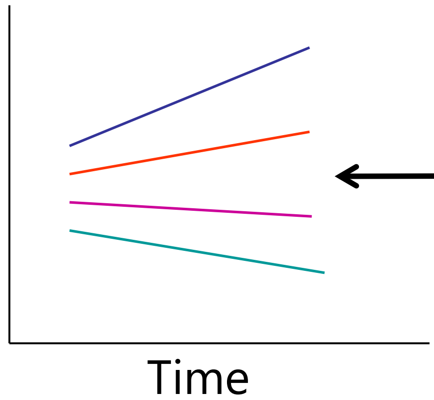
Describing Within-Person Fluctuation over Time using Alternative Covariance Structures

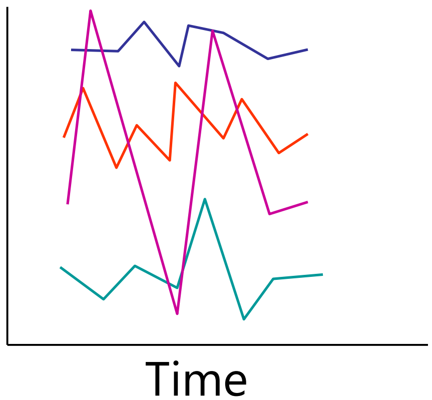
- Today's Class:
  - The Big Picture
  - ACS models using the **R** matrix only
  - Introducing the **G**, **Z**, and **V** matrices
  - ACS models combining the **G** and **R** matrices

Modeling Change vs. Fluctuation

Pure WP Change



Our focus right now → Pure WP Fluctuation



Model for the Means:

- WP Change → describe pattern of *average* change (over "time")
- **WP Fluctuation** → *may* not need anything (if no systematic change)

Model for the Variance:

- WP Change → describe *individual differences* in change (random effects)
  → this allows variances and covariances to differ over time
- **WP Fluctuation** → describe pattern of variances and covariances over time

Big Picture Framework: Models for the Variance in Longitudinal Data

Compound Symmetry (CS)

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

Most useful model: likely somewhere in between!

Unstructured (UN)

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

Parsimony ⟵⟶ **Good fit**

Univariate RM ANOVA

NAME …THAT … STRUCTURE!

Multivariate RM ANOVA

What is the pattern of variance and covariance over time?

CS and UN are just two of the many, many options available within MLM, including ***random effects models*** (for change) and ***alternative covariance structure models*** (for fluctuation).

Relative Model Fit by Model Side

- Nested models (i.e., in which one is a subset of the other) can now differ from each other in two important ways

- **Model for the Means** → which predictors and which fixed effects of them are included in the model
  - Does not require assessment of relative model fit using LL or −2LL (can still use univariate or multivariate Wald tests for this)

- **Model for the Variance** → what the pattern of variance and covariance of residuals from the same unit should be
  - DOES require assessment of relative model fit using LL or −2LL
  - Cannot use the Wald test *p*-values that show up on the output for testing significance of variances because those *p*-values are use a two-sided sampling distribution for what the variance could be (but variances cannot be negative, so those *p*-values are not valid)

Comparing Models for the Variance

- ACS models require assessment of **relative model fit**: how well does the model fit relative to other possible models?

- 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 main questions in choosing a model for the variance:**
  - How does the variance of the residuals differ across occasions?
  - How are the residuals from the same sampling unit correlated?

- 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

- What your *p*-value for the −2ΔLL test means:
  - 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

N = # level-2 units

Alternative Covariance Structure Models

- Useful in predicting patterns of variance and covariance that arise from fluctuation in the outcome across time:
  - **Variances**: Same (homogeneous) or different (heterogeneous)?
  - **Covariances**: Same or different? If different, what is the pattern?
    - Models with heterogeneous variances predict correlation instead of covariance because covariances will differ when variances differ
  - Often don't need any fixed effects for systematic effects of time in the model for the means (although this is always an empirical question)

- Limitations for most of the ACS models:
  - Require **equal-interval** occasions (if they use the idea of "time lag")
  - Require **balanced** time across persons (no intermediate time values)
  - But **do not require complete data** (unlike when CS and UN are estimated via least squares in ANOVA instead of ML/REML in MLM)

- ACS models do require some new terminology to introduce…

Two Families of ACS Models

- So far, we've referred to the variance and covariance matrix of the multivariate (longitudinal) outcomes as the **R** matrix
  - We now refer to these as **"R-only models"** (use **REPEATED** statement only)
  - Although the **R** matrix is actually specified per individual, ACS models usually assume the same **R** matrix for everyone
  - **R** matrix is symmetric with dimensions *n* x *n*, in which *n* = # occasions per person (although people can have missing data, the same set of *possible* occasions is required across people to use most **R**-only models)

