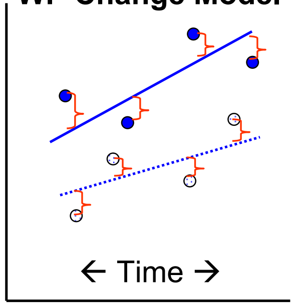
Lecture 6: Time-Varying Predictors in Longitudinal Models

- **Effects of Time-Varying Predictors**
- Person-Mean-Centering (PMC)
  - PMC Examples in SAS, SPSS, and STATA
- Grand-Mean-Centering (GMC)
  - GMC Examples in SAS, SPSS, and STATA
- Model Extensions and Model Evaluation

The JOY of Time-Varying Predictors

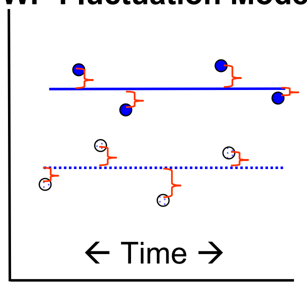
- TV predictors predict leftover **WP (residual) variation**:

WP Change Model



If model for time works, then residuals should look like this →

WP Fluctuation Model



- Modeling time-varying predictors is complicated because they represent an **aggregated effect**:
  - Effect of *between-person* variation in x_{ti} on Y
  - Effect of *within-person* variation in x_{ti} on Y
  - Here we are assuming x_{ti} fluctuates over time…
    - *Will need multivariate model if x_{ti} changes systematically over time*

The JOY of Time-Varying Predictors

- Time-varying predictors usually carry 2 kinds of effects because they are really 2 predictor variables, not 1
- Example: Stress measured daily
  - Some days are worse than others:
    - **WP variation in stress** *(represented as deviation from own mean)*
  - Some people just have more stress than others all the time:
    - **BP variation in stress** *(represented as mean predictor over time)*
- Can quantify each source of variation with an ICC
  - ICC = (BP var) / (BP + WP var)
  - ICC > 0? There is BP variation in the time-varying predictor
  - ICC < 1? There is WP variation in the time-varying predictor

Between- vs. Within-Person Effects

- Between- & within-person effects in <u>SAME</u> direction
  - Stress → Health?
    - (BP) People with more stressful lives may have <u>worse</u> general health than people with less stressful lives
    - (WP) People may feel <u>worse</u> than usual when they are currently under more stress than usual (regardless of what "usual" is)
- Between- & within-person effects in <u>OPPOSITE</u> direction
  - Exercise → Blood pressure?
    - (BP) People who exercise more often generally have <u>lower</u> blood pressure than people who are more sedentary.
    - (WP) During exercise, blood pressure is <u>higher</u> than during rest.
- Variables have different **meanings** at different levels!
- Variables have different **scales** at different levels!

3 Pieces of Information about Effects of Time-Varying Predictors

- **Is the Between-Person (BP) effect of x_{ti} significant?**
  - Are people higher on X *on average* higher on Y *on average*?
  - Does person mean X account for L2 random intercept U_{0i} variance?
- **Is the Within-Person (WP) effect of x_{ti} significant?**
  - If you have higher x_{ti} values than usual *right now*, do you also have higher y_{ti} values than predicted *right now*?
  - Does within-person deviation x_{ti} account for L1 residual e_{ti} variance?
- **Are the BP and WP Effects of x_{ti} of the same magnitude?**
  - Is there *an additional bonus/decrement for predicted y_{ti}* that comes from having high average scores on PMx_i *above and beyond* (controlling for) just having a high x_{ti} value right now?
  - In other words, is there a *contextual effect*?
  - Do the BP and WP effects of x_{ti} show *convergence*, such that we only need one parameter for the effect of x_{ti} instead of two?

Modeling Time-Varying Predictors

- <u>**Level-2 effect of x_{ti}:**</u>
  - The level-2 effect of x_{ti} is usually represented by the person's mean of time-varying x_{ti} across time
  - $PersonMeanX_i$ should be centered (at grand mean, other constant) so that 0 is meaningful, just like any other time-invariant predictor
- <u>**Level-1, Within-Person effect of X_{ti}:**</u>
  - 2 alternative ways of representing level-1 effect of x_{ti}:
    - "**Group-mean-centering**" → "**person-mean-centering**" in longitudinal
    - "**Grand-mean-centering**" → can center around any meaningful constant, not necessarily the grand mean (it's just called that)
  - Note that these 2 choices do NOT apply to the level-2 effect of x_{ti}!
    - But the interpretation of the level-2 effect of x_{ti} WILL DIFFER based on which level-1 centering method for x_{ti} you choose!

Lecture 6: Time-Varying Predictors in Longitudinal Models

- Effects of Time-Varying Predictors
- **Person-Mean-Centering (PMC)**
  - PMC Examples in SAS, SPSS, and STATA
- Grand-Mean-Centering (GMC)
  - GMC Examples in SAS, SPSS, and STATA
- Model Extensions and Model Evaluation

Person-Mean-Centering (PMC)

- In PMC, we decompose the time-varying predictor x_{ti} into the 2 variables it really is, and include those instead:
- **Level-2, PM predictor = person mean of x_{ti}**
  → $PMx_i = \overline{x}_i - C$
  - PMx_i is centered at C, a constant so 0 is meaningful (e.g., sample mean)
  - PMx_i is positive? Above sample mean → "more than other people"
  - PMx_i is negative? Below sample mean → "less than other people"
- **Level-1, WP predictor = deviation from person mean of x_{ti}**
  → $WPx_{ti} = x_{ti} - \overline{x}_i$
  - WPx_{ti} is NOT centered at a constant; is centered at a VARIABLE
  - WPx_{ti} is positive? Above your own mean → "more than usual"
  - WPx_{ti} is negative? Below your own mean → "less than usual"

Within-Person Fluctuation Model with **Person-Mean-Centered Level-1 X:** WP and BP Effects are **Separate** Parameters

L1 is **person**-mean-centered into WPx_{ti}, with PMx_i at L2:

$$\text{L1:} \quad y_{ti} = \beta_{0i} + \beta_{1i}(\mathbf{WPx_{ti}}) + \mathbf{e_{ti}}$$

$$\text{L2:} \quad \beta_{0i} = \gamma_{00} + \gamma_{01}(\mathbf{PMx_i}) + \mathbf{U_{0i}}$$

$$\beta_{1i} = \gamma_{10}$$

WPx_{ti} contains only L1 variation (=x_{ti} – PMx_i)

