Random Intercept and Slope Models

Edps/Psych/Soc 587

Carolyn J. Anderson

Department of Educational Psychology
ILLINOIS

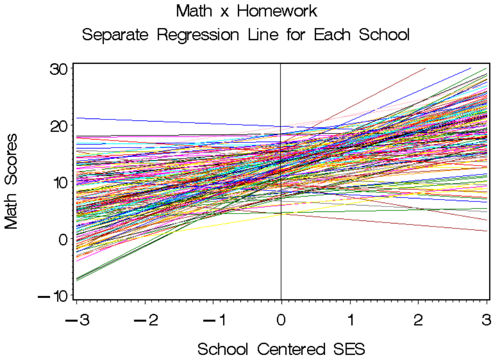
Outline

- Random intercepts and slopes with one mirco-variable.
- Multiple Micro-level variables.
- Discrete Random Effects.
- NELS Example.
- Cross-level interactions (macro level variables).
- Centering Level 1 Variables.
- SAS/MIXED.
- Computer Lab Session 2.

Snijders & Bosker (2012) Chapter 5

A Little Motivation

From the HSB data that we looked at before. . .



Random Intercepts & Slopes (1 x_{ij})

The model and it's properties:

Level 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ $i.i.d.$.

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \longleftarrow \text{ this is new} \end{aligned}$$

where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} = \boldsymbol{U}_j \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} \\ \tau_{10} & \tau_1^2 \end{pmatrix} \right) \quad iid$$

and independent with R_{ij} .

Level 2 Models

Notes:

- τ_0^2 is the variance of the level 2 residuals U_{0j} from predicting the level 1 intercept (i.e, β_{0j}).
- τ_1^2 is the variance of the level 2 residuals U_{1j} from predicting the level 1 slope (i.e, β_{1j}).
- τ_{10} is the covariance between U_{0j} and U_{1j}.

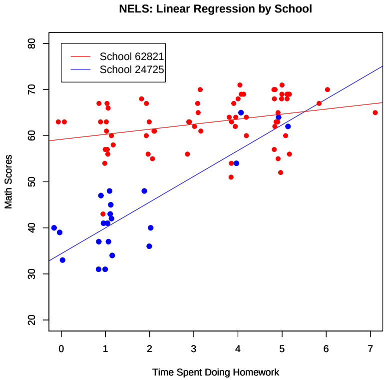
  $\text{cov}(U_{0j}, U_{1j}) = \text{cov}(\beta_{0j}, \beta_{1j}) = \tau_{10}$.

  How to interpret τ_{10}?

Interpreting τ_{10}

Suppose that the following two schools from the 1003 NELS88 are typical of what's in the population.

Would you expect $\tau_{10} <, =,$or > 0?



The Linear Mixed Model

$$\begin{aligned} Y_{ij} &= (\gamma_{00} + U_{0j}) + (\gamma_{10} + U_{1j})x_{ij} + R_{ij} \\ &= \underbrace{\gamma_{00} + \gamma_{10}x_{ij}}_{\text{fixed}} + \underbrace{U_{0j} + U_{1j}x_{ij} + R_{ij}}_{\text{random}} \end{aligned}$$

What's the variance of Y_{ij}?

The Variance of Y_{ij}

$$
\begin{aligned}
\text{var}(Y_{ij}) &\equiv \mathsf{E}[(Y_{ij} - \mathsf{E}(Y_{ij}))^2] \\
&= \mathsf{E}[((\gamma_{00} + \gamma_{10}x_{ij} + U_{0j} + U_{1j}x_{ij} + R_{ij}) - (\gamma_{00} + \gamma_{10}x_{ij}))^2] \\
&= \mathsf{E}[U_{0j}^2 + U_{1j}^2 x_{ij}^2 + R_{ij}^2 + 2U_{0j}U_{1j}x_{ij} + 2U_{0j}R_{ij} + 2U_{1j}x_{ij}R_{ij}] \\
&= \mathsf{E}[U_{0j}^2] + \mathsf{E}[U_{1j}^2 x_{ij}^2] + \mathsf{E}[R_{ij}^2] + \mathsf{E}[2U_{0j}U_{1j}x_{ij}] \\
&= \mathsf{E}[(U_{0j} - 0)^2] + x_{ij}^2\mathsf{E}[(U_{1j} - 0)^2] + \mathsf{E}[(R_{ij}^2 - 0)^2] \\
&\quad + 2x_{ij}\mathsf{E}[(U_{0j} - 0)(U_{1j} - 0)] \\
&= \tau_0^2 + 2\tau_{10}x_{ij} + \tau_1^2 x_{ij}^2 + \sigma^2
\end{aligned}
$$

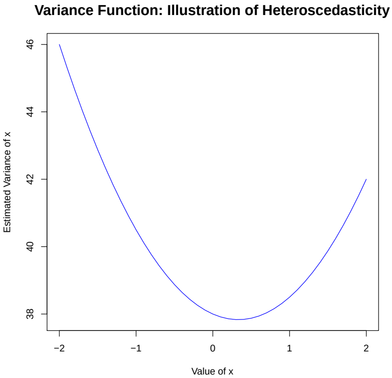
The Marginal Model

$$Y_{ij} \sim \mathcal{N}((\gamma_{00} + \gamma_{10}x_{ij}), (\tau_0^2 + 2\tau_{10}x_{ij} + \tau_1^2 x_{ij}^2 + \sigma^2))$$

- For normally distributed data, all we have to do is estimate the mean and variance.
- The variance is not constant over $x_{ij} \rightarrow$ "heteroscedastic."
- var(Y_{ij}) is a quadratic function of x_{ij}.