- **3 other matrices we'll see in "G and R combined" ACS models:**
  - **G** = matrix of random effects variances and covariances (stay tuned)
  - **Z** = matrix of values for predictors that have random effects (stay tuned)
  - **V** = symmetric *n* x *n* matrix of **total** variance and covariance over time
    - If the model includes random effects, then **G** and **Z** get combined with **R** to make **V** as $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^{\mathrm{T}} + \mathbf{R}$ (accomplished by adding the **RANDOM** statement)
    - If the model does NOT include random effects in **G**, then $\mathbf{V} = \mathbf{R}$… *so, **R**-only*

Review: Covariances and Correlations

$$\text{Correlation}_{y1,y2} = \frac{\text{Covariance}_{y1,y2}}{\sqrt{\text{Variance}_{y1}} * \sqrt{\text{Variance}_{y2}}}$$

$$\text{Covariance}_{y1,y2} = \text{Correlation}_{y1,y2} * \sqrt{\text{Variance}_{y1}} * \sqrt{\text{Variance}_{y2}}$$

- Given the standard deviation (as $\sqrt{\text{Variance}}$) at each occasion, either the correlation and covariance can be calculated given the other
- ACS models with **homogeneous variances** tend to be specified in terms of **variance and covariance**
  - Given same variance over time, same covariance → same correlation
- ACS models with **heterogeneous variance** tend to be specified in terms of **variance and correlation**
  - Different variances over time → different covariances over time, <u>even if the correlation is the same</u> (so only correlation is estimated directly)

R-Only ACS Models

- The **R-only** models to be presented next are all specified using the **REPEATED** statement only (no RANDOM statement)
- They are explained by showing their predicted **R** matrix, which provides the **total** variances and covariances across occasions
  - Total variance per occasion on diagonal
  - Total covariances across occasions on off-diagonals
  - I've included in " " the labels SAS uses for each parameter
- Correlations across occasions can be calculated given variances and covariances, which would be shown in the **RCORR** matrix (available in SAS PROC MIXED)
  - 1's on diagonal (standardized variables), correlations on off-diagonal
- **Unstructured (TYPE=UN) will always fit best by −2LL**
  - All ACS models are nested within Unstructured (UN = the data)
  - Goal: find an ACS model that is **simpler** but **not worse fitting** than UN

R-Only ACS Models: CS/CSH

- **Compound Symmetry: TYPE=CS**
  - 2 parameters:
    - **1 "residual" variance σ_e^2**
    - **1 "CS" covariance** across occasions
  - Constant total variance: $CS + \sigma_e^2$
  - Constant total covariance: CS

$$\begin{bmatrix} CS+\sigma_e^2 & CS & CS & CS \\ CS & CS+\sigma_e^2 & CS & CS \\ CS & CS & CS+\sigma_e^2 & CS \\ CS & CS & CS & CS+\sigma_e^2 \end{bmatrix}$$

- **Compound Symmetry Heterogeneous: TYPE=CSH**
  - *n*+1 parameters:
    - ***n* separate "Var(*n*)" total variances σ_{Tn}^2**
    - **1 "CSH" total correlation across occasions**
  - Separate total variances are estimated directly
  - Still constant total correlation: CSH (but has non-constant covariances)

$$\begin{bmatrix} \sigma_{T1}^2 & CSH\sigma_{T1}\sigma_{T2} & CSH\sigma_{T1}\sigma_{T3} & CSH\sigma_{T1}\sigma_{T4} \\ CSH\sigma_{T2}\sigma_{T1} & \sigma_{T2}^2 & CSH\sigma_{T2}\sigma_{T3} & CSH\sigma_{T2}\sigma_{T4} \\ CSH\sigma_{T3}\sigma_{T1} & CSH\sigma_{T3}\sigma_{T2} & \sigma_{T3}^2 & CSH\sigma_{T3}\sigma_{T4} \\ CSH\sigma_{T4}\sigma_{T1} & CSH\sigma_{T4}\sigma_{T2} & CSH\sigma_{T4}\sigma_{T3} & \sigma_{T4}^2 \end{bmatrix}$$

R-Only ACS Models: AR1/ARH1

- **1^{st} Order Auto-Regressive: TYPE=AR(1)**
  - 2 parameters:
    - **1 constant total variance σ_T^2 (mislabeled "residual")**
    - **1 "AR1" total auto-correlation r_T** across occasions
    - r_T^1 is lag-1 correlation, r_T^2 is lag-2 correlation, r_T^3 is lag-3 correlation….

