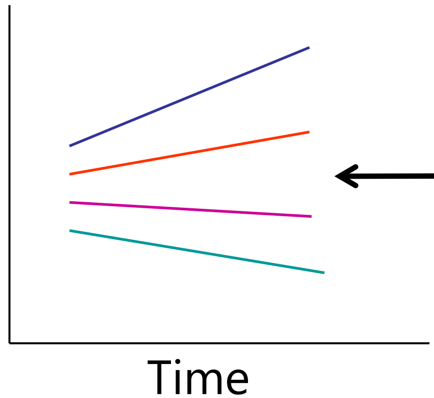
Describing Within-Person Fluctuation over Time using Alternative Covariance Structures

- Topics:
  - 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

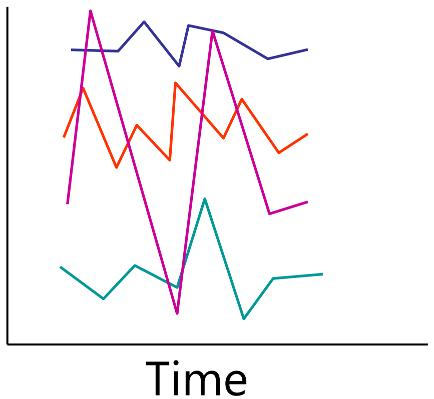
Modeling Change vs. Fluctuation

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 −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 −2LL
  - Cannot use any Wald test *p*-values that may be in the output for testing significance of variances because those *p*-values use a two-sided sampling distribution for what the variance could be (but variances cannot be negative, so those *p*-values are not valid)

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, R, 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?
  - We will answer both questions using model comparisons!

- 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 numerator DF = ΔDF**
4. Get *p*-value (from CHIDIST in excel, LRTEST in STATA, ANOVA in R, or the %FitTest custom macro program I wrote in SAS)

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(#size)*(#parameters)
  - What "#parameters" means depends on estimator in SAS: ML = ALL parameters; REML = variance model parameters only
    - In R nlme and STATA, #parameters = ALL parameters regardless
  - What "#size" means differs by program: #size = level-2 N in SAS and R mmrm, but #size is total number of observations in R nlme and STATA

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 will directly predict correlations instead of covariances because covariances will differ when variances differ
    - In R GLS and LME, "structures" are always specified using correlations
  - May not 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 (except AR1):
  - 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 using only the **R matrix**
  - We will refer to these as **"R-only models"** (use **REPEATED** only), which can be estimated in SAS MIXED, STATA MIXED, or R GLS or MMRM
  - Although the **R** matrix *can* be allowed to differ across individuals, 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 of the **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** = *n* x *n* matrix of **total** (marginal) variance and covariance over time
    - If the model also 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, it's **R**-only*

R-Only ACS Models

- The **R-only** models to be presented next are all specified using:
  - STATA **RESIDUALS** option (with "noconstant" option in random part of model)
  - R **MMRM** (more choices than GLS; because LME and LMER require random effects)
  - SAS **REPEATED** statement only (no RANDOM statement)
- They are explained by showing their predicted **R** matrix, which tries to recreate the **total** (marginal) variances and covariances across occasions
  - Total variance per occasion on diagonal (leftover variance if predictors are included)
  - Total covariances across occasions on off-diagonals (or leftover covariances)
  - I've included in " " the labels SAS uses for each parameter, but it varies by program
- Correlations across occasions can be calculated given variances and covariances, which would be shown in the **RCORR** matrix (or correlation given directly in GLS)
  - 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
  - Btw—UN models take forever to fit in STATA MIXED! They also seem to result in incorrect denominator DF (for testing fixed effects) using GLS in R (but not MMRM in R)

R-Only ACS Models: CS/CSH

- **Compound Symmetry: TYPE=CS** (you know this one already)
  - 2 parameters:
    - **1 "residual" variance σ_e^2**
    - **1 "CS" covariance** across occasions
  - Constant total variance: $CS + \sigma_e^2$ (what is given by STATA instead)
  - 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

- **1st 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….

