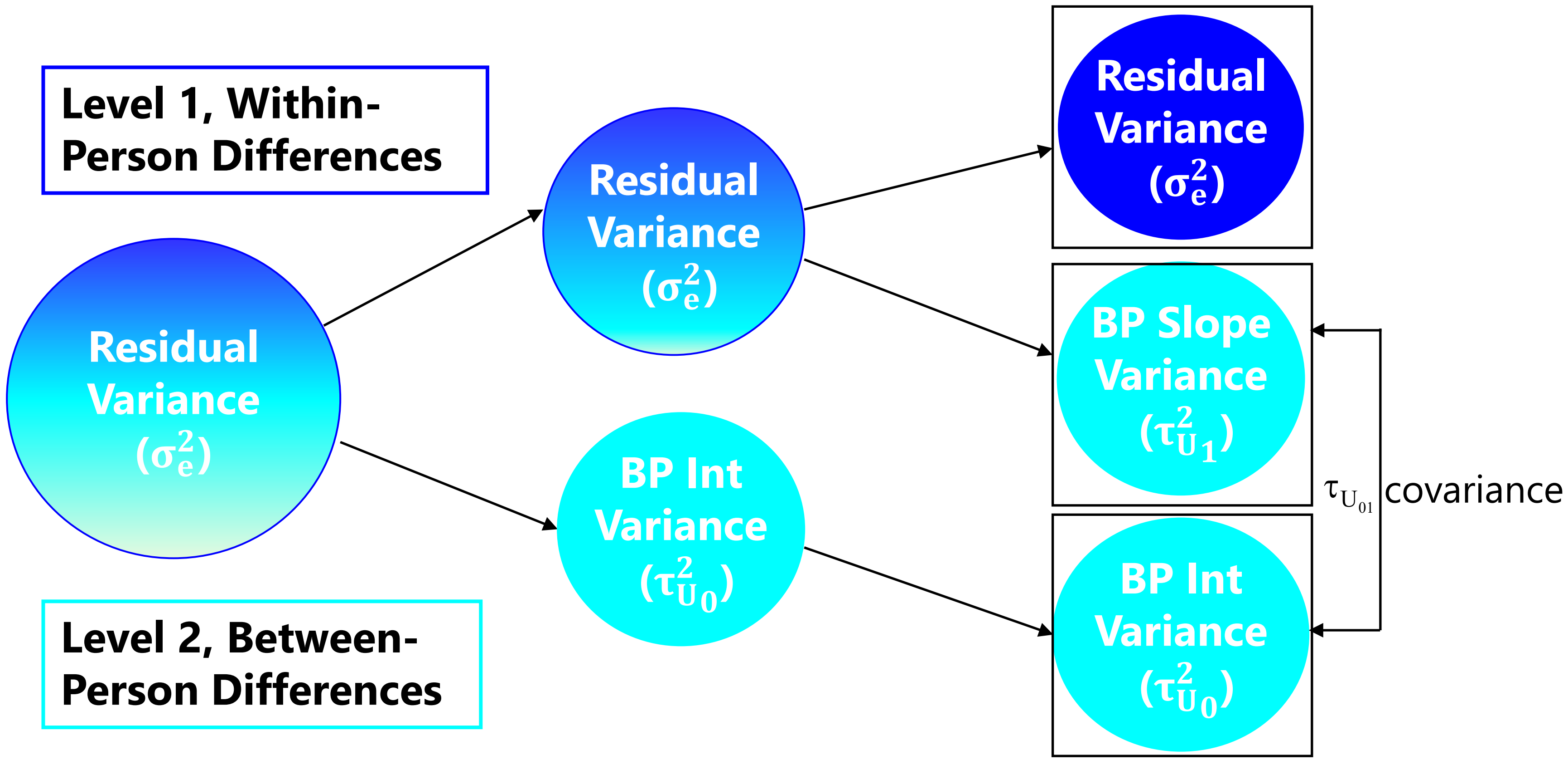
- **<u>ICC for level 3 relative to level 2 (ignoring level 1):</u>**

- $\text{ICC}_{\text{L3}} = \frac{\text{Between}-\text{Group}}{\text{Between}-\text{Person}} = \frac{\text{L3}}{\text{L3+L2}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{00}}}}{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}}}$

  → This ICC expresses similarity of persons from same group (ignoring within-person variation over time) → of ***that total between-person variation in Y***, how much of that is actually ***between groups***?