$$\begin{bmatrix} \sigma_T^2 & r_T^1\sigma_T^2 & r_T^2\sigma_T^2 & r_T^3\sigma_T^2 \\ r_T^1\sigma_T^2 & \sigma_T^2 & r_T^1\sigma_T^2 & r_T^2\sigma_T^2 \\ r_T^2\sigma_T^2 & r_T^1\sigma_T^2 & \sigma_T^2 & r_T^1\sigma_T^2 \\ r_T^3\sigma_T^2 & r_T^2\sigma_T^2 & r_T^1\sigma_T^2 & \sigma_T^2 \end{bmatrix}$$

- **1^{st} Order Auto-Regressive Heterogeneous: TYPE=ARH(1)**
  - *n*+1 parameters:
    - ***n* separate "Var(*n*)" total variances σ_{Tn}^2**
    - **1 "ARH1" total auto-correlation r_T** across occasions
    - r_T^1 is lag-1 correlation, r_T^2 is lag-2 correlation, r_T^3 is lag-3 correlation….

$$\begin{bmatrix} \sigma_{T1}^2 & r_T^1\sigma_{T1}\sigma_{T2} & r_T^2\sigma_{T1}\sigma_{T3} & r_T^3\sigma_{T1}\sigma_{T4} \\ r_T^1\sigma_{T2}\sigma_{T1} & \sigma_{T2}^2 & r_T^1\sigma_{T2}\sigma_{T3} & r_T^2\sigma_{T2}\sigma_{T4} \\ r_T^2\sigma_{T3}\sigma_{T1} & r_T^1\sigma_{T3}\sigma_{T2} & \sigma_{T3}^2 & r_T^1\sigma_{T3}\sigma_{T4} \\ r_T^3\sigma_{T4}\sigma_{T1} & r_T^2\sigma_{T4}\sigma_{T2} & r_T^1\sigma_{T4}\sigma_{T3} & \sigma_{T4}^2 \end{bmatrix}$$

R-Only ACS Models: TOEP*n*/TOEPH*n*

- **Toeplitz(*n*): TYPE=TOEP(*n*)**
  - *n* parameters:
    - **1 constant total variance σ^2_T (mislabeled "residual")**
    - ***n*−1 "TOEP(*lag*)" c_{Tn} banded total covariances** across occasions
    - c_{T1} is lag-1 covariance, c_{T2} is lag-2 covariance, c_{T3} is lag-3 covariance….

$$\begin{bmatrix} \sigma^2_T & & & \\ c_{T1} & \sigma^2_T & & \\ c_{T2} & c_{T1} & \sigma^2_T & \\ c_{T3} & c_{T2} & c_{T1} & \sigma^2_T \end{bmatrix}$$

- **Toeplitz Heterogeneous(*n*): TYPE=TOEPH(*n*)**
  - *n* + (*n*−1) parameters:
    - ***n* separate "Var(*n*)" total variances σ^2_{Tn}**
    - ***n*−1 "TOEPH(*lag*)" r_{Tn} banded total correlations** across occasions
    - $r_{T1}$is lag-1 correlation, r_{T2} is lag-2 correlation, r_{T3} is lag-3 correlation….

$$\begin{bmatrix} \sigma^2_{T1} & r_{T1}\sigma_{T1}\sigma_{T2} & r_{T2}\sigma_{T1}\sigma_{T3} & r_{T3}\sigma_{T1}\sigma_{T4} \\ r_{T1}\sigma_{T2}\sigma_{T1} & \sigma^2_{T2} & r_{T1}\sigma_{T2}\sigma_{T3} & r_{T2}\sigma_{T2}\sigma_{T4} \\ r_{T2}\sigma_{T3}\sigma_{T1} & r_{T1}\sigma_{T3}\sigma_{T2} & \sigma^2_{T3} & r_{T1}\sigma_{T3}\sigma_{T4} \\ r_{T3}\sigma_{T4}\sigma_{T1} & r_{T2}\sigma_{T4}\sigma_{T2} & r_{T1}\sigma_{T4}\sigma_{T3} & \sigma^2_{T4} \end{bmatrix}$$

Comparing **R**-only ACS Models

- Baseline models: **CS =simplest**, **UN = most complex**
  - Relative to CS, more complex models fit "better" or "not better"
  - Relative to UN, less complex models fit "worse" or "not worse"

- Other rules of nesting and model comparisons:
  - Homogeneous variance models are nested within heterogeneous variance models (e.g., CS in CSH, AR1 in ARH1, TOEP in TOEPH)
  - CS and AR1 are each nested within TOEP (i.e., TOEP can become CS or AR1 through restrictions of its covariance patterns)
  - CS and AR1 are not nested (because both have 2 parameters)
  - **R**-only models differ in unbounded parameters, so can be compared using regular $-2\Delta LL$ tests (instead of mixture $-2\Delta LL$ tests)
  - Good idea to start by assuming heterogeneous variances until you settle on the covariance pattern, then test if het. var. are still necessary
  - When in doubt, just compare AIC and BIC (useful even with $-2\Delta LL$ tests)