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

- **1st 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*

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

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

- **<u>Toeplitz Heterogeneous(*n*): TYPE=TOEPH(*n*)</u>**
  - *n* + (*n*−1) parameters:
    - ***n* separate "Var(*n*)" total variances σ^2_{Tn}**
    - ***n*−1 "TOEPH(?)" 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 −2ΔLL tests (instead of mixture −2ΔLL tests, stay tuned)
  - Helpful to start by assuming heterogeneous variances until you settle on the covariance pattern, then test if het. variances are still needed
  - When in doubt, just compare AIC and BIC (useful even with −2ΔLL tests)

The Other Family of ACS Models

- **R**-only models try to *directly* recreate the pattern of **total** (marginal) variance and covariance over time (without separation into "levels")
- **G** and **R** models *indirectly* recreate 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 (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 (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** (marginal) variance and covariance across time that results from putting **G** and **R** together: $\mathbf{V} = \mathbf{ZGZ}^{\mathrm{T}} + \mathbf{R}$
    - **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 Recreated 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

Independent (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 recreated data matrix is called V matrix, built 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 of 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 recreated 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 **R** 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 highest-lag **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 level-2 random intercept $\mathbf{U_{0i}}$)
  - Resulting **R**-only model (of BP+WP combined) may require lots of estimated parameters as a result (e.g., 8 occasions? Pry need a 7-lag Toeplitz 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 $\boldsymbol{\tau}^2_{\mathbf{U_0}}$ in **G** → control for person mean differences
  - THEN recreate just the **residual** variance and 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 $\boldsymbol{\tau}^2_{\mathbf{U_0}}$ 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)
  - Can be estimated in SAS MIXED, STATA MIXED, or LME in R (only some)

Random Intercept + Diagonal **R** Models

RI and DIAG: **V** is built 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 built from **G [TYPE=UN]** and **R [TYPE=UN(1)]**:
heterogeneous *residual variances;* ***no*** *residual correlations*

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

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

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

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

<u>RI and ARH1</u>: V is built from G [TYPE=UN] and R [TYPE=ARH(1)]:
heterogeneous *residual variances;* ***auto-regressive lagged*** *residual correlations*

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

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

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

RI and TOEP*n−1*: **V** is built 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

RI and TOEPH*n−1*: **V** is built from **G [TYPE=UN]** and **R [TYPE=TOEPH(*n−1*)]:**
heterogeneous *residual variances;* ***banded*** *residual correlations*

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 + TOEP2 **R** Models

RI and TOEP2: V is built from **G [TYPE=UN] and R [TYPE=TOEP(2)]:**
homogeneous *residual variances;* ***banded*** *residual covariance at* ***lag1*** *only*