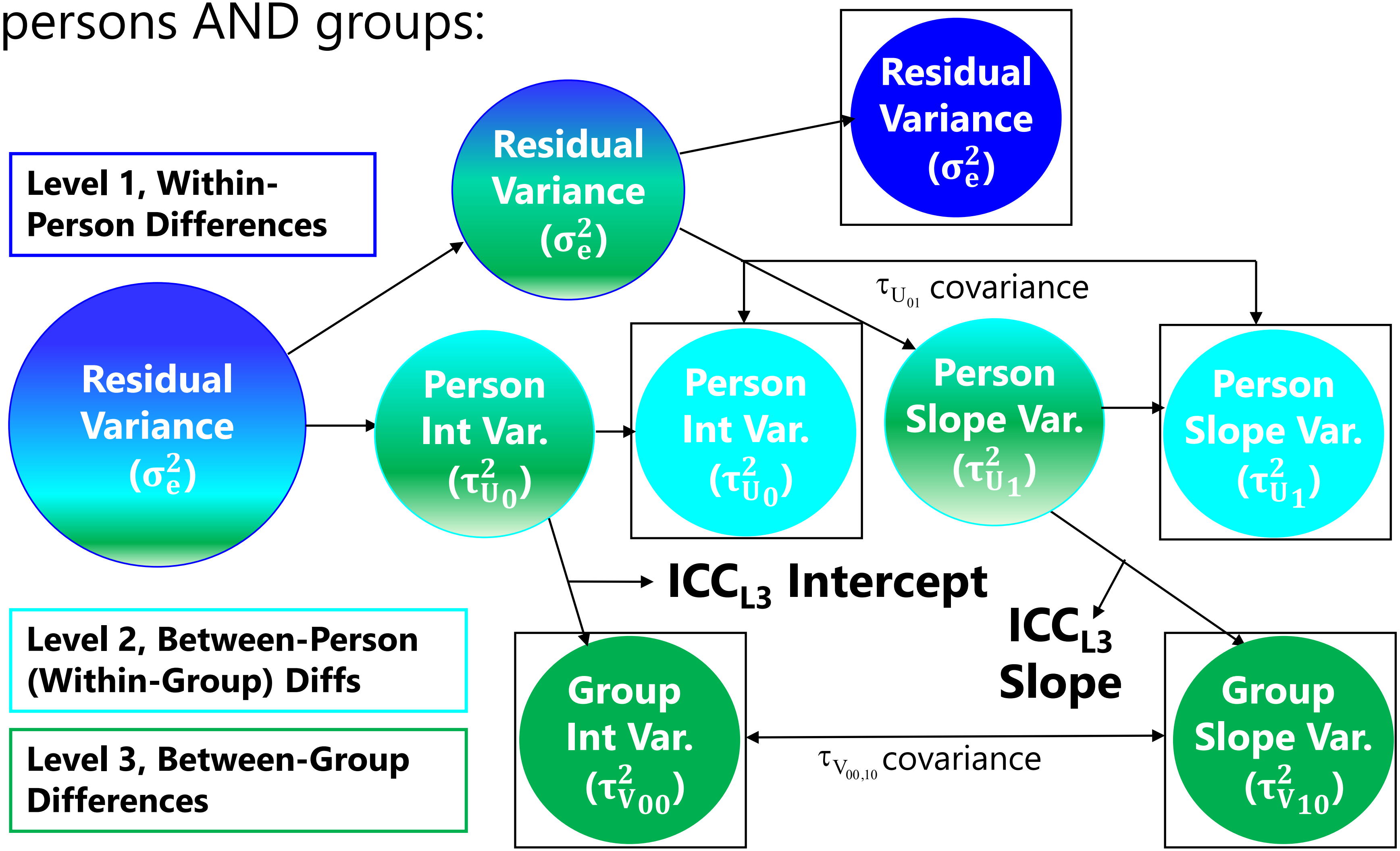
2-Level Random Slope Model

- What about time? After adding fixed effects of time, we can add random effects of time over persons in a 2-level model:



3-Level Random Slope Model

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



3-Level Random Time Slope Model

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

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

Residual = time-specific deviation from person's predicted growth line (σ_e^2)

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

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

Person Random Intercept and Slope = person-specific deviations from group's predicted intercept, slope ($\tau_{U_0}^2$, $\tau_{U_1}^2$, $\tau_{U_{01}}$)