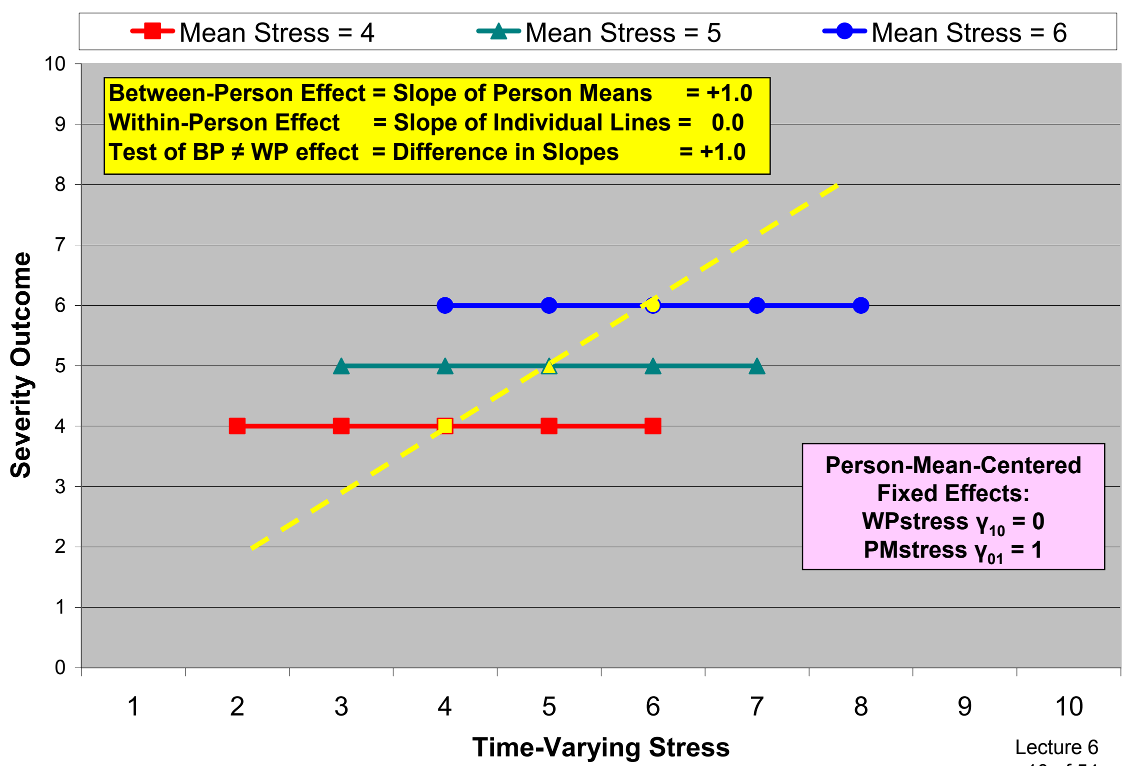
PMx_i contains only L2 variation (=mean of x_{ti})

γ_{10} = WP effect = effect of higher X than usual

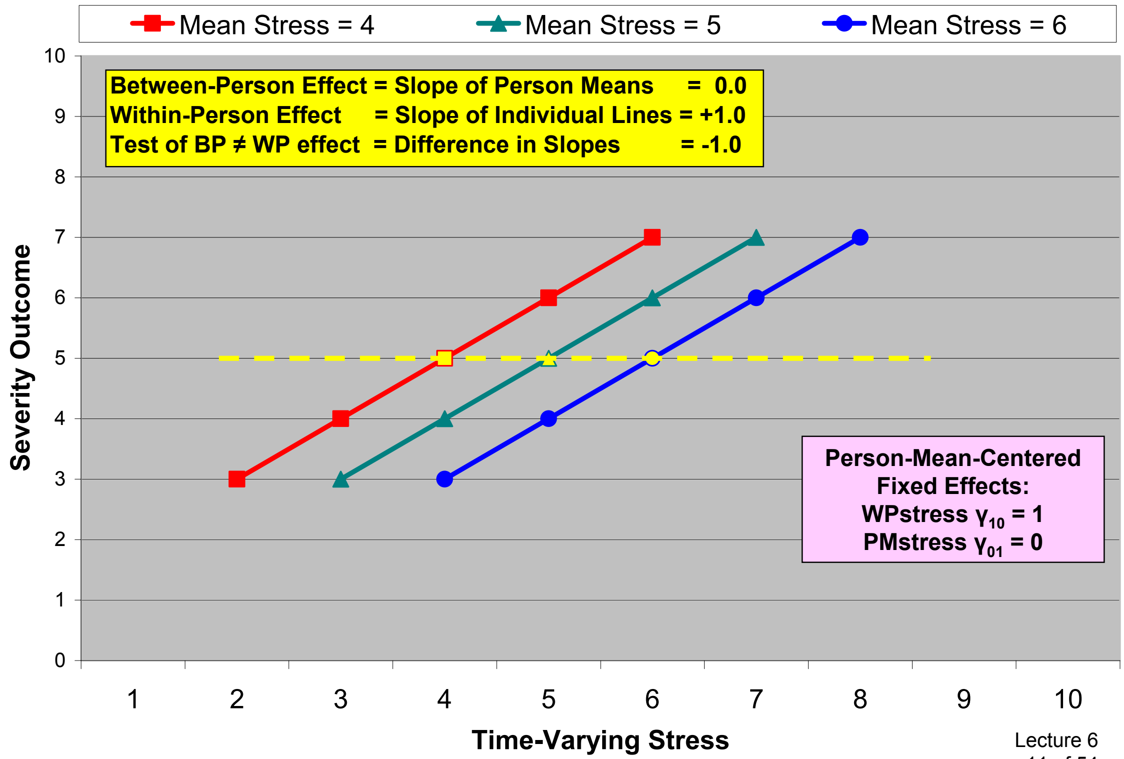
γ_{01} = BP effect = effect of being a "high X" person

Because WPx_{ti} and PMx_i are uncorrelated, each gets the total effect for its level (WP=L1, BP=L2)

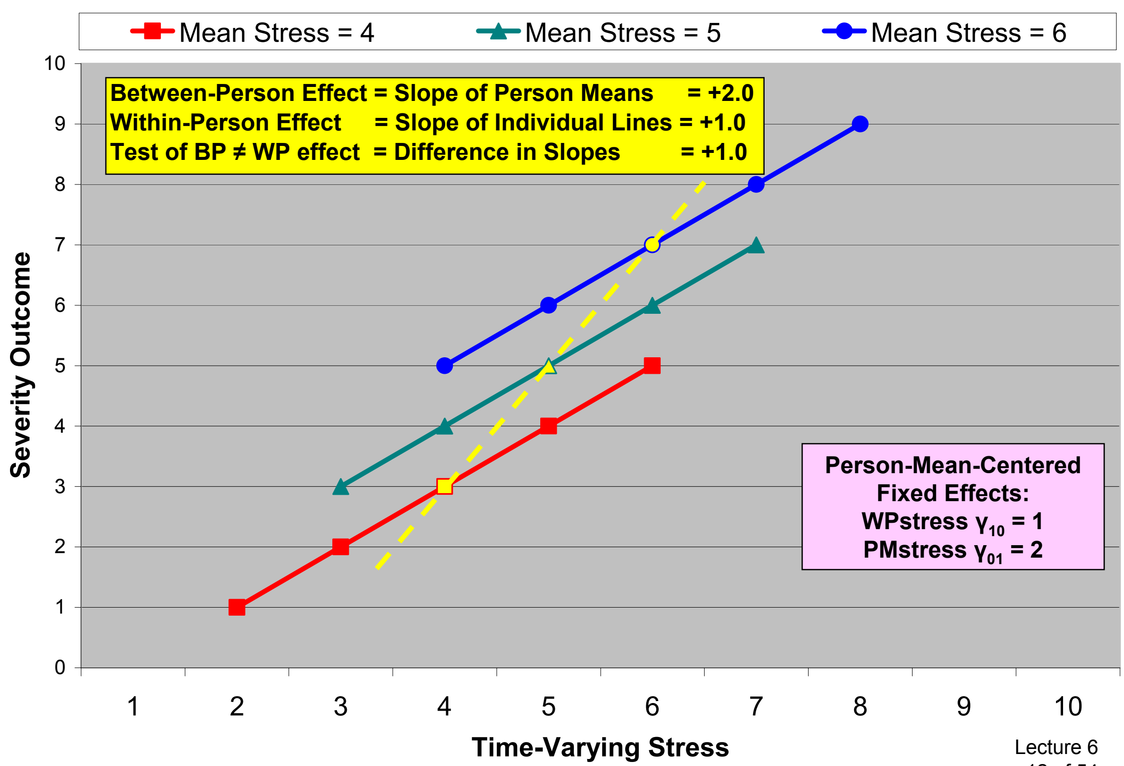
ALL Between-Person, NO Within-Person



NO Between-Person, ALL Within-Person



Between-Person > Within-Person



Within-Person Fluctuation Model with **Person-Mean-Centered Level-1 X:** WP and BP Effects are **Separate** Parameters