Heteroscedasticity

Example of variance of Y_{ij} plotted versus x_{ij} where $\tau_0^2 = 8$, $\tau_1^2 = 1.5$, $\tau_{10} = -.5$, & $\sigma^2 = 30$.



Within Group Covariance

What about the covariance of observations within the same group?

$$\begin{aligned}
\text{cov}(Y_{ij}, Y_{i'j}) &\equiv \mathsf{E}[((Y_{ij} - \mathsf{E}(Y_{ij}))(Y_{i'j} - \mathsf{E}(Y_{i'j})))] \\
&= \mathsf{E}[(\gamma_{00} + \gamma_{10}x_{ij} + U_{0j} + U_{1j}x_{ij} + R_{ij}) - (\gamma_{00} + \gamma_{10}x_{ij}) \\
&\quad \times(\gamma_{00} + \gamma_{10}x_{i'j} + U_{0j} + U_{1j}x_{i'j} + R_{i'j}) - (\gamma_{00} + \gamma_{10}x_{i'j})] \\
&= \mathsf{E}[(U_{0j} + U_{1j}x_{ij} + R_{ij})(U_{0j} + U_{1j}x_{i'j} + R_{i'j})] \\
&\quad \vdots \\
&= \tau_0^2 + \tau_{10}(x_{ij} + x_{i'j}) + \tau_1^2 x_{ij}x_{i'j}
\end{aligned}$$

It depends on x.

Within Group Correlation

- For random intercept models,

$$\text{cov}(Y_{ij}, Y_{i'j}) = \tau_0^2$$

and the intra-class correlation,

$$\rho_I = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$$

- With random intercepts & slopes, the (intra-class) correlation is not constant,

$$\rho(Y_{ij}, Y_{i'j}) = \frac{\tau_0^2 + \tau_{10}(x_{ij} + x_{i'j}) + \tau_1^2 x_{ij} x_{i'j}}{\sqrt{(\tau_0^2 + 2\tau_{10} x_{ij} + \tau_1^2 x_{ij}^2 + \sigma^2)}\sqrt{(\tau_0^2 + 2\tau_{10} x_{i'j} + \tau_1^2 x_{i'j}^2 + \sigma^2)}}$$

. . . in Matrix Form

$$Y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + U_{0j} + U_{1j}x_{ij} + R_{ij}$$

$$\begin{pmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{n_j j} \end{pmatrix} = \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \\ \vdots & \vdots \\ 1 & x_{n_j j} \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \end{pmatrix} + \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \\ \vdots & \vdots \\ 1 & x_{n_j j} \end{pmatrix} \begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} + \begin{pmatrix} R_{1j} \\ R_{2j} \\ \vdots \\ R_{n_j j} \end{pmatrix}$$

$$\boldsymbol{Y}_j = \boldsymbol{X}_j\boldsymbol{\Gamma} + \boldsymbol{Z}_j\boldsymbol{U}_j + \boldsymbol{R}_j$$

$$\text{and} \qquad \text{var}(Y_{ij}) = \boldsymbol{Z}_j\boldsymbol{T}\boldsymbol{Z}_j' + \sigma^2\boldsymbol{I}$$

The elements of $\boldsymbol{Z}_j$ are not the z_j's in the level 2 models. $\boldsymbol{Z}_j$ is a sub-matrix of $\boldsymbol{X}_j$.

The Marginal Model

$$Y_{ij} \sim \mathcal{N}(\gamma_{00} + \gamma_{10} x_{ij}, (\tau_0^2 + 2\tau_{01} x_{ij} + \tau_2^2 x_{ij}^2 + \sigma^2))$$

. . . in terms of matrix notation,

$$\boldsymbol{Y}_j \sim \mathcal{N}(\boldsymbol{X}_j \boldsymbol{\Gamma}, (\boldsymbol{Z}_j \boldsymbol{T} \boldsymbol{Z}_j' + \sigma^2 \boldsymbol{I}))$$

Example 1: High School and Beyond

Y_{ij} = Math achievement (outcome/response).
x_{ij} = Group centered SES (micro-level variable).

Level 1:

$$\text{math}_{ij} = \beta_{0j} + \beta_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ $i.i.d.$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \end{aligned}$$

where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} \\ \tau_{10} & \tau_1^2 \end{pmatrix}\right)$$

or in matrix terms $\boldsymbol{U}_j \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$ and $\boldsymbol{U}_j$ is independent over j and of R_{ij}'s.