The Other Family of ACS Models

- **R**-only models *directly* predict the **total** variance and covariance
- **G** and **R** models *indirectly* predict the total variance and covariance through **between-person (BP)** and **within-person (WP)** sources of variance and covariance → So, for this model: $\mathbf{y_{ti} = \beta_0 + U_{0i} + e_{ti}}$
  - **BP** = **G** matrix of **level-2 random effect ($\mathbf{U_{0i}}$)** variances and covariances
    - Which effects get to be random (whose variance and covariances are then included in **G**) is specified using the **RANDOM** statement (always TYPE=UN)
    - Our ACS models have a random intercept only, so **G** is 1x1 scalar of $[\boldsymbol{\tau}^2_{\mathbf{U_0}}]$
  - **WP** = **R** matrix of **level-1 ($\mathbf{e_{ti}}$) residual** variances and covariances
    - The *n* x *n* **R** matrix of **residual** variances and covariances **that remain** after controlling for random intercept variance is then modeled with **REPEATED**
  - **Total** = **V** = *n* x *n* matrix of **total** variance and covariance over time that results from putting **G** and **R** together: $\mathbf{V = ZGZ^T + R}$
    - **Z** is a matrix that holds the values of predictors with random effects, but **Z** will be an *n* x 1 column of 1's for now (random intercept only)

A "Random Intercept" (**G** and **R**) Model

Total Predicted Data Matrix is called V Matrix

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

Level 2, BP Variance

Unstructured **G Matrix**

(RANDOM statement)

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

Random Intercept Variance only

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

Level 1, WP Variance

Diagonal (VC) **R Matrix**

(REPEATED statement)

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

Residual Variance only

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

CS as a "Random Intercept" Model

<u>RI and DIAG</u>: Total predicted data matrix is called V matrix, created from the G [TYPE=UN] and R [TYPE=VC] matrices as follows:

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

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

Does the end result V look familiar? It should: CS = τ^2_{U0}

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

So if the **R-only CS model** (the simplest baseline) can be specified equivalently using **G and R**, can we do the same for the **R-only UN model** (the most complex baseline)?

Absolutely! ...*with one small catch*

UN via a "Random Intercept" Model

<u>RI and UN*n*−1</u>: Total predicted data matrix is called V matrix, created from the G [TYPE=UN] and R [TYPE=UN(*n*−1)] matrices as follows:

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_{e1} & \sigma_{e12} & \sigma_{e13} & 0\\ \sigma_{e21} & \sigma^2_{e2} & \sigma_{e23} & \sigma_{e24}\\ \sigma_{e31} & \sigma_{e32} & \sigma^2_{e3} & \sigma_{e34}\\ 0 & \sigma_{e42} & \sigma_{e43} & \sigma^2_{e4}\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_{e1} & \tau^2_{U_0}+\sigma_{e12} & \tau^2_{U_0}+\sigma_{e13} & \tau^2_{U_0}\\ \tau^2_{U_0}+\sigma_{e21} & \tau^2_{U_0}+\sigma^2_{e2} & \tau^2_{U_0}+\sigma_{e23} & \tau^2_{U_0}+\sigma_{e24}\\ \tau^2_{U_0}+\sigma_{e31} & \tau^2_{U_0}+\sigma_{e32} & \tau^2_{U_0}+\sigma^2_{e3} & \tau^2_{U_0}+\sigma_{e34}\\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma_{e42} & \tau^2_{U_0}+\sigma_{e43} & \tau^2_{U_0}+\sigma^2_{e4}\end{bmatrix}$$

This **RI and UN*n*−1 model** is equivalent to (makes same predictions as) the **R-only UN model**. But it shows the *residual* (not total) covariances.

Because we can't estimate all possible variances and covariances in the **R** matrix and also estimate the random intercept variance $\tau^2_{U_0}$ in the **G** matrix, <u>we are eliminating the last **R** matrix covariance by setting it to 0</u>.

Accordingly, in the **RI and UN*n*−1** model, the random intercept variance $\tau^2_{U_0}$ takes on the value of the covariance for the first and last occasions.