Level 3: $\delta_{00j} = \gamma_{000} + V_{00j}$

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

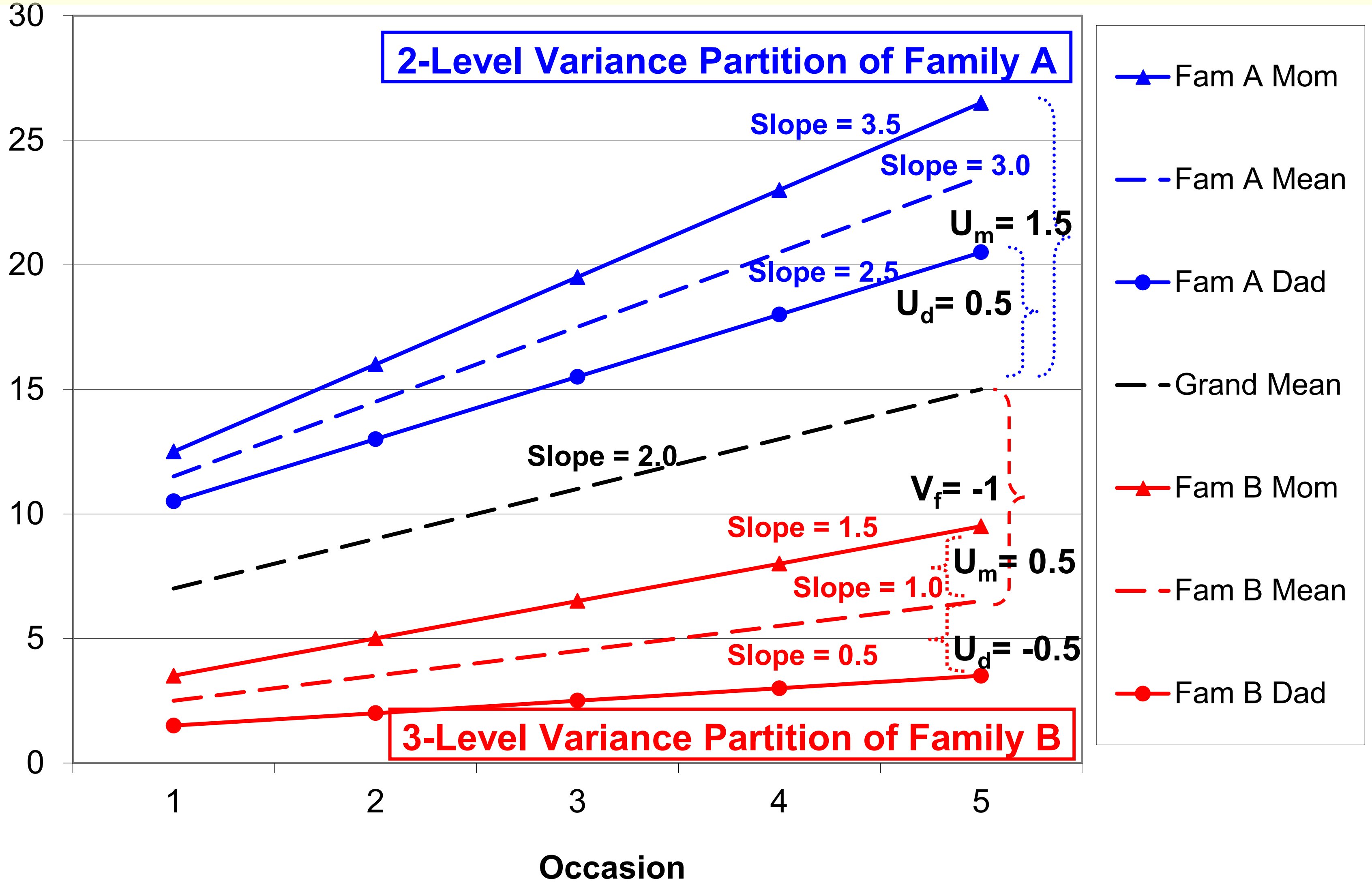
Group Random Intercept and Slope = group-specific deviations from fixed intercept, slope ($\tau_{V_{00}}^2$, $\tau_{V_{10}}^2$, $\tau_{V_{00,10}}$)

Fixed Intercept, Fixed Linear Time Slope

Composite equation (9 parameters):

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

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



ICCs for Random Intercepts and Slopes

- Once random slopes are included at both level-3 and level-2, ICCs can be computed for the random intercepts and slopes specifically (which would be the level-3 type of ICC)

$$\mathrm{ICC_{Int}} = \frac{\text{Between} - \text{Group}}{\text{Between} - \text{Person}} = \frac{\text{L3 Int}}{\text{L3 Int} + \text{L2 Int}} = \frac{\boldsymbol{\tau^2_{V_{00}}}}{\boldsymbol{\tau^2_{V_{00}} + \tau^2_{U_0}}}$$

$$\mathrm{ICC_{Slope}} = \frac{\text{Between} - \text{Group}}{\text{Between} - \text{Person}} = \frac{\text{L3 Slope}}{\text{L3 Slope} + \text{L2 Slope}} = \frac{\boldsymbol{\tau^2_{V_{10}}}}{\boldsymbol{\tau^2_{V_{10}} + \tau^2_{U_1}}}$$