L1 is **person**-mean-centered into WPx_{ti}, with PMx_i at L2:

L1: $y_{ti} = \beta_{0i} + \beta_{1i}(WPx_{ti}) + e_{ti}$

L2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

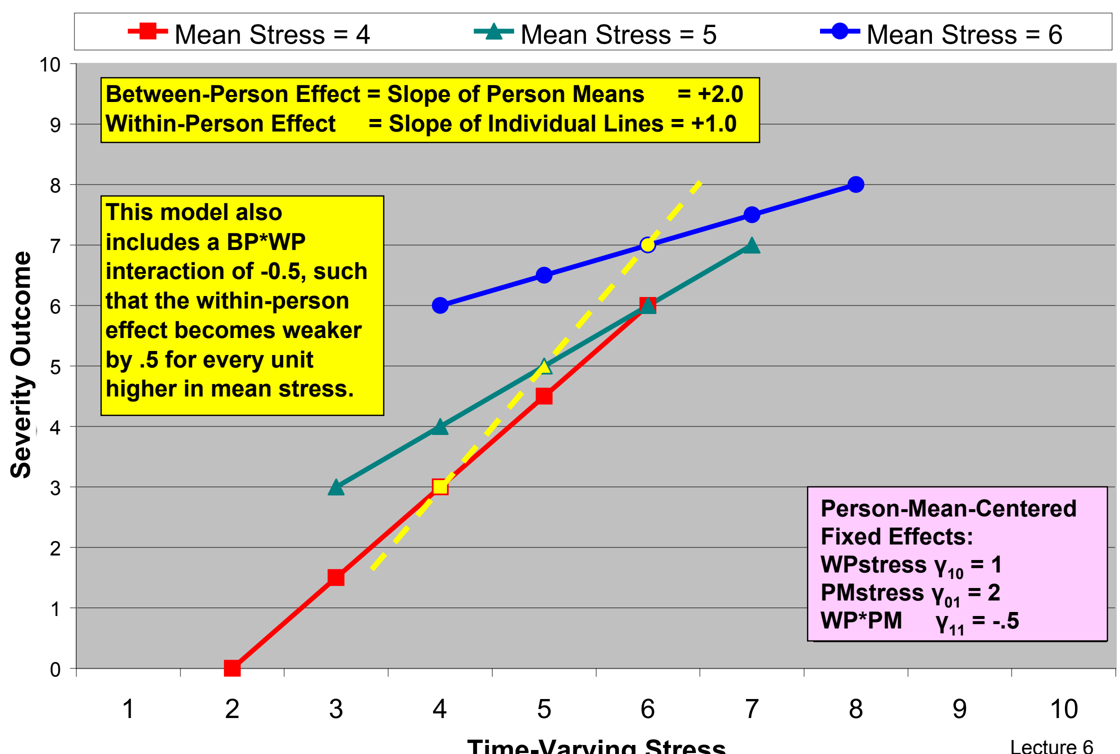
$\beta_{1i} = \gamma_{10} + \gamma_{11}(PMx_i) + U_{1i}$

Model Parameters:
4 Fixed Effects,
2 Random Effects Variances,
1 Random Effects Covariance,
1 Residual variance

Could add a cross-level interaction of WPx_{ti} and PMx_i

Could also add a random slope of WPx_{ti}

Between-Person **x** Within-Person Interaction



Lecture 6: Time-Varying Predictors in Longitudinal Models

- Effects of Time-Varying Predictors
- Person-Mean-Centering (PMC)
  - **PMC Examples in SAS, SPSS, and STATA**
- **Grand-Mean-Centering (GMC)**
  - GMC Examples in SAS, SPSS, and STATA
- Model Extensions and Model Evaluation

3 Pieces of Information about Effects of Time-Varying Predictors

What Person-Mean-Centering tells us directly:

- **Is the Between-Person (BP) effect of x_{ti} significant?**
  - Are people higher on x_{ti} *on average* (PMx_i) higher on y_{ti} *on average*?
    - Note this is NOT controlling for current levels of x_{ti}
  - Does person mean x_{ti} (PMx_i) account for L2 random intercept variance?
    - Is the effect of PMx_i significant in the model?
- **Is the Within-Person (WP) effect of x_{ti} significant?**
  - If you have higher x_{ti} values *than usual right now*, do you also have higher y_{ti} values than predicted *right now*?
    - Note this does NOT depend on absolute value of x_{ti}, just relative value of x_{ti}
  - Does within-person deviation x_{ti} (WPx_{ti}) account for L1 residual variance?
    - Is the effect of WPx_{ti} significant in the model?

3 Pieces of Information about Effects of Time-Varying Predictors

What Person-Mean-Centering does NOT tell us directly:

- **Are the BP and WP Effects of x_{ti} of the same magnitude?**
  - Is there *an additional bonus/decrement for predicted y_{ti}* that comes from having high average scores on x_{ti} (PMx_i) *above and beyond* (controlling for) just having a high x_{ti} score right now?
  - In other words, is there a ***contextual effect***?
  - If you know current x_{ti}, does it matter what PMx_i is, too?
  - Do the BP and WP effects of x_{ti} show *convergence (equality)*, such that we only need one parameter for the effect of x_{ti} instead of two?

- **To answer these questions, we have two options:**
  - Ask for explicit model-implied contrasts using ESTIMATE statements
  - Use **grand-mean-centering** for time-varying X instead
    → **$TVx_{ti} = x_{ti} - C$ x_{ti} is centered at a CONSTANT, not a variable**

Back to Regression for a Moment: $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + e$

- If X_1 and X_2 **ARE NOT** correlated:
  - β_1 is **ALL the relationship** between X_1 and Y
  - β_2 is **ALL the relationship** between X_2 and Y

- If X_1 and X_2 **ARE** correlated:
  - β_1 is **less than** the full relationship between X_1 and Y
    - "Unique" effect of X_1 *controlling for X_2* or *holding X_2 constant*
  - β_2 is **less than** the full relationship between X_2 and Y
    - "Unique" effect of X_2 *controlling for X_1* or *holding X_1 constant*

- Hang onto that principle…

PMC vs. GMC for Time-Varying Predictors

Level 2		Original	PMC Level 1	GMC Level 1
PMx_i	PMx_i - 5	x_{ti}	$WPx_{ti} = x_{ti} - PMx_i$	$TVx = x_{ti} - 5$
3	-2	2	-1	-3
3	-2	4	1	-1
7	2	6	-1	1
7	2	8	1	3

Same PMx_i goes into either model…

Under **PMC**, **WPx_{ti}** DOES NOT contain level 2, BP variation, so PMx_i is NOT correlated with WPx_{ti}

Under **GMC**, **TVx_{ti}** DOES contain level 2 BP variation, so PMx_i IS correlated with TVx_{ti}

***** This means that the effects of PMx_i and TVx_{ti} together under GMC will be different than their effects by themselves…**

Within-Person Fluctuation Model with **x_{ti} represented at Level 1 Only:**

WP and BP Effects are **Smushed Together**

L1 **grand**-mean centered into TVx_{ti}, **WITHOUT** PMx_i at L2:

L1: $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

TVx_{ti} contains BOTH L1 and L2 variation

L2: $\beta_{0i} = \gamma_{00} + U_{0i}$

$\beta_{1i} = \gamma_{10}$

γ_{10} = smushed* WP and BP effect

***aka, *convergence, conflated, composite*, or *aggregate* effect**

Because TVx_{ti} actually contains 2 different variables (WPx and PMx), its 1 parameter has to do the work of 2 predictors

Convergence (Smushed) Effect of a Time-Varying Predictor

$$\text{Convergence Effect: } \gamma_{conv} \approx \frac{\dfrac{\gamma_{BP}}{SE_{BP}^2} + \dfrac{\gamma_{WP}}{SE_{WP}^2}}{\dfrac{1}{SE_{BP}^2} + \dfrac{1}{SE_{WP}^2}}$$

Adapted from Raudenbush & Bryk (2002, p. 138)

- Convergence effect will often be closer to the **within-person effect** (due to larger level-1 sample size and thus smaller SE)
- **It is the rule, rather than the exception**, that between and within effects differ (Snijders & Bosker, p. 52-56, and personal experience!)
- However – you don't have to assume convergence in order to use grand-mean-centering for a time-varying predictor…

Within-Person Fluctuation Model with **Grand-Mean-Centered Level-1 X:**
Tests Difference of WP and BP effects (It's been fixed!)