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

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

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

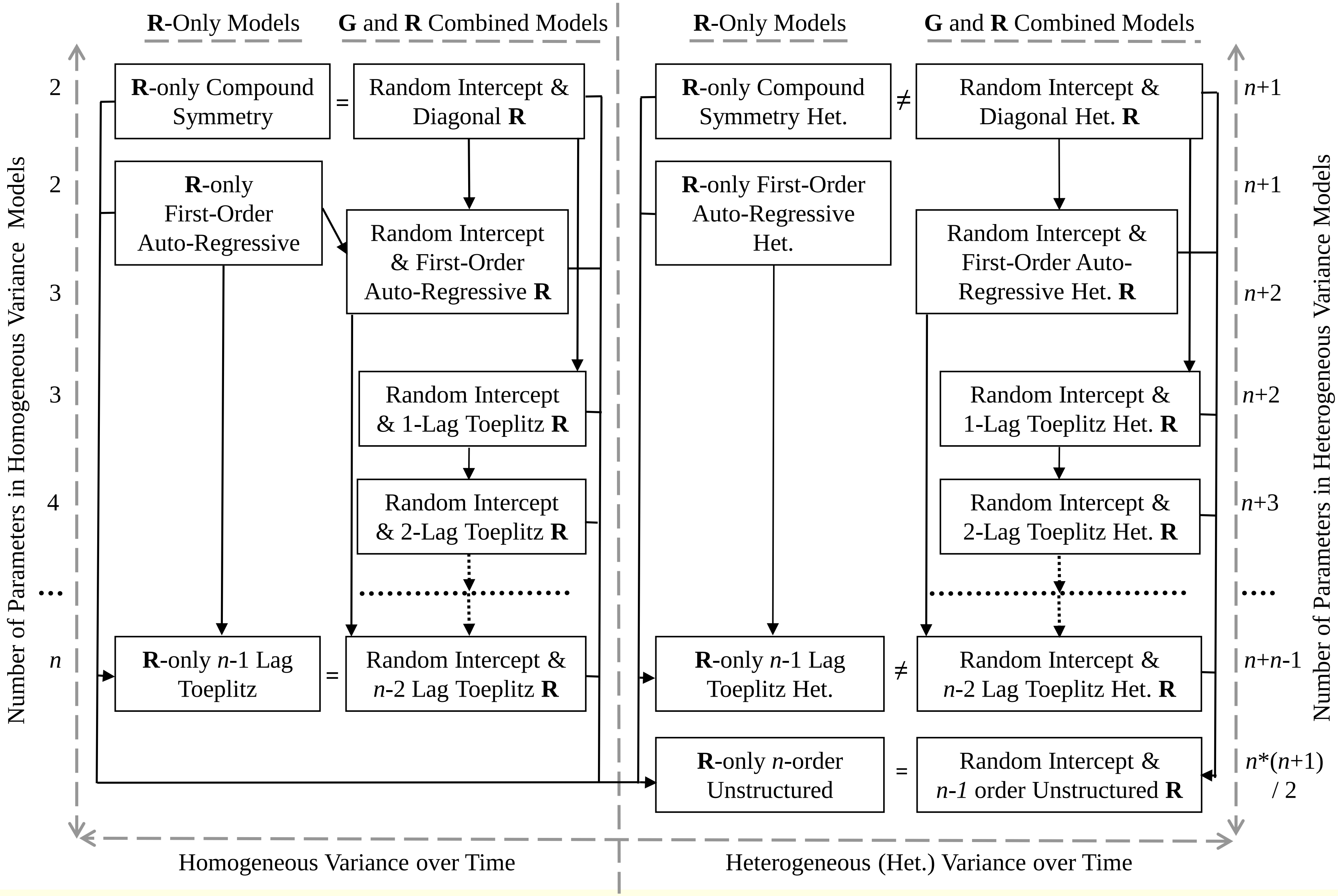
RI and TOEPH2: V is built from **G [TYPE=UN] and R [TYPE=TOEPH(2)]:**
heterogeneous *residual variances;* ***banded*** *residual correlation at* ***lag1*** *only*

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

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

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

Hoffman (2015) Figure 4.1: 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 variance in **G**…
  - 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 **quantitative** in the fixed effects, you will need another variable for **time** that is **categorical** to use in SAS syntax:
  - "Quantitative Time" → predictor on MODEL (and RANDOM) statements
  - "Categorical Time" → ID variable 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 or covariances of same lag equal doesn't make sense
  - Other structures can be used for unbalanced time
    - In SAS, SP(POW)(time) = AR1; In R GLS or LME, corCAR1 = AR1

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)
  - Marginal **R**: *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** combine to make **V**):
  - Specify random intercept variance $\tau^2_{U_0}$ in **G** using RANDOM (or ||), then specify **R** model using SAS REPEATED (or STATA residual or GLS CORR=)
  - **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** predicts leftover variance and covariance after accounting for BP mean differences (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 **SAS RANDOM** (or STATA ||) statement → **G matrix**
  - Random intercept models BP interindividual differences in mean outcome
- What about **residual variance σ^2_e (for e_{ti})** ?
  - Use the **SAS REPEATED** (or STATA residuals) statement → **R matrix**
    - **WITHOUT a RANDOM statement: R is BP and WP variance together** = σ^2_T
      → Total (marginal) 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
- In all programs, the **REPEATED** statement is always there implicitly…
  - Any model **always** has at least one residual variance in **R** matrix
- In SAS, the **RANDOM** statement is only there if you write it
  - **G** matrix isn't always necessary (don't always need random intercept)
  - In STATA MIXED, random intercepts are included by default
  - In R, the default for random effects varies across packages…