The Linear Mixed & Marginal Models

The Linear Mixed Model:

$$\text{math}_{ij} = \underbrace{\gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j)}_{\text{Fixed}} + \underbrace{U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}}_{\text{Random}}$$

The Marginal Model:

$$Y_{ij} \sim \mathcal{N}(\mu_{ij}, \text{var}(Y_{ij}))$$

where $\mu_{ij} = \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j)$, and

$$\text{var}(Y_{ij}) = \tau_0^2 + 2\tau_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \tau_1^2(\text{SES}_{ij} - \overline{\text{SES}}_j)^2 + \sigma^2$$

(Edited) Summary R/lmer Results

```
>summary(model2 ← lmer(mathach ∼ 1 + ses.centered
                 + (1 + ses.centered| id), data=hsb,
                 REML=FALSE))
```

```
Linear mixed model fit by maximum likelihood.
t-tests use Satterthwaite's method [lmerModLmerTest]
Formula:  mathach ∼ 1 + ses.centered +
                 (1 + ses.centered | id)
Data:  hsb

     AIC       BIC    logLik  deviance  df.resid
 46723.0  46764.3  -23355.5   46711.0      7179
```

(Edited) Summary R/lmer Results

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
id	(Intercept)	8.6213	2.9362	
	ses.centered	0.6783	0.8236	0.02
Residual		36.7000	6.0581	

Number of obs: 7185, groups: id, 160

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> $\|t\|$)	
(Intercept)	12.6494	0.2437	157.7206	51.90	<2e-16	***
ses.centered	2.1931	0.1278	156.1499	17.15	<2e-16	***

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

R/lmer Results

Using texreg (most of you could use screenreg in the texreg package)

	RI & S cses
(Intercept)	$12.65\ (0.24)^{***}$
ses.centered	$2.19\ (0.13)^{***}$
AIC	46722.98
BIC	46764.26
Log Likelihood	-23355.49
Num. obs.	7185
Num. groups: id	160
Var: id (Intercept)	8.62
Var: id ses.centered	0.68
Cov: id (Intercept) ses.centered	0.05
Var: Residual	36.70

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

(Edited) SAS/MIXED Results

Dimensions

Covariance Parameters	4	$\longrightarrow$ what are they?
Columns in X	2	
Columns in Z Per Subject	2	
Subjects	160	
Max Obs Per Subject	67	
Observations Used	7185	
Observations Not Used	0	
Total Observations	7185	

Convergence criteria met.

Covariance Parameter Estimates

Covariance Parm		Subject	Estimate	Standard Error	Z Value	Pr Z
τ_{00}	UN(1,1)	id	8.6161	1.0684	8.06	$< .0001$
τ_{01}	UN(2,1)	id	0.05041	0.4024	0.13	.9003
τ_{11}	UN(2,2)	id	0.6782	0.2780	2.44	.0073
σ^2	Residual		36.7005	0.6257	58.65	$< .0001$

Solution for Fixed Effects

	Effect	Estimate	Standard Error	DF	t Value	Pr> $|t|$
γ_{00}	Intercept	12.6494	0.2437	159	51.91	$< .0001$
γ_{10}	cses	2.1931	0.1278	159	17.15	$< .0001$

The average regression line (i.e., fixed part of the model):

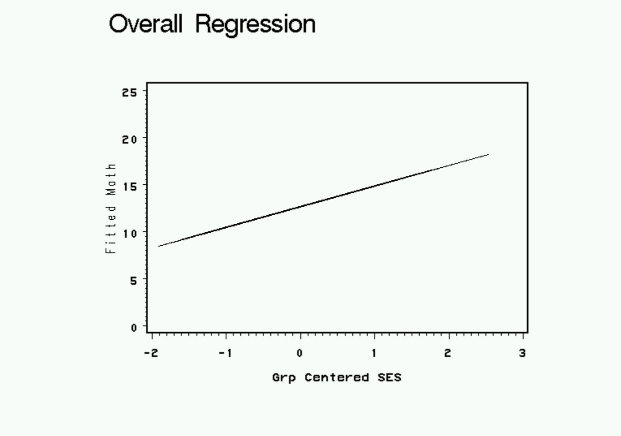
$$\widehat{\text{math}}_{ij} = 12.6494 + 2.1931(\text{SES}_{ij} - \overline{\text{SES}}_j),$$

which is almost the same as the random intercept model:

$$\widehat{\text{math}}_{ij} = 12.6494 + 2.1912(\text{SES}_{ij} - \overline{\text{SES}}_j),$$

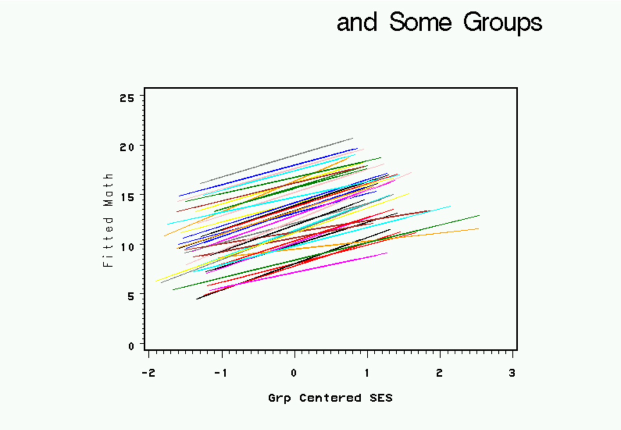
The Estimated Average Regression Line

$$\widehat{\text{math}}_{ij} = 12.65 + 2.19(\text{SES}_{ij} - \overline{\text{SES}}_j)$$