L1 x is **grand**-mean centered into TVx_{ti}, **WITH** PMx_i at L2:

L1: $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

L2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10}$

TVx_{ti} contains BOTH L1 and L2 variation

PMx_i contains only L2 variation (=mean of L1)

Because L2 variance is still in TVx_{ti}, PMx_i takes the <u>unique</u> part of the L2 effect that TVx_{ti}'s L2 variance didn't cover

γ_{10} *becomes* WP effect → pure L1 effect (now that PMx is included)

γ_{01} *is contextual (incremental)* L2 effect → tests difference of BP and WP effects

Equivalence of PMC and GMC (Fixed effects; Main effects only)

Person-Mean-Centering (uses WPx_{ti}):

L1: $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti} - PMx_i) + e_{ti}$

L2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10}$

→ $y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(TVx_{ti} - PMx_i) + U_{0i} + e_{ti}$ ← **In terms of WPx_{ti}**

→ $y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PMx_i) + \gamma_{10}(TVx_{ti}) + U_{0i} + e_{ti}$ ← **In terms of TVx_{ti}**

Grand-Mean-Centering (uses TVx_{ti}):

L1: $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

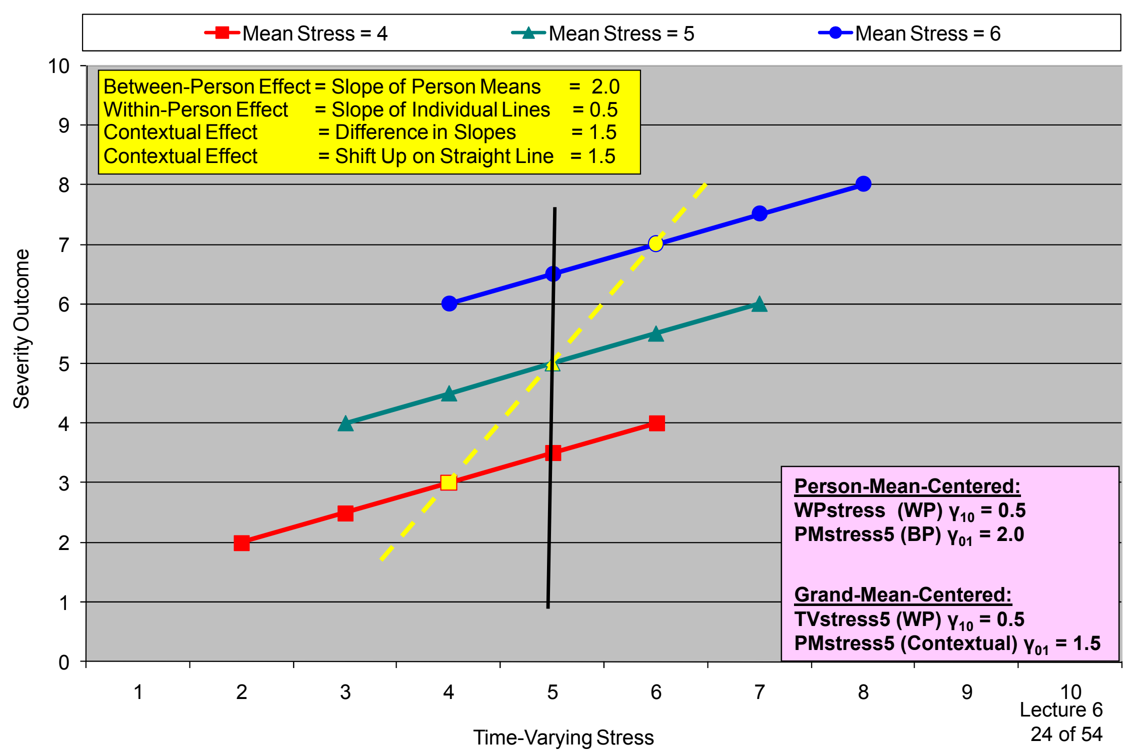
L2: $\beta_{0i} = \gamma_{00} + \gamma_{01}^{*}(PMx_i) + U_{0i}$

$\beta_{1i} = \gamma_{10}$

→ $y_{ti} = \gamma_{00} + \gamma_{01}^{*}(PMx_i) + \gamma_{10}(TVx_{ti}) + U_{0i} + e_{ti}$

Term	P-MC	G-MC
Intercept	γ_{00}	γ_{00}
WP Effect	γ_{10}	γ_{10}
Contextual	$\gamma_{01} - \gamma_{10}$	γ_{01}^{*}
BP Effect	γ_{01}	$\gamma_{01}^{*} + \gamma_{10}$

PMC vs. GMC: Interpretation Example



3 Pieces of Information about Effects of Time-Varying Predictors

- **Is the Between-Person (BP) effect of x_{ti} significant?**
  - Are people higher on x_{ti} *on average* (PMx_i) higher on y_{ti} *on average*?
  - Only **PERSON**-mean-centering gives you this directly in the model
- **Is the Within-Person (WP) effect of x_{ti} significant?**
  - If you have higher x_{ti} values *than usual right now*, do you also have higher y_{ti} values than predicted *right now*?
  - **Either PERSON- or GRAND-**mean-centering gives you this directly
- **Are the BP and WP Effects of x_{ti} of the same magnitude?**
  - Is there *an additional bonus/decrement for predicted y_{ti}* (contextual effect) that comes from having high average scores on x_{ti} (PMx_i) *above and beyond* (controlling for) just having a high x_{ti} value right now?
  - Only **GRAND**-mean-centering gives you this directly in the model

** Can use **ESTIMATE** (in SAS) or **TEST** (in SPSS) or **LINCOM** (in STATA) or **NEW** (in Mplus) to get any implied effects not directly provided by the model

Parameter Interpretation across Methods of Centering Level-1 X_{ti}

- **Person-mean-centering → subtract a VARIABLE**
  - L1 predictor "WPx_{ti}" = time-varying x_{ti} – original person mean x_i
    - *Directly represents Within-Person (WP) effect of X (regardless of whether PMx is in the model at L2)*
  - L2 predictor "PMx_i" = person mean x_i – constant
    - *Directly represents Between-Person (BP) effect of x_{ti} (regardless of whether WPx_{ti} is in the model, too)*
- **Grand-mean-centering → subtract a CONSTANT**
  - L1 predictor "TVx_{ti}" = time-varying x_{ti} – constant
    - *WITHOUT PMx_i at L2, is combined BP and WP effects*
    - *WITH PMx_i at L2, becomes WP effect*
  - L2 predictor "PMx_i" = person-mean x_i – constant
    - *WITHOUT TVx_{ti} at L1, is BP effect (like above)*
    - *WITH TVx_{ti} at L1, becomes difference of BP and WP effects*