- Can be computed for any level-1 slope that is random at both levels (e.g., linear and quadratic time, time-varying predictors)
- Be careful when the model is uneven across levels, though

$$\frac{\text{Random Level 2: int, linear, quad}}{\text{Random Level 3: int, linear}} \rightarrow \frac{\text{Linear is when time} = 0}{\text{Linear is at any occasion}}$$

More on Random Slopes in 3-Level Models

- Any level-1 predictor can have a random effect over level 2, level 3, or over both levels, but I recommend working your way **UP the higher levels** for assessing random effects…
  - e.g., Does the effect of time vary over level-2 persons?
  - If so, does the effect of time vary over level-3 groups, too? → Is there a commonality in how people from the same group change over time?
- … because random effects at level 3 only are possible but unlikely (e.g., means everyone in the group changes the same)

- Level-2 predictors can also have random effects over level 3
  - e.g., Does the effect of a L2 person characteristic vary over L3 groups?
- Level-1, level-2, and level-1 by level-2 cross-level interactions can all have random effects over level 3, too, in theory
  - But tread carefully! The more random effects you have, the more likely you are to have convergence problems ("**G** matrix not positive definite")

What about Time-Varying Group Membership?

- Students are nested within classes at each occasion…
- But if students move into different classes across time…
  - Time at level 1 is nested within student AND within classes
  - Students are crossed with classes at level 2
- How to model a **time-varying classroom effect**?
  - Btw, this is the basis of so-called "value-added models"

- Two example options (both via fixed or random effects):
  - **"Acute" effect**: Class effect active only when students are in that class
    - e.g., class effect ← teacher bias
    - Once a student is out of the class, class effect is no longer present
  - **"Transfer" effect**: Effect is active when in class AND in the future…
    - e.g., class effect ← differential learning
    - Effect stays with the student in the future (i.e., a "layered" value-added model)

Time (t), Students (s), and Classes (c)

- Custom-built intercepts for time-varying effects of classes
  - An intercept is usually a column of 1's, but ours will be 0's and 1's to serve as switches that turn on/off class effects

				Per-Year Class ID (−99 = missing)			Intercepts for Acute Effects			Intercepts for Transfer Effects		
Student ID	Class ID	Grade	Year	Year 0 Class	Year 1 Class	Year 2 Class	Year 0 Intercept	Year 1 Intercept	Year 2 Intercept	Year 0 Effect	Year 1 Effect	Year 2 Effect
101	1	3	0	1	−99	43	1	0	0	1	0	0
101	−99	4	1	1	−99	43	0	0	0	0	0	0
101	43	5	2	1	−99	43	0	0	1	1	0	1
102	3	3	0	3	21	42	1	0	0	1	0	0
102	21	4	1	3	21	42	0	1	0	1	1	0
102	42	5	2	3	21	42	0	0	1	1	1	1

Time (t), Students (s), and Classes (c)

- Hoffman (2015) Equation 11.3: **fixed effects model** for classroom as a categorical time-varying predictor:
  - Allows for control of classroom differences only….

$$\begin{aligned}\text{Effort}_{tsc} &= \gamma_{000} + \gamma_{100}\left(\text{Year01}_{tsc}\right) + \gamma_{200}\left(\text{Year12}_{tsc}\right) + U_{0s0} + e_{tsc} \\ &+ \gamma^{0}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int0}_{tsc}\right) + \gamma^{0}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int0}_{tsc}\right)\cdots + \gamma^{0}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int0}_{tsc}\right) \\ &+ \gamma^{1}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int1}_{tsc}\right) + \gamma^{1}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int1}_{tsc}\right)\cdots + \gamma^{1}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int1}_{tsc}\right) \\ &+ \gamma^{2}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int2}_{tsc}\right) + \gamma^{2}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int2}_{tsc}\right)\cdots + \gamma^{2}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int2}_{tsc}\right)\end{aligned}$$

- Hoffman (2015) Equation 11.4: classrooms as a random effect crossed with students (as a random effect) at level 2:
  - Controls and quantifies classroom variance so it can be predicted!

$$\begin{aligned}\text{Effort}_{tsc} &= \gamma_{000} + \gamma_{100}\left(\text{Year01}_{tsc}\right) + \gamma_{200}\left(\text{Year12}_{tsc}\right) + U_{0s0} + e_{tsc} \\ &+ U^{0}_{00c}\left(\text{Int0}_{tsc}\right) + U^{1}_{00c}\left(\text{Int1}_{tsc}\right) + U^{2}_{00c}\left(\text{Int2}_{tsc}\right)\end{aligned}$$