Rationale for **G** and **R** ACS models

- Modeling WP fluctuation traditionally involves using **R** only (no **G**) → **Total** BP + WP variance described by just **R** matrix (so **R**=**V**)
  - Correlations would still be expected even at distant time lags because of constant individual differences (i.e., the BP random intercept)
  - Resulting **R**-only model may require lots of estimated parameters as a result e.g., 8 time points? Pry need a 7-lag Toeplitz(8) model

- **Why not take out the primary reason for the covariance across occasions (the random intercept variance) and see what's left?**
  - Random intercept variance $\boldsymbol{\tau}_{\mathbf{U}_0}^2$ in **G** → control for person mean differences
  - THEN predict just the **residual** variance/covariance in **R**, not the **total**
  - Resulting model may be more parsimonious (e.g., maybe only lag1 or lag2 occasions are still related after removing $\boldsymbol{\tau}_{\mathbf{U}_0}^2$ as a source of covariance)
  - Has the advantage of still distinguishing BP from WP variance (useful for descriptive purposes and for calculating effect sizes later)

Random Intercept + Diagonal **R** Models

RI and DIAG: **V** is created from **G [TYPE=UN] and R [TYPE=VC]:**
homogeneous *residual variances;* ***no*** *residual covariances*

Same fit as R-only CS

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

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

RI and DIAGH: **V** is created from **G [TYPE=UN] and R [TYPE=UN(1)]:**
heterogeneous *residual variances;* ***no*** *residual covariances*

NOT same fit as R-only CSH

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_{e1} & 0 & 0 & 0\\ 0 & \sigma^2_{e2} & 0 & 0\\ 0 & 0 & \sigma^2_{e3} & 0\\ 0 & 0 & 0 & \sigma^2_{e4}\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_{e1} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_{e2} & \tau^2_{U_0} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_{e3} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+\sigma^2_{e4}\end{bmatrix}$$

Random Intercept + AR1 **R** Models

RI and AR1: **V** is created from **G [TYPE=UN] and R [TYPE=AR(1)]:**
homogeneous *residual variances;* ***auto-regressive lagged*** *residual covariances*

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_e & r^1_e\sigma^2_e & r^2_e\sigma^2_e & r^3_e\sigma^2_e \\ r^1_e\sigma^2_e & \sigma^2_e & r^1_e\sigma^2_e & r^2_e\sigma^2_e \\ r^2_e\sigma^2_e & r^1_e\sigma^2_e & \sigma^2_e & r^1_e\sigma^2_e \\ r^3_e\sigma^2_e & r^2_e\sigma^2_e & r^1_e\sigma^2_e & \sigma^2_e\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+r^1_e\sigma^2_e & \tau^2_{U_0}+r^2_e\sigma^2_e & \tau^2_{U_0}+r^3_e\sigma^2_e \\ \tau^2_{U_0}+r^1_e\sigma^2_e & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+r^1_e\sigma^2_e & \tau^2_{U_0}+r^2_e\sigma^2_e \\ \tau^2_{U_0}+r^2_e\sigma^2_e & \tau^2_{U_0}+r^1_e\sigma^2_e & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+r^1_e\sigma^2_e \\ \tau^2_{U_0}+r^3_e\sigma^2_e & \tau^2_{U_0}+r^2_e\sigma^2_e & \tau^2_{U_0}+r^1_e\sigma^2_e & \tau^2_{U_0}+\sigma^2_e\end{bmatrix}$$

RI and ARH1: **V** is created from **G [TYPE=UN] and R [TYPE=ARH(1)]:**
heterogeneous *residual variances;* ***auto-regressive lagged*** *residual covariances*

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_{e1} & r^1_e\sigma_{e1}\sigma_{e2} & r^2_e\sigma_{e1}\sigma_{e3} & r^3_e\sigma_{e1}\sigma_{e4} \\ r^1_e\sigma_{e2}\sigma_{e1} & \sigma^2_{e2} & r^1_e\sigma_{e2}\sigma_{e3} & r^2_e\sigma_{e2}\sigma_{e4} \\ r^2_e\sigma_{e3}\sigma_{e1} & r^1_e\sigma_{e3}\sigma_{e2} & \sigma^2_{e3} & r^1_e\sigma_{e3}\sigma_{e4} \\ r^3_e\sigma_{e4}\sigma_{e1} & r^2_e\sigma_{e4}\sigma_{e2} & r^1_e\sigma_{e4}\sigma_{e3} & \sigma^2_{e4}\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_{e1} & \tau^2_{U_0}+r^1_e\sigma_{e1}\sigma_{e2} & \tau^2_{U_0}+r^2_e\sigma_{e1}\sigma_{e3} & \tau^2_{U_0}+r^3_e\sigma_{e1}\sigma_{e4} \\ \tau^2_{U_0}+r^1_e\sigma_{e2}\sigma_{e1} & \tau^2_{U_0}+\sigma^2_{e2} & \tau^2_{U_0}+r^1_e\sigma_{e2}\sigma_{e3} & \tau^2_{U_0}+r^2_e\sigma_{e2}\sigma_{e4} \\ \tau^2_{U_0}+r^2_e\sigma_{e3}\sigma_{e1} & \tau^2_{U_0}+r^1_e\sigma_{e3}\sigma_{e2} & \tau^2_{U_0}+\sigma^2_{e3} & \tau^2_{U_0}+r^1_e\sigma_{e3}\sigma_{e4} \\ \tau^2_{U_0}+r^3_e\sigma_{e4}\sigma_{e1} & \tau^2_{U_0}+r^2_e\sigma_{e4}\sigma_{e2} & \tau^2_{U_0}+r^1_e\sigma_{e4}\sigma_{e3} & \tau^2_{U_0}+\sigma^2_{e4}\end{bmatrix}$$