Modeling Time-Varying Categorical Predictors

- Person- and grand-mean-centering really only apply to *continuous* level-1 predictors, but concerns about separation of WP effects from BP effects applies to *categorical* level-1 predictors as well.
- I wouldn't recommend trying to create PMx of a categorical variable. Try to think about types of people, and code BP effects accordingly.
- Example: Dementia present/not at each time point?
  - **BP effects** → Ever diagnosed with dementia (no, yes)?
    - People who will eventually be diagnosed may be on a different trajectory prior to the point of diagnosis as well as after (a BP effect)
  - **TV effect** → Diagnosed with dementia at each time point (no, yes)?
    - Scores may be lower after diagnosis than before (a WP effect)
- Other examples: Mentor Status, Father Absence, Type of Shift Work

Lecture 6: Time-Varying Predictors in Longitudinal Models

- Effects of Time-Varying Predictors
- Person-Mean-Centering (PMC)
  - PMC Examples in SAS, SPSS, and STATA
- Grand-Mean-Centering (GMC)
  - **GMC Examples in SAS, SPSS, and STATA**
- **Model Extensions and Model Evaluation**

Summary: Time-Varying Predictors

- Time-varying predictors carry at least two potential effects:
  - Some people are higher/lower all the time → BP, level 2 effect
  - Some occasions are higher/lower than usual → WP, level 1 effect
- TV predictors can be used to predict <u>intraindividual deviation</u> from predicted growth (or deviation from a flat line in a WP fluctuation model)
  - HOWEVER: If their ICC is ≠ 0 or 1, they contain both BP AND WP variance, each of which could be related to Y differently
    - GMC → TVx_{ti} + PMx_i → used to test WP effect=0? and L2 contextual effect=0?
    - PMC → WPx_{ti} + PMx_i → used to test WP effect=0? and L2 BP effect=0?
- Another alternative is "Time0 centering" (not pure BP/WP separation):
  - Level 1 = stress – **stressTime0** → longitudinal effect
    - L1 represents *change from baseline*, not deviation from own mean
  - Level 2 = **stressTime0** – c → cross-sectional effect
    - L2 represents effect of *baseline level*, not effect of mean level

What about "Multilevel SEM"?

- In order to get BP and WP effects, so far we've separated the BP and WP variance in a time-varying predictor by brute force (e.g., by computing a PMx_i variable to use with TVx_{ti} or WPx_{ti})
- An alternative is "multilevel SEM" (which isn't really SEM if it doesn't have other kinds of latent variables besides the MLM-based random effects, but whatever)
- Multivariate model → the variance in TV predictors is decomposed **by the model** into random intercept (BP) vs. residual (WP), the same as if it were an outcome (thus predictors = outcomes)
  - Pros:
    - Can have missing data on TV predictors (because are outcomes then)
    - Can be used to test multilevel mediation (currently impossible in MIXED)
    - May have less biased level-2 effects because there is no observed PMx_i variable assumed perfectly reliable (see Lüdke et al. 2008 Psych Methods)
  - Cons:
    - Greater estimation demands → more likely to blow up (only available in Mplus)
    - Different (but equivalent) syntax → BP or contextual effects (be careful)
    - Good luck fitting interaction terms! (→ latent variable interactions)

Interactions Involving Level-1 Predictors Belong at Both Levels…

Example: Does the effect of time-varying stress (x_{ti}) on time-varying illness (y_{ti}) interact with sex (Z_i)?

- **Person-Mean-Centering for stress (x_{ti}):**
  - WPx_{ti} by Sex_i → Does the WP stress effect differ by sex?
  - PMx_i by Sex_i → Does the BP stress effect differ by sex?
    - Moderation of total person stress effect (not controlling for current stress)
    - If forgotten, then sex moderates the stress effect only at level 1 (pry **weird**)
- **Grand-Mean-Centering for stress (x_{ti}):**
  - TVx_{ti} by Sex_i → Does the WP stress effect differ by sex?
  - PMx_i by Sex_i → Does the contextual stress effect differ by sex?
    - Moderation of <u>incremental</u> person stress effect <u>controlling for</u> current stress (moderation of the "boost" in person illness from being chronically stressed)
    - If forgotten, then although the main effect of time-varying stress has been un-smushed, the interaction of TVx_{ti} by sex would still be smushed, which assumes that sex moderates the WP and BP stress effects equally (pry **wrong**)

Interactions Involving Level-1 Predictors Belong at Both Levels…

<u>Person-Mean-Centering (uses WPx_{ti}):</u>

$$y_{ti} = \gamma_{00} + U_{0i} + e_{ti} + \gamma_{10}(TVx_{ti} - PMx_i) + \gamma_{01}(PMx_i) + \gamma_{02}(Z_i) + \gamma_{03}(PMx_i)(Z_i) + \gamma_{11}(TVx_{ti} - PMx_i)(Z_i)$$

← **In terms of WPx_{ti}**

$$y_{ti} = \gamma_{00} + U_{0i} + e_{ti} + \gamma_{10}(TVx_{ti}) + (\gamma_{01} - \gamma_{10})(PMx_i) + (\gamma_{02})(Z_i) + \gamma_{11}(TVx_{ti})(Z_i) + (\gamma_{03} - \gamma_{11})(PMx_i)(Z_i)$$

← **In terms of TVx_{ti}**

<u>Grand-Mean-Centering (uses TVx_{ti}):</u>

$$y_{ti} = \gamma_{00} + U_{0i} + e_{ti} + \gamma_{10}(TVx_{ti}) + \gamma_{01}^{*}(PMx_i) + \gamma_{02}(Z_i) + \gamma_{03}^{*}(PMx_i)(Z_i) + \gamma_{11}(TVx_{ti})(Z_i)$$

After adding an interaction with L2 Z_i at both levels, then the Person-MC and Grand-MC models are equivalent

---

Intercept: $\gamma_{00} = \gamma_{00}$ **BP Effect:** $\gamma_{01} = \gamma_{01}^{*} + \gamma_{10}$ **Contextual:** $\gamma_{01}^{*} = \gamma_{01} - \gamma_{10}$

WP Effect: $\gamma_{10} = \gamma_{10}$ **BP*Z Effect:** $\gamma_{03} = \gamma_{03}^{*} + \gamma_{11}$ **Contextual*Z:** $\gamma_{03}^{*} = \gamma_{03} - \gamma_{11}$

Z Effect: $\gamma_{20} = \gamma_{20}$ **BP*WP or Contextual*WP is the same:** $\gamma_{11} = \gamma_{11}$

The same is true for BP*WP intra-variable interactions (e.g., WPx_{ti}*PMx_i or TVx_{ti}*PMx_i)

Person-Mean-Centering (uses WPx_{ti}):

$$y_{ti} = \gamma_{00} + U_{0i} + e_{ti} + \gamma_{10}(TVx_{ti} - PMx_i) + \gamma_{01}(PMx_i) + \gamma_{02}(PMx_i)^2 + \gamma_{11}(TVx_{ti} - PMx_i)(PMx_i)$$

← In terms of WPx_{ti}

$$y_{ti} = \gamma_{00} + U_{0i} + e_{ti} + \gamma_{10}(TVx_{ti}) + (\gamma_{01} - \gamma_{10})(PMx_i) + (\gamma_{02} - \gamma_{11})(PMx_i)^2 + \gamma_{11}(TVx_{ti})(PMx_i)$$

← In terms of TVx_{ti}

Grand-Mean-Centering (uses TVx_{ti}):

$$y_{ti} = \gamma_{00} + U_{0i} + e_{ti} + \gamma_{10}(TVx_{ti}) + \gamma_{01}^{*}(PMx_i) + \gamma_{02}^{*}(PMx_i)^2 + \gamma_{11}(TVx_{ti})(PMx_i)$$

Must also add PMx_i^2 to allow moderation at both levels, but the PMC and GMC models are then equivalent

---

Intercept: $\gamma_{00} = \gamma_{00}$ **BP Effect:** $\gamma_{01} = \gamma_{01}^{*} + \gamma_{10}$ **Contextual:** $\gamma_{01}^{*} = \gamma_{01} - \gamma_{10}$