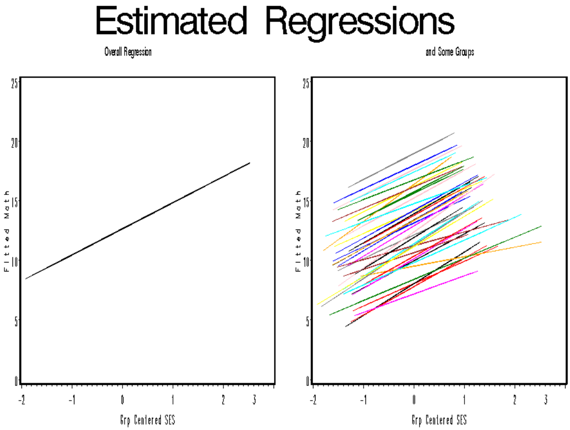


The Estimated Group Regression Lines

$$\widehat{\mathsf{math}}_{ij} = 12.65 + 2.19(\mathsf{SES}_{ij} - \overline{\mathsf{SES}}_j) + \hat{U}_{0j} + \hat{U}_{1j}(\mathsf{SES}_{ij} - \overline{\mathsf{SES}}_j)$$



Overall & Group Regressions



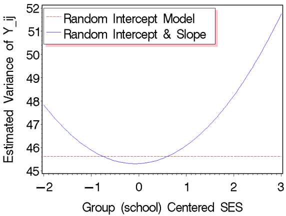
Estimated Variance of Y_{ij} (math scores)

Random Intercept Model:

$$\widehat{\text{var}}(\text{math}_{ij}) = 8.61 + 37.01 = 45.62$$

Random Intercept and Slope Model

$$\widehat{\text{var}}(\text{math}_{ij}) = 8.61 + 2(.05)x_{ij} + .68x_{ij}^2 + 36.70$$



Summary & Comparison (R)

	Null	RI	RI r & S
(Intercept)	12.64 (0.24)	12.65 (0.24)	12.65 (0.24)
ses.centered		2.19 (0.11)	2.19 (0.13)
AIC	47121.81	46728.41	46722.98
BIC	47142.45	46755.93	46764.26
Log Likelihood	-23557.91	-23360.21	-23355.49
Num. obs.	7185	7185	7185
Num. groups: id	160	160	160
Var: id (Intercept)	8.55	8.61	8.62
Var: Residual	39.15	37.01	36.70
Var: id ses.centered			0.68
Cov: id (Intercept) ses.centered			0.05

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Summary & Comparison (SAS)

	(Baseline) Null model		Random intercept		Random intercept and slope	
	Value	SE	Value	SE	Value	SE
Fixed Effects						
γ_{00}	12.64	.24	12.65	.24	12.65	.24
γ_{10}	—	—	2.19	.11	2.19	.13
Random Effects						
τ_0^2	8.55	1.07	8.61	1.07	8.61	1.07
τ_{10}	—	—	—	—	.05	.40
τ_1^2	—	—	—	—	.68	.28
σ^2	39.15	.66	37.01	.63	36.70	.63

Example 2: NELS88

- Sample $N = 23$ schools from the full 1003 schools.
- Y_{ij}, Math achievement is the response/outcome variable.
- x_{ij}, Homework (time spent doing math homework) is the explanatory variable.

The models fit to these data:

- Null model (Baseline).
- Random intercept.
- Random intercept and slope models (new one).

NELS88: Null Model

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z Per Subject	1
Subjects	23
Max Obs Per Subject	67
Observations Used	519
Observations Not Used	0
Total Observations	519

Convergence criteria met.

NELS88: Null Model

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	SCHOOL	24.8503	8.4079	2.96	0.0016
Residual		81.2374	5.1530	15.77	$< .0001$

So the intra-class correlation is

$$\hat{\rho}_I = \frac{24.8503}{24.8503 + 81.2374} = .23$$

NELS88: Null Model

Solution for the fixed effect:

$$\hat{\gamma}_{00} = 50.7574$$

with standard error $= 1.1267$ ($df = 22$, $t = 45.05$, $p < .0001$).

NELS88: Random Intercept Model

Add time spent doing homework,

Level 1:

$$(\text{math})_{ij} = \beta_{0j} + \beta_{1j}(\text{homew})_{ij} + R_{ij} \qquad R_{ij} \sim N(0, \sigma^2)$$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} \end{aligned}$$

Linear mixed model is

$$\text{math}_{ij} = \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + U_{0j} + R_{ij}$$

where

$$\begin{pmatrix} U_{0j} \\ R_{ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right) \qquad i.i.d.$$

Solution: Random Intercept Model

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	SCHOOL	20.2251	7.0708	2.86	0.0021
Residual		71.1440	4.5172	15.75	$< .0001$

The reduction of τ_0^2 & σ^2 relative to null model (no level 1 predictor):

For τ_0^2, 20.23 versus 24.85 $\longrightarrow$ level 1 predictor reduced τ_0^2

For σ^2, 71.14 versus $81.24.$ $\longrightarrow$ level 1 predictor reduced σ^2

Solution: Random Intercept Model

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	$Pr > \|t\|$
Intercept	46.3494	1.1411	22	40.62	$< .0001$
HOMEW	2.4020	0.2768	495	8.68	$< .0001$

The estimated overall regression,

$$\widehat{(\text{math})}_{ij} = 46.35 + 2.40(\text{homew})_{ij}$$

Random Intercept & Slope Model

Question: Is the effect of time spent doing homework different between schools?

- In some schools, don't need to rely on homework (tutoring, curriculum, class size, etc.).
- The effect of $(\text{homew})_{ij}$ is different in different schools.
- Add random slope for time spent doing homework.