Random Intercept + TOEP$n-1$ **R** Models

<u>RI and TOEP*n*−*1*</u>: V is created from **G [TYPE=UN] and R [TYPE=TOEP(*n*−*1*)]:**
homogeneous *residual variances;* ***banded*** *residual covariances*

Same fit as R-only TOEP(*n*)

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_e & c_{e1} & c_{e2} & 0\\ c_{e1} & \sigma^2_e & c_{e1} & c_{e2}\\ c_{e2} & c_{e1} & \sigma^2_e & c_{e1}\\ 0 & c_{e2} & c_{e1} & \sigma^2_e\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+c_{e2} & \tau^2_{U_0}\\ \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+c_{e2}\\ \tau^2_{U_0}+c_{e2} & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+c_{e1}\\ \tau^2_{U_0} & \tau^2_{U_0}+c_{e2} & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+\sigma^2_e\end{bmatrix}$$

Because of $\tau^2_{U_0}$, highest lag covariance in **R** must be set to 0 for model to be identified

<u>RI and TOEPH*n*−*1*</u>: V is created from **G [TYPE=UN] and R [TYPE=TOEPH(*n*−*1*)]:**
heterogeneous *residual variances;* ***banded*** *residual covariances*

NOT same fit as R-only TOEPH(*n*)

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_{e1} & r_{e1}\sigma_{e1}\sigma_{e2} & r_{e2}\sigma_{e1}\sigma_{e3} & 0\\ r_{e1}\sigma_{e2}\sigma_{e1} & \sigma^2_{e2} & r_{e1}\sigma_{e2}\sigma_{e3} & r_{e2}\sigma_{e2}\sigma_{e4}\\ r_{e2}\sigma_{e3}\sigma_{e1} & r_{e1}\sigma_{e3}\sigma_{e2} & \sigma^2_{e3} & r_{e1}\sigma_{e3}\sigma_{e4}\\ 0 & r_{e2}\sigma_{e4}\sigma_{e2} & r_{e1}\sigma_{e4}\sigma_{e3} & \sigma^2_{e4}\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_{e1} & \tau^2_{U_0}+r_{e1}\sigma_{e1}\sigma_{e2} & \tau^2_{U_0}+r_{e2}\sigma_{e1}\sigma_{e3} & \tau^2_{U_0}\\ \tau^2_{U_0}+r_{e1}\sigma_{e2}\sigma_{e1} & \tau^2_{U_0}+\sigma^2_{e2} & \tau^2_{U_0}+r_{e1}\sigma_{e2}\sigma_{e3} & \tau^2_{U_0}+r_{e2}\sigma_{e2}\sigma_{e4}\\ \tau^2_{U_0}+r_{e2}\sigma_{e3}\sigma_{e1} & \tau^2_{U_0}+r_{e1}\sigma_{e3}\sigma_{e2} & \tau^2_{U_0}+\sigma^2_{e3} & \tau^2_{U_0}+r_{e1}\sigma_{e3}\sigma_{e4}\\ \tau^2_{U_0} & \tau^2_{U_0}+r_{e2}\sigma_{e4}\sigma_{e2} & \tau^2_{U_0}+r_{e1}\sigma_{e4}\sigma_{e3} & \tau^2_{U_0}+\sigma^2_{e4}\end{bmatrix}$$

Random Intercept + TOEP*2* **R** Models

<u>RI and TOEP*2*</u>: V is created from **G [TYPE=UN] and R [TYPE=TOEP(*2*)]:**
homogeneous *residual variances;* ***banded*** *residual covariance at* ***lag1*** *only*

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_e & c_{e1} & 0 & 0\\ c_{e1} & \sigma^2_e & c_{e1} & 0\\ 0 & c_{e1} & \sigma^2_e & c_{e1}\\ 0 & 0 & c_{e1} & \sigma^2_e\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0} & \tau^2_{U_0}\\ \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+\sigma^2_e & \tau^2_{U_0}+c_{e1}\\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+c_{e1} & \tau^2_{U_0}+\sigma^2_e\end{bmatrix}$$