WP Effect: $\gamma_{10} = \gamma_{10}$ **BP² Effect:** $\gamma_{02} = \gamma_{02}^{*} + \gamma_{11}$ **Contextual²:** $\gamma_{02}^{*} = \gamma_{02} - \gamma_{11}$

BP*WP or Contextual*WP is the same either way: $\gamma_{11} = \gamma_{11}$

When PMC vs. GMC Matters:
Random Slopes across Models

Person-Mean Centering:

$$y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(WPx_{ti}) + U_{0i} + U_{1i}(WPx_{ti}) + e_{ti} \rightarrow$$

$$y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(TVx_{ti} - PMx_i) + U_{0i} + \boxed{U_{1i}(TVx_{ti}) - U_{1i}(PMx_i)} + e_{ti}$$

Grand-Mean Centering:

$$y_{ti} = \gamma_{00} + \gamma^{*}_{01}(PMx_i) + \gamma_{10}(TVx_{ti}) + U_{0i} + \boxed{U_{1i}(TVx_{ti})} + e_{ti}$$

Both centerings yield equivalent models if the L1 effect is fixed, but NOT if the L1 effect is random.

The variance in PMx is NOT subtracted out of the random slope in Grand MC. Therefore, models with random slopes are not equivalent.

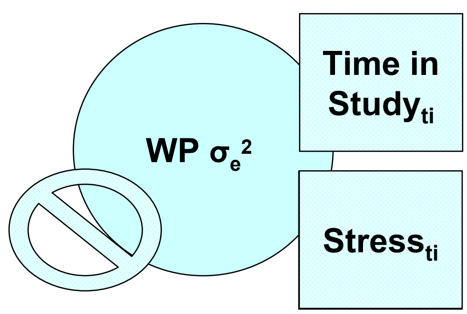
So which do you choose?

Adding Random Slopes of L1 x_{ti}

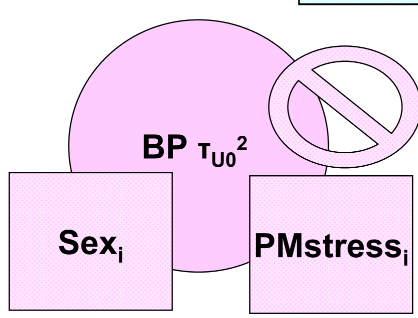
- **Random intercepts** mean different things under each model:
  - **Person-MC** → Individual differences at **WPx_{ti} =0** (everyone has)
  - **Grand-MC** → Individual differences at **TVx_{ti}=0** (not everyone has)
- **Differential shrinkage of the random intercepts** results from differential reliability of the intercept data across models:
  - Person-MC → Won't affect shrinkage of slopes unless highly correlated
  - Grand-MC → Will affect shrinkage of slopes due to forced extrapolation
- As a result, the **random slope variance may be smaller** under grand-MC than under person-MC
  - Problem worsens with greater BP variation in X (more extrapolation)
  - Anecdotal example using clustered data was presented in Raudenbush & Bryk (2002; chapter 6)
- Now let's review what happens to variance components….

Example Longitudinal Two-Level Model
Predicting Severity: Time, Stress, Sex

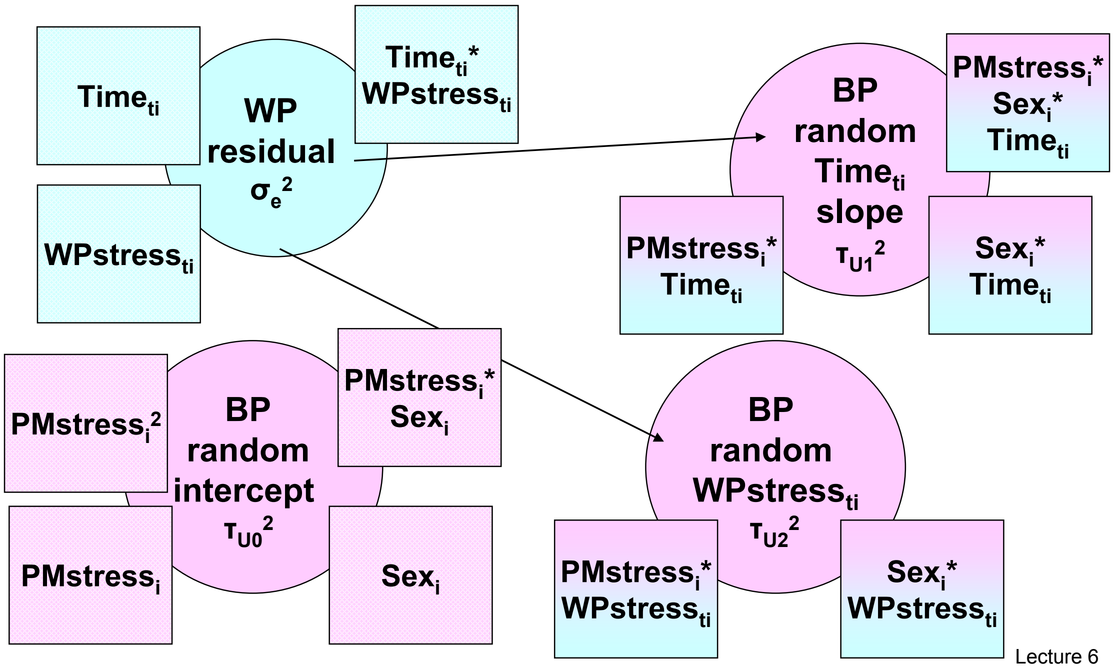
Level-1
Residual Variance:
Within-Person (WP) or Across-Time Variation from Own Mean Severity



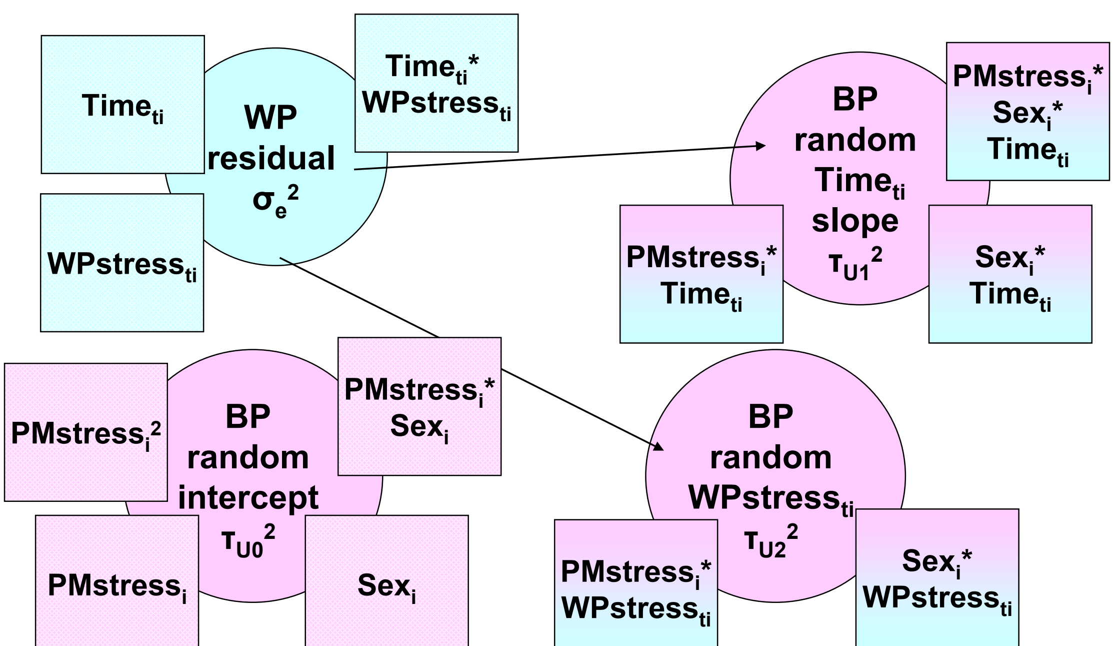
Level-2 Random
Intercept Variance:
Between-Person (BP) Mean Variation from Grand Mean Severity