Random Intercept & Slope Model

The model for the NELS88 ($N = 23$ schools):

Level 1: $\text{math}_{ij} = \beta_{0j} + \beta_{1j}(\text{homew})_{ij} + R_{ij}$ and $R_{ij} \sim N(0, \sigma^2)$.

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \end{aligned}$$

where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ R_{ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} & 0 \\ \tau_{10} & \tau_1^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix} \right) \qquad i.i.d.$$

Random Intercept & Slope Model

Linear Mixed Model:

$$(\text{math})_{ij} = \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + U_{0j} + U_{1j}(\text{homew})_{ij} + R_{ij}$$

What's the distribution of random effects?

Marginal Model:

$$(\text{math})_{ij} \sim \mathcal{N}(\mu_{ij}, \text{var}((\text{math})_{ij}))$$

where

$$\mu_{ij} = (\gamma_{00} + \gamma_{10}(\text{homew})_{ij})$$

and

$$\text{var}((\text{math})_{ij}) = \tau_0^2 + 2\tau_{10}(\text{homew})_{ij} + \tau_1^2(\text{homew})_{ij}^2 + \sigma^2$$

Solution: Random Intercept & Slope

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	SCHOOL	59.2439	19.9598	2.97	0.0015
UN(2,1)	SCHOOL	−26.1242	9.8467	−2.65	0.0080
UN(2,2)	SCHOOL	16.7681	5.8301	2.88	0.0020
Residual		53.3000	3.4665	15.38	$<.0001$

Solution: Random Intercept & Slope

So

$$\hat{\boldsymbol{T}} = \begin{pmatrix} 59.24 & -26.12 \\ -26.12 & 16.77 \end{pmatrix}$$

and note that the correlation matrix for U_{0j} and U_{1j} is

$$\boldsymbol{R}_\tau = \begin{pmatrix} 1 & -.83 \\ -.83 & 1 \end{pmatrix}$$

i.e., $r(U_{0j}, U_{1j}) = \frac{-26.12}{\sqrt{59.24}\sqrt{16.77}} = -.83$

Solution:Random Intercept & Slope

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > $\|t\|$
Intercept	46.3225	1.7190	22	26.95	$< .0001$
HOMEW	1.9870	0.9056	22	2.19	0.0391

The estimated overall regression model is

$$\widehat{(\text{math})}_{ij} = 46.32 + 1.98(\text{homew})_{ij}$$

NELS88 : Summary & Comparison

	Random intercept		Random intercept and slope	
	Value	SE	Value	SE
Fixed Effects				
γ_{00}	46.35	1.4	46.32	1.72
γ_{10} $(\text{homew})_{ij}$	2.40	.28	1.99	.91
Random Effects				
τ_0^2	20.23	7.07	59.24	19.96
τ_{10}			-26.12	9.85
τ_1^2			16.77	5.83
σ^2	71.14	4.52	53.30	3.47

NELS88 : Summary & Comparison

Questions:

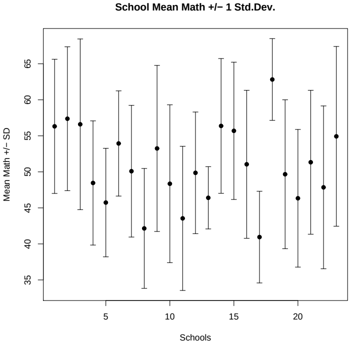
1. In the null model $\hat{\tau}_0^2 = 24.85$ and $\hat{\sigma}^2 = 81.23$, why is there a drop in both of these variances when we add $(\text{homew})_{ij}$ to the model?

2. Do we need to include $(\text{homew})_{ij}$ in the random intercept and slope model (i.e., $\hat{\gamma}_{10}$ is small relative to it's standard error)?

3. Why are the estimates of τ_0^2 and σ^2 so different in the random intercept and slope model relative to the random intercept model?

Question 1

School mean time spent doing homework by school...between & within variability. (note: model with school mean and mean centered is fit when we talk about centering).

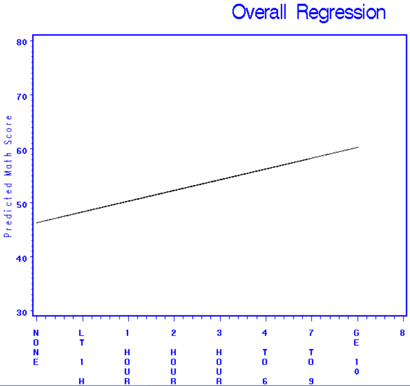


Overall Regression

...in the random intercept and slope model...

NELS88 ($N = 23$): Overall (average) regression line — fixed effects

$$\widehat{\text{math}}_{ij} = 46.32 + 1.99(\text{homew})_{ij}$$

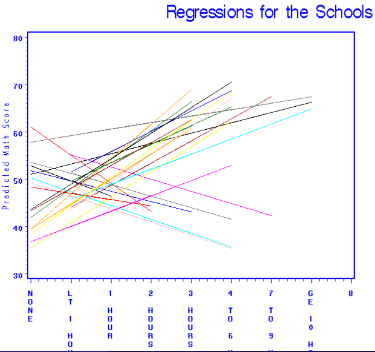


Schools' Regressions

. . . in the random intercept and slope model. . .