Now we can test the need for residual covariances at higher lags

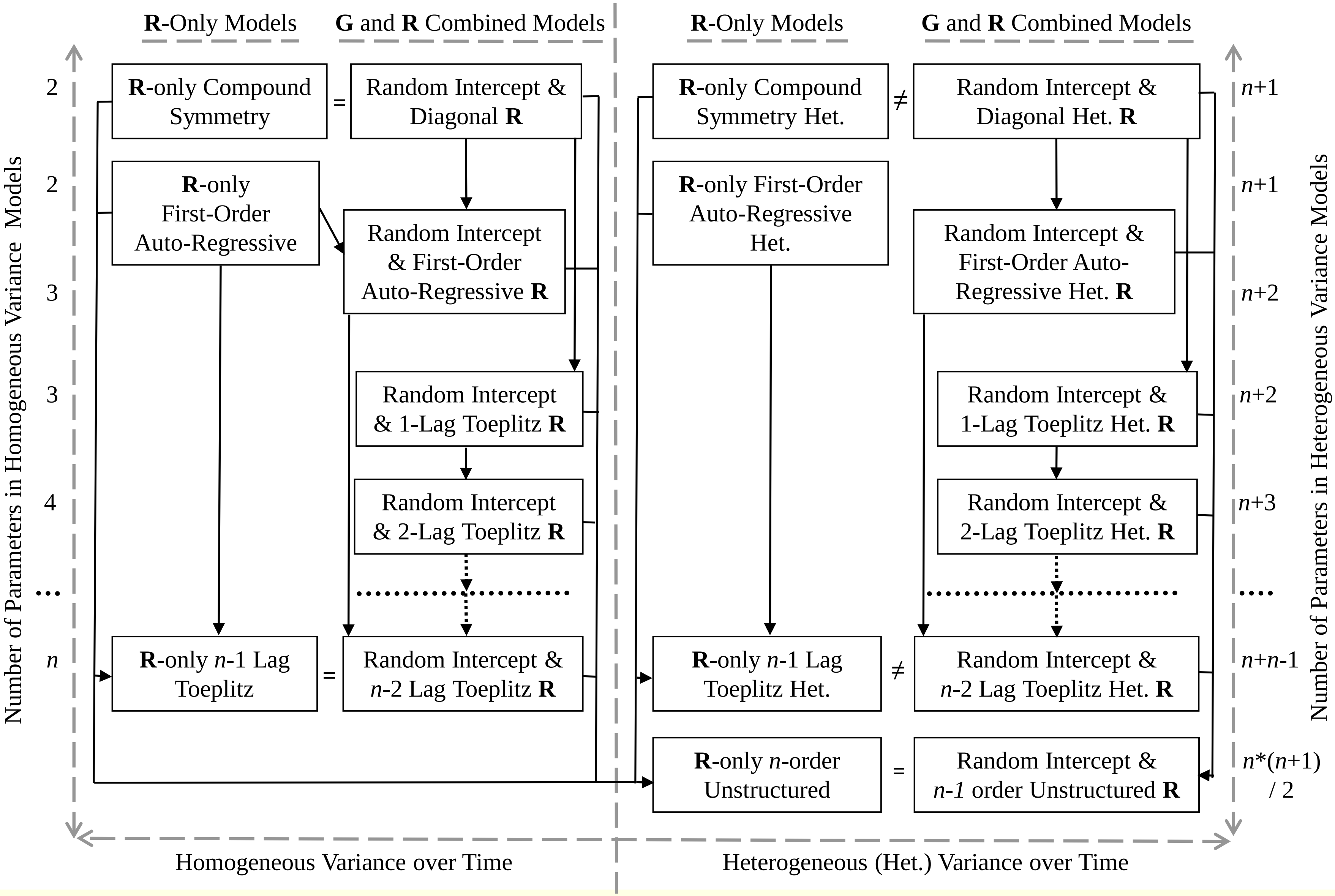
<u>RI and TOEPH*1*</u>: V is created from **G [TYPE=UN] and R [TYPE=TOEPH(*2*)]:**
heterogeneous *residual variances;* ***banded*** *residual covariance at* ***lag1*** *only*