Adding Cross-Level Interactions to a Random **Slopes** Model



Adding Cross-Level Interactions to a Random **Intercept** Model



Variance Accounted For By Level-2 Time-Invariant Predictors

- **Fixed effects of level 2 predictors *by themselves*:**
  - L2 (BP) main effects (e.g., Sex) reduce L2 random intercept variance $\boldsymbol{\tau_{U0}^2}$
  - L2 (BP) interactions (e.g., Sex*PMstress; PMstress*PMstress) also reduce L2 (BP) random intercept variance $\boldsymbol{\tau_{U0}^2}$
- **Fixed effects of *cross-level interactions* (level 1* level 2):**
  - If the interacting level 1 predictor is <u>random</u>, any cross-level interaction with it will reduce its corresponding L2 BP random slope variance
    - e.g., if *time* is random, then Sex**time,* PMstress**time*, and Sex*PMstress**time* can each reduce the random linear time slope variance $\boldsymbol{\tau_{U1}^2}$
  - If the interacting level 1 predictor <u>not random</u>, any cross-level interaction with it will reduce the L1 WP residual variance $\boldsymbol{\sigma_e^2}$ instead
    - e.g., if *WPstress* is not random, then Sex**WPstress,* PMstress**WPstress*, and Sex*PMstress**WPstress* will reduce the L1 (WP) residual variance $\boldsymbol{\sigma_e^2}$ → Different *WPstress* slopes from Sex and PMstress will allow better *WPstress* trajectories, reduce the variance around trajectories

Variance Accounted For (R^2) By Level-1 Time-Varying Predictors

- **Fixed effects of level 1 predictors *by themselves*:**
  - L1 (WP) main effects (e.g., linear time, quadratic time, WPstress) reduce L1 (WP) residual variance $\boldsymbol{\sigma_e^2}$
  - L1 (WP) interactions (e.g., time*WPstress) also reduce L1 (WP) residual variance $\boldsymbol{\sigma_e^2}$
- **What happens at level 2 depends on what kind of variance the level-1 predictor has:**
  - If the level-1 predictor ALSO has L2 variance (e.g., GMC predictors, age in accelerated longitudinal designs), then that L2 variance will also likely reduce the L2 random intercept variance $\boldsymbol{\tau_{U0}^2}$
  - If the level-1 predictor DOES NOT have L2 variance (e.g., PMC predictors, time in balanced designs), then its reduction in the L1 residual variance $\boldsymbol{\sigma_e^2}$ will cause an INCREASE in L2 random intercept variance $\boldsymbol{\tau_{U0}^2}$
    - Same thing happens in other case, but you don't generally see it

More on Pseudo-R^2 Effect Size

- Pseudo-R^2 is not calculated when adding **random** effects
  - Does not apply: fixed effects reduce variance, but random effects only **re-partition variance** (random effect = new pile of variance)
  - Calculate random effects confidence intervals instead!

- Pseudo-R^2 is only calculable across models with same piles of variance (meaning of each variance changes if others are added)
  - Another problem: Adding level-2 predictors of one random effect may cause other random effect variances to decrease through their correlation

- **A simple alternative: Total R^2**
  - Generate model-predicted y values from *fixed effects only* (using OUTPM in SAS, FIXPRED in SPSS, or PREDICT XB in STATA) and correlate with observed y values (then square that correlation → total R^2)
  - Total R^2 = total reduction in overall variance of y across levels
  - Can be “unfair” in models with large unexplained sources of variance
    - Such as in cross-classified models… stay tuned

Checking for Violations of Model Assumptions: Why should we care?

- “Fitting a model with untenable assumptions is as senseless as fitting a model to data that are knowingly flawed” (Singer & Willett, pg. 127)

- HOWEVER:
  - We don’t actually know the true population relationships, so we don’t know when our estimates, SE’s, and *p*-values are off
  - Recommended strategy: “check assumptions of several initial models and any model you cite or interpret explicitly”
  - Mostly informal inspection – requires judgment call
    - “We prefer visual inspection of residual distributions” (S & W pg. 128)
  - Some things are fixable, some things are not
  - **End goal: Analyze the data the least wrong way possible** (because all models are wrong; some are useful)

General Consequences of Violating Model Assumptions

2 parts of the model to be concerned with:

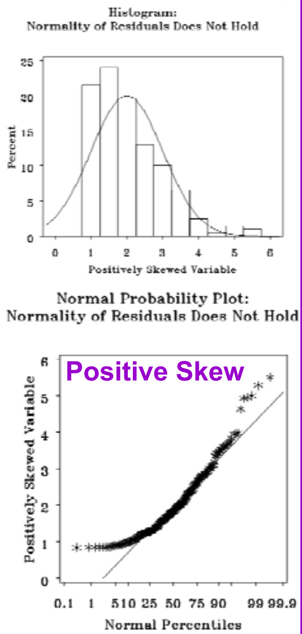
- **Model for means = fixed effects**
  - Estimates depend on having the "right" model for the means → all relevant predictors, measured with as little error as possible
  - To the extent that predictors are missing or their effects are specified incorrectly, **fixed effect estimates will be biased**
- **Model for the variances = random effects and residuals**
  - SE and *p*-values of fixed effects depend on having the "right" model for the variances → most closely approximate actual data
  - To the extent that the model for the variances is off, **fixed effects SE and *p*-values will be off, too (biased)**
  - Because the general linear mixed model is estimated using a multivariate normal distribution for the **V** matrix, certain assumptions follow…

Basic Model Assumptions

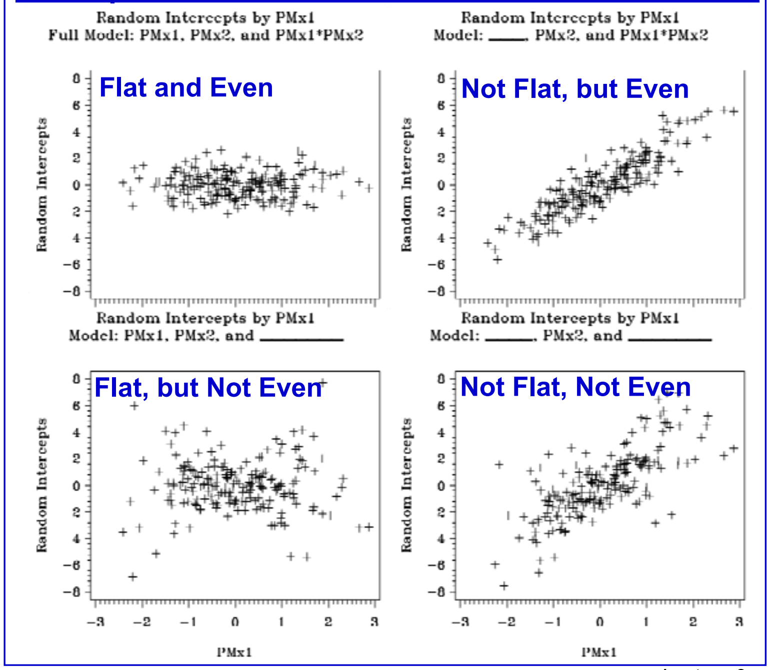
- GLM Assumptions:
  - Normality of **residuals** (not outcomes)
  - Independence and constant variance of **residuals**
    - Across sampling units
    - Across predictors
- MLM Assumptions are the same, except:
  - Apply **at** each level and **across** levels
  - More general options are available for changing the model to accommodate violations of assumptions if needed (**transform the model**, not the data)
  - ML also assumes MAR for any missing outcomes