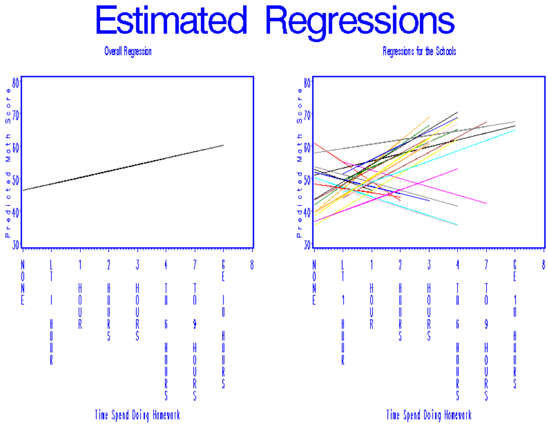
NELS88 ($N = 23$): Overall (average) regression line — fixed effects

$$\widehat{\text{math}}_{ij} = 46.32 + 1.99(\text{homew})_{ij} + \hat{U}_{oj} + \hat{U}_{1j}(\text{homew})_{ij}$$



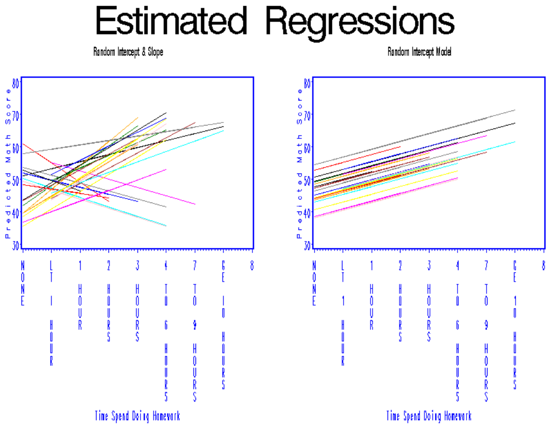
Overall and School Regressions

NELS88 ($N = 23$) Random Intercept & Slope



Random β_{0j} & β_{1j} versus Random β_{0j}

NELS88 ($N = 23$): Why are $\hat{\tau}_0^2$'s so different?



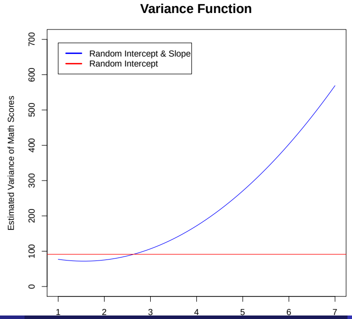
Estimated Variance of Math Scores

Random Intercept Model:

$$\widehat{\mathsf{var}}(Y_{ij}) = 20.2251 + 71.1440$$

Random Intercept and Slope Model:

$$\widehat{\mathsf{var}}(Y_{ij}) = 59.2439 + 2(-26.1242)(\mathsf{hmwk}_{ij}) + 16.7681(\mathsf{hmwk}_{ij})^2 + 53.3000$$



Multiple Micro-Level Variables

Mini-outline

1. The model and it's properties.
2. Example 1: High School & Beyond.
3. Example 2: NELS88.

Multiple Micro-Level Variables

Trying to account for within group differences
(i.e., variability of R_{ij}).

Level 1:

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \ldots + \beta_{pj}x_{pij} + R_{ij} \\ &= \beta_{0j} + \sum_{k=1}^{p} \beta_{kj}x_{kij} + R_{ij} \end{aligned}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ and independent.

Multiple Micro-Level Variables

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \\ \beta_{2j} &= \gamma_{20} + U_{2j} \\ \vdots &\quad \vdots \\ \beta_{pj} &= \gamma_{p0} + U_{pj} \end{aligned}$$

or more compactly,...

$$\beta_{kj} = \gamma_{k0} + U_{kj} \quad \text{for } k = 1, 2, \ldots, p$$

Level 2 Assumptions

where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \\ \vdots \\ U_{pj} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} & \tau_{20} & \cdots & \tau_{p0} \\ \tau_{10} & \tau_1^2 & \tau_{21} & \cdots & \tau_{p1} \\ \tau_{20} & \tau_{21} & \tau_2^2 & \cdots & \tau_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tau_{p0} & \tau_{p1} & \tau_{p2} & \cdots & \tau_p^2 \end{pmatrix} \right),$$

and independent (between macro units) and independent of R_{ij}.

Linear Mixed Model

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{10}x_{1ij} + \gamma_{20}x_{2ij} + \ldots + \gamma_{p0}x_{pij}}_{\text{fixed}}$$

$$+ \underbrace{U_{0j} + U_{1j}x_{1ij} + U_{2j}x_{2ij} + \ldots + U_{pj}x_{pij} + R_{ij}}_{\text{random}}$$

$$= \underbrace{\gamma_{00} + \sum_{k=1}^{p} \gamma_{k0}x_{kij}}_{\text{fixed}} + \underbrace{U_{0j} + \sum_{k=1}^{p} U_{kj}x_{kij} + R_{ij}}_{\text{random}}$$