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

$$\mathbf{V} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}\tau^2_{U_0}\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} + \begin{bmatrix}\sigma^2_{e1} & r_{e1}\sigma_{e1}\sigma_{e2} & 0 & 0\\ r_{e1}\sigma_{e2}\sigma_{e1} & \sigma^2_{e2} & r_{e1}\sigma_{e2}\sigma_{e3} & 0\\ 0 & r_{e1}\sigma_{e3}\sigma_{e2} & \sigma^2_{e3} & r_{e1}\sigma_{e3}\sigma_{e4}\\ 0 & 0 & r_{e1}\sigma_{e4}\sigma_{e3} & \sigma^2_{e4}\end{bmatrix} = \begin{bmatrix}\tau^2_{U_0}+\sigma^2_{e1} & \tau^2_{U_0}+r_{e1}\sigma_{e1}\sigma_{e2} & \tau^2_{U_0} & \tau^2_{U_0}\\ \tau^2_{U_0}+r_{e1}\sigma_{e2}\sigma_{e1} & \tau^2_{U_0}+\sigma^2_{e2} & \tau^2_{U_0}+r_{e1}\sigma_{e2}\sigma_{e3} & \tau^2_{U_0}\\ \tau^2_{U_0} & \tau^2_{U_0}+r_{e1}\sigma_{e3}\sigma_{e2} & \tau^2_{U_0}+\sigma^2_{e3} & \tau^2_{U_0}+r_{e1}\sigma_{e3}\sigma_{e4}\\ \tau^2_{U_0} & \tau^2_{U_0} & \tau^2_{U_0}+r_{e1}\sigma_{e4}\sigma_{e3} & \tau^2_{U_0}+\sigma^2_{e4}\end{bmatrix}$$

Map of **R**-only and **G** and **R** ACS Models

Arrows indicate nesting (end is more complex model)



Stuff to Watch Out For…

- **If using a random intercept, don't forget to drop 1 parameter in:**
  - ***n*-1 order UN R**: Can't get all possible elements in **R**, plus $\tau^2_{U_0}$ in **G**
  - **TOEP*n*−1**: Have to eliminate last lag covariance

- If using a random intercept…
  - Can't do RI + CS **R**: Can't get a constant in **R**, and then another constant in **G**
  - Can often test if random intercept helps (e.g., AR1 is nested within RI + AR1)

- If "**time**" is treated as **continuous** in the fixed effects, you will need another variable for **time** that is **categorical** to use in the syntax:
  - "Continuous Time" → on MODEL statement
  - "Categorical Time" → on CLASS and REPEATED statements

- Most alternative covariance structure models assume **time is balanced across persons with equal intervals across occasions**
  - If not, holding correlations of same lag equal doesn't make sense
  - Other structures can be used for unbalanced time
    - SP(POW)(time) = AR1 for unbalanced time (see SAS REPEATED statement for others)

Summary: Two Families of ACS Models

- **R**-only models:
  - Specify **R** model on REPEATED statement without any random effects variances in **G** (so no RANDOM statement is used)
  - Include UN, CS, CSH, AR1, AR1H, TOEP*n*, TOEPH*n* (among others)
  - *Total* variance and *total* covariance kept in **R**, so **R** = **V**
  - Other than CS, does not partition total variance into BP vs. WP

- **G** and **R** combined models (so **G** and **R** → **V**):
  - Specify random intercept variance $\tau^2_{U_0}$ in **G** using RANDOM statement, then specify **R** model using REPEATED statement
  - **G** matrix = Level-2 BP variance and covariance due to U_{0i}, so **R** = Level-1 WP variance and covariance of the e_{ti} residuals
  - **R** models what's left after accounting for mean differences between persons (via the random intercept variance $\tau^2_{U_0}$ in **G**)

Syntax for Models for the Variance

- Does your model include r**andom intercept variance** $\tau^2_{U_0}$ **(for U_{0i})** ?
  - Use the **RANDOM** statement → **G matrix**
  - Random intercept models BP interindividual differences in mean Y
- What about **residual variance** σ^2_e **(for e_{ti})** ?
  - Use the **REPEATED** statement → **R matrix**
    - **WITHOUT a RANDOM statement: R is BP and WP variance together** = σ^2_T
      → Total variances and covariances (to model all variation, so **R** = **V**)
    - **WITH a RANDOM statement: R is WP variance only =** σ^2_e
      → Residual variances and covariances to model WP intraindividual variation
      → **G** and **R** put back together = **V matrix** of total variances and covariances

- The **REPEATED** statement is always there implicitly…
  - Any model **always** has at least one residual variance in **R** matrix
- But the **RANDOM** statement is only there if you write it
  - **G** matrix isn't always necessary (don't always need random intercept)

Wrapping Up: ACS Models

- Even if you just expect fluctuation over time rather than change, you still should be concerned about accurately predicting the variances and covariances across occasions

- Baseline models (from ANOVA least squares) are CS & UN:
  - Compound Symmetry: Equal variance and covariance over time
  - Unstructured: All variances & covariances estimated separately
  - CS and UN via ML or REML estimation allows missing data

- MLM gives us choices in the middle
  - Goal: Get as close to UN as parsimoniously as possible
  - **R**-only: Structure TOTAL variation in one matrix (**R** only)
  - **G**+**R**: Put constant covariance due to random intercept in **G**, then structural RESIDUAL covariance in **R** (so that **G** and **R** → **V** TOTAL)