Plots to Assess Assumptions:

Normality



Independence & Constant Variance



MLM Assumptions: **Normality**
Multiple 'residuals' to consider:

Level-1 e_{ti} residuals → (multivariate) normal distribution

→ e_{ti} ~ N(0, **R**) where **R** = σ_e^2

→ e_{ti} has a mean = 0 and some estimated variance(s) and potentially covariances as well (an empirical question)
(if separate variances, each is assumed normally distributed)

Level-2 $\mathbf{U}_i$'s → multivariate normal distribution

→ $U_{0i}, U_{1i}, \ldots$ ~ N(0, **G**)

→ If random intercept: $G = \begin{pmatrix} \tau^2_{U0} \end{pmatrix}$ If random slopes: $G = \begin{pmatrix} \tau^2_{U0} & \\ \tau_{U01} & \tau^2_{U1} \end{pmatrix}$

→ U's EACH have a mean = 0 and some estimated variance, with estimated covariances between them

- The actual mean of U has another name: ___________
- Covariances not included by default: added with TYPE=UN

3 Solutions for Non-Normality

1. <u>Pick a new model for the level-1 e_{ti} residuals</u>
   - **General*ized* linear mixed models** to the rescue!
     - Binary → Logit or Probit, Ordinal → Cumulative Logit
     - Count → Poisson or Negative Binomial (+ Zero-Inflated versions)
   - Unfortunately, level-2 **U**'s are still assumed multivariate normal
     - Problems with skewness → random effects CI's go out of bounds
   - Tricky to estimate – should use ML with numeric integration when possible (try to avoid "pseudo" or "quasi" ML options)

2. <u>Transform your data… carefully if at all...</u>
   - Assumptions apply to residuals, not to data!
   - Complicates interpretations (linear relationships → nonlinear)
   - Inherently subjective (especially "outlier" removal)
     - Check for extreme leverage on solution instead via INFLUENCE options after / on MODEL statement in PROC MIXED

3. Robust ML for Non-Normality

- **MLR in Mplus**: ≈ Yuan-Bentler T_2 (permits MCAR or MAR missing)
  - Same estimates and -2LL, corrected standard errors for all model parameters
- **χ^2-based fit statistics** are adjusted based on an estimated **scaling factor:**
  - Scaling factor = 1.000 = perfectly multivariate normal = same as ML
  - Scaling factor > 1.000 = leptokurtosis (too-fat tails; fixes too big χ^2)
  - Scaling factor < 1.000 = platykurtosis (too-thin tails; fixes too small χ^2)
- **SEs** computed with Huber-White 'sandwich' estimator → uses an information matrix from the variance of the partial first derivatives to correct the information matrix from the partial second derivatives
  - Leptokurtosis (too-fat tails) → increases information; fixes too small SEs
  - Platykurtosis (too-thin tails) → lowers information; fixes too big SEs
- In **SAS**: use "EMPIRICAL" option in PROC MIXED line
  - SEs are computed the same way but for fixed effects only, but can be unstable in unbalanced data, especially in small samples
  - SAS does not provide the needed scaling factor to adjust -2ΔLL test (not sure if this is a problem if you just use the fixed effect *p*-values)

Independence of Residuals At Level 1:

- Level-1 e_{ti} residuals are uncorrelated across level-1 units
  - Once random effects are modeled, residuals of the occasions from the same person are no longer correlated
- Solution for clustered or longitudinal models**:**
  - Choose the 'right' specification of random effects
    - Random effects go in **G**; **what's left in R is uncorrelated across obs**
- Another solution for longitudinal models:
  - Choose the 'right' alternative for the structure of the residual variances and covariances over time
  - Use **R** matrix or **G** and **R** matrices to better approximate observed data:
    - Are the residuals still correlated (AR1, TOEP) after random effects?
    - Are the variances over time homogeneous or heterogeneous?
      - This falls under the "constant variance" assumption – more on that later

Independence of Residuals At Level 2:

- Level-2 $\mathbf{U}_i$'s are uncorrelated across level-2 units
  - Implies no additional effects of clustering/nesting across persons after controlling for person-level predictors
- Two alternatives to deal with additional clustering/nesting:
  - Via fixed effects: Add dummy codes as level-2 predictors
    - Adjusts model for mean differences,
      but DOES NOT allow you to predict those mean differences
  - Via random effects: Add more levels (e.g., for family, group)
    - Adjusts model for mean differences,
      and it DOES allow you to predict those mean differences
    - Like adding another part to **G**

Independence of Residuals Across Levels:

- Level-1 e_{ti} residuals and Level-2 $\mathbf{U}_i$'s are also uncorrelated
  - Implies that what's left over at level-1 is not related to what's left over at level 2
  - Could be violated if level-2 effects are not modeled separately from level-1 effects (i.e., if convergence of level-1 predictors is assumed when it shouldn't be)

- Solution: Don't smush anything!
  - Allow different effects across upper levels for any lower-level predictor with respect to both main effects and interactions

Independence and Constant Variance of Residuals Across Predictors:

- **Level-1 e_{ti} residuals** are flat with constant variance across **level-1 X's**
  - Implies no remaining relationship of X-Y **at level 1**
  - Specific example: level-1 residuals are flat and even across time after fixed and random effects (but we can fit separate variances by time if needed)
  - Check for potential nonlinear effects of level-1 predictors
- **Level-2 U_i's** are flat with constant variance across **level-1 X's**
  - Only possible relation between level-2 U_i and level-1 X is through relationship between level-2 PMx and level-2 U_i (so include PMx to avoid smushing)
- **Level-1 e_{ti} residuals** are flat with constant variance across **level-2 X's**
  - If not, we can fit a heterogeneous variance model instead (stay tuned)
- **Level-2 U_i's** are flat with constant variance across **level-2 X's**
  - Implies no remaining relationship of X-Y **at level 2**
  - Check for potential nonlinear effects of level-2 predictors
  - If not, we can fit a heterogeneous variance model instead (stay tuned)

Constant Variance of Residuals Across Sampling Units:

- **Level-2 U_i's** have constant variance across **level-2 units**
  - Implies no subgroups of individuals or groups that are more or less variable in terms of their distributions of random effects
  - If not, we can fit a heterogeneous variance model instead (stay tuned)
- **Level-1 e_{ti} residuals** have constant variance across **level-2 units***
  - Implies equal unexplained within-person variability across persons
  - Check for missing random effects of level-1 X's or cross-level interactions
  - If not, we can fit a heterogeneous variance model instead (stay tuned)
- **Level-1 e_{ti} residuals** have constant variance across **level-1 units**
  - Implies equal unexplained within-person variability across occasions
  - Can add additional random slopes for time or fit a heterogeneous variance model instead (e.g., TOEPH instead of TOEP, data permitting)
- * Test for heterogeneity of level-1 residuals applicable sometimes (see Snijders & Bosker, 1999, p. 126-7)

Assumptions of MLM: Summary

- Because model estimates, SEs, and fit statistics are derived from likelihood estimation using the multivariate normal distribution, their accuracy depends on its assumptions being met:
  - Residuals at each level (level 1 = e_{ti} values, level 2 = $\mathbf{U}_i$ values) are
    (1) normally distributed,
    (2) uncorrelated at each level and across levels,
    ($\mathbf{U}_i$ values can be correlated within their level), and
    (3) equally distributed across X's at each level and across levels.
- If not:
  (1) transform the data (meh) or pick a generalized model for non-linear outcomes (better when possible), or use robust ML for corrected SE's
  (2) add fixed or random effects (or a correlation over time),
  (3) make sure predictive relationships are correctly specified, and then consider heterogeneous variance models if needed.