Linear Mixed Model

in terms of vectors and matrices:

$$\begin{pmatrix} Y_{1j} \\ Y_{2j} \\ \dots \\ Y_{n_jj} \end{pmatrix} = \begin{pmatrix} 1 & x_{11j} & x_{11j} & \dots & x_{p1j} \\ 1 & x_{12j} & x_{12j} & \dots & x_{p2j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_{1n_jj} & x_{1n_jj} & \dots & x_{pn_jj} \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{p0} \end{pmatrix}$$

$$+ \begin{pmatrix} 1 & x_{11j} & x_{11j} & \dots & x_{p1j} \\ 1 & x_{12j} & x_{12j} & \dots & x_{p2j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_{1n_jj} & x_{1n_jj} & \dots & x_{pn_jj} \end{pmatrix} \begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \\ \vdots \\ U_{pj} \end{pmatrix} + \begin{pmatrix} R_{1j} \\ R_{2j} \\ \vdots \\ R_{n_jj} \end{pmatrix}$$

$$\boldsymbol{Y}_j = \boldsymbol{X}_j\boldsymbol{\Gamma} + \boldsymbol{Z}_j\boldsymbol{U}_j + \boldsymbol{R}_j$$

The Marginal Model

$$Y_{ij} \sim \mathcal{N}\left(\mu_{ij}, \text{var}(Y_{ij})\right)$$

where

$$\mu_{ij} = \gamma_{00} + \sum_{k=1}^{p} \gamma_{k0} x_{kij}$$

$$\text{var}(Y_{ij}) = \tau_0^2 + 2\sum_{k=1}^{p} \tau_{k0} x_{kij} + \sum_{k=1}^{p} \tau_k^2 x_{kij}^2 + 2\sum_{k<l} \tau_{kl} x_{kij} x_{lij} + \sigma^2$$

In terms of vectors and matrices:

$$\boldsymbol{Y}_j \sim \mathcal{N}\left((\boldsymbol{X}_j \boldsymbol{\Gamma}), (\boldsymbol{Z}_j \boldsymbol{T} \boldsymbol{Z}_j' + \sigma^2 \boldsymbol{I})\right).$$

- Where $\boldsymbol{T}$ is the covariance matrix for $\boldsymbol{U}_j$.
- R_{ij}'s and U_{kj} are independent of each other (i.e., within and between group residuals are independent)— Assumed this to derive the marginal model.
- $\text{cov}(Y_{ij}, Y_{ij'}) = 0$ (i.e, observations from different groups are independent).

Example 1: HSB again

Y_{ij} = math test scores = response variable.

The x_{kij}'s:

- Student's SES relative to average of school;

  i.e., Group centered SES: $(\text{SES}_{ij} - \overline{\text{SES}}_j)$ ... numerical

- Student's Gender: $\text{gender}_{ij} = 1$ if female, $= 0$ if male ... discrete

- Student's Ethnicity: ethnic $= 1$ if minority, $= 0$ if not (discrete).

Example 1: HSB again

The Plan:

1. Add the variables gender and ethnicity as fixed effects.
2. Make the slopes random.

HSB: Fixed Gender & Ethnicity

Level 1:

$$\begin{aligned} \text{math}_{ij} &= \beta_{0j} + \beta_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \beta_{2j}(\text{gender})_{ij} \\ &\quad + \beta_{3j}(\text{ethnic})_{ij} + R_{ij} \end{aligned}$$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \end{aligned}$$

Assumptions:

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ R_{ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} & 0 \\ \tau_{10} & \tau_1^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix} \right) \qquad i.i.d.$$

HSB: Fixed Gender & Ethnicity

Linear Mixed Model:

$$
\begin{aligned}
\text{math}_{ij} &= \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) \\
&\quad + \gamma_{20}(\text{gender})_{ij} + \gamma_{30}(\text{ethnic})_{ij} \\
&\quad + U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}
\end{aligned}
$$

What's the Marginal Model?

HSB: Model Information

Data Set	WORK.HSBCENT
Dependent Variable	mathach
Covariance Structure	Unstructured
Subject Effect	id
Estimation Method	ML
Residual Variance Method	Profile
Fixed Effects	SE
Method	Model-Based
Degrees of Freedom Method	Containment

HSB: Dimensions

Covariance Parameters	4	$\longrightarrow$ what are they?
Columns in X	4	$\longrightarrow$ what are they?
Columns in Z Per Subject	2	$\longrightarrow$ what are they?
Subjects	160	
Max Obs Per Subject	67	
Observations Used	7185	
Observations Not Used	0	
Total Observations	7185	

Convergence criteria met.

HSB: Covariance Parameters Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	6.2575	0.8054	7.77	$< .0001$
UN(2,1)	id	-0.3014	0.3314	-0.91	0.3631
UN(2,2)	id	0.4798	0.2521	1.90	0.0285
Residual		35.6639	0.6083	58.63	$< .0001$

HSB: Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr $> \|t\|$
Intercept	14.1423	0.2350	159	60.18	$< .0001$
cses	1.8945	0.1224	159	15.47	$< .0001$
female	-1.2199	0.1643	6863	-7.42	$< .0001$
minority	-3.1191	0.2102	6863	-14.84	$< .0001$

All Look Good.