ACS Models in R

- I have found through trial and error that no one R function can do everything that SAS and STATA MIXED can do, so here are (some of) the R choices and their pros and cons:
  - `mmrt` package has `mmrt` function for R-only models
    - More and easier choices than `gls`, Satterthwaite DDF work well
    - No built-in functions to use for LRTs (so I did them in code) or RCORR
  - `nlme` package has `gls` and `lme`
    - Neither works consistently well with Satterthwaite DDF
    - `gls` has a few built-in R-only models; `lme` has a few built-in G+R models
    - Can use `anova` for LRT model comparisons
  - lme4 package has lmer
    - `lmerTest` package adds Satterthwaite DDF
    - NO choices for R matrix (= diagonal always)
    - Can use `ranova` function for LRT model comparisons

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 (CS): Equal variance and covariance over time
  - Unstructured (UN): All variances & covariances estimated separately
  - Fitting CS or UN via ML or REML estimation allows missing occasions
- MLM gives us choices in the middle
  - Goal: Get as close to UN as parsimoniously as possible
  - **R**-only: Structure TOTAL (marginal) variation in one matrix (**R** only)
  - **G**+**R**: Put constant covariance due to random intercept in **G**, then structure residual covariance in **R** (so that **G** and **R** build **V** TOTAL)

Bonus: Continuous Time AR1

- In SAS, use REPEATED TYPE=SP(POW)(*time*): $\boxed{\sigma^2 \rho^{d_{ij}}}$
  - SP(POW) = "spatial power" structure (up to 4 dimensions!), using a continuous time variable as the 2 dimensions here
  - 2 parameters build the covariance between any occasions i and j:
    - $\boldsymbol{\sigma^2}$ = constant variance over time (total if R-only; of residuals if G and R)
    - $\boldsymbol{\rho}$ = AR1 correlation (total if R-only; of residuals if G and R)
  - $\boldsymbol{d}$ = distance in time between occasions i and j creates the exponent
    - I *think* this means time can be unbalanced and unequal-interval!
- Example continuous time AR1 R matrix using integer time

$$\sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \rho^{d_{34}} \\ \rho^{d_{41}} & \rho^{d_{42}} & \rho^{d_{43}} & 1 \end{bmatrix}$$

But each person could have their own version of the R matrix for their particular occasions, still built from 2 common ingredients!

Bonus: Continuous Time AR1

- In R GLS or LME, use TYPE=corCAR1
  - From p. 48 of nlme manual

Examples

```
## covariate is Time and grouping factor is Mare
cs1 <- corCAR1(0.2, form = ~ Time | Mare)

# Pinheiro and Bates, pp. 240, 243
fm1Ovar.lme <- lme(follicles ~
           sin(2*pi*Time) + cos(2*pi*Time),
   data = Ovary, random = pdDiag(~sin(2*pi*Time)))
```

Usage

```
corCAR1(value, form, fixed)
```

Arguments

`value`	the correlation between two observations one unit of time apart. Must be between 0 and 1. Defaults to 0.2.
`form`	a one sided formula of the form `~ t`, or `~ t | g`, specifying a time covariate `t` and, optionally, a grouping factor `g`. Covariates for this correlation structure need not be integer valued. When a grouping factor is present in `form`, the correlation structure is assumed to apply only to observations within the same grouping level; observations with different grouping levels are assumed to be uncorrelated. Defaults to `~ 1`, which corresponds to using the order of the observations in the data as a covariate, and no groups.
`fixed`	an optional logical value indicating whether the coefficients should be allowed to vary in the optimization, or kept fixed at their initial value. Defaults to `FALSE`, in which case the coefficients are allowed to vary.

Value

an object of class `corCAR1`, representing an autocorrelation structure of order 1, with a continuous time covariate.