HSB: Summary

		Null Model value	Random Intercept value	Random Intercept and Slope Models value	SE	value	SE
Fixed Effects							
γ_{00}	intercept	12.64	12.65	12.65	.24	14.14	.24
γ_{10}	cses		2.19	2.19	.13	1.89	.12
γ_{20}	female					-1.22	.16
γ_{20}	minority					-3.12	.21
Random Effects							
τ_0^2	intercept	8.55	8.61	8.61	1.07	6.26	0.81
τ_{10}				.05	.40	-0.30	0.33
τ_1^2	cses			.68	.28	0.48	0.25
σ^2		39.15	37.01	36.70	.63	35.66	0.61

HSB: Interpretation

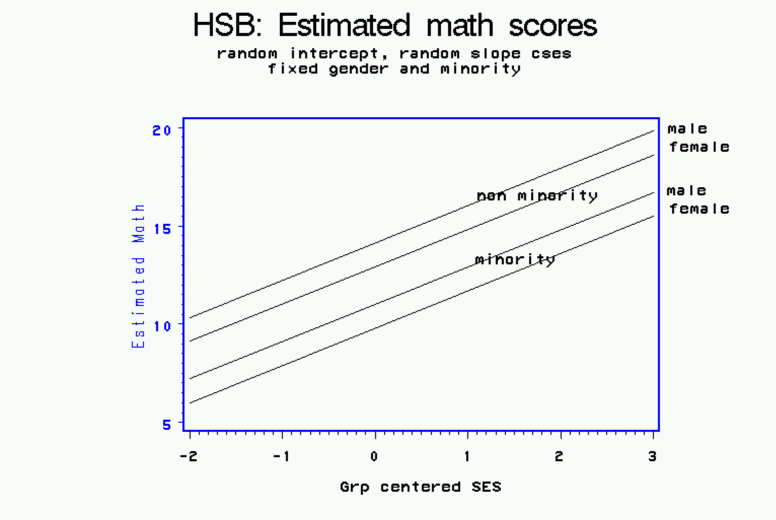
Random Effects : The larger a school's intercept (i.e., U_{0j}), the smaller the school's slope (i.e., U_{1j}). Is this a valid conclusion??

Fixed Effects: Linear model,

$$\widehat{(\text{math})}_{ij} = 14.14 + 1.89(\text{SES}_{ij} - \overline{\text{SES}}_j) - 1.22(\text{gender})_{ij} - 3.12(\text{ethnic})_{ij}$$

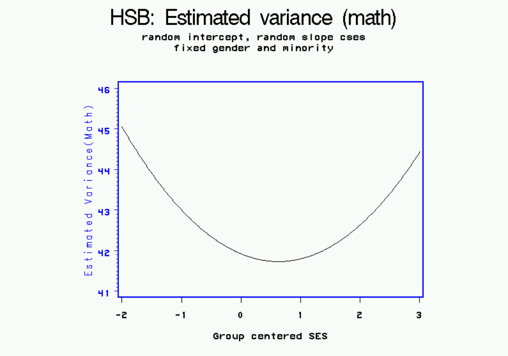
$$\widehat{\text{math})}_{ij} = \begin{cases} 9.8 + 1.89(\text{SES}_{ij} - \overline{\text{SES}}_j) & \text{for minority females} \\ 11.02 + 1.89(\text{SES}_{ij} - \overline{\text{SES}}_j) & \text{for minority males} \\ 12.92 + 1.89(\text{SES}_{ij} - \overline{\text{SES}}_j) & \text{for non-minority females} \\ 14.14 + 1.89(\text{SES}_{ij} - \overline{\text{SES}}_j) & \text{for non-minority males} \end{cases}$$

HSB: Interpretation (continued)



Estimated Variance of Math Scores

$$\begin{aligned}\widehat{\text{var}}(\text{math}_{ij}) &= \hat{\tau}_0^2 + 2\hat{\tau}_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \hat{\tau}_1^2(\text{SES}_{ij} - \overline{\text{SES}}_j)^2 + \sigma^2 \\ &= 6.26 - .60(\text{SES}_{ij} - \overline{\text{SES}}_j) + .48(\text{SES}_{ij} - \overline{\text{SES}}_j)^2 + 35.66\end{aligned}$$



Minimum occurs at $\frac{-\tau_{10}}{\tau_1^2} = \frac{.30}{.48} = .625$

Random Categorical Predictors

- In the HSB data set we saw that gender and minority are potential predictors of math.
- Adding qualitative predictors is same as in regular linear regression,

$$\text{Female}_{ij} = \begin{cases} 1 & \text{if female} \\ 0 & \text{if male} \end{cases} \quad \& \quad \text{Minority}_{ij} = \begin{cases} 1 & \text{if minority} \\ 0 & \text{if non-minority} \end{cases}$$

- Making a qualitative predictor random can be tricky.
  - How does program handles categorical variables?
  - In SAS, you can fit equivalent models with random categorical predictors that either include or do not include a "random intercept".
  - Always check 'log' file.

Different Linear Mixed Models

Define an additional variable,

$$\text{Male}_{ij} = \begin{cases} 1 & \text{if male} \\ 0 & \text{if female} \end{cases}$$

Consider the following models:

$$\text{(a)} \quad \underline{\text{math}}_{ij} = \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{11}\overline{\text{SES}}_j + \gamma_{20}\text{Female}_{ij} + \underline{u}_{0j} + \underline{u}_{1j}\text{Female}_{ij} + \underline{u}_{2j}\text{Male}_{ij} + \sigma^2$$

$$\text{(b)} \quad \underline{\text{math}_{ij}} = \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{11}\overline{\text{SES}}_j + \gamma_{20}\text{Female}_{ij} + \underline{u}_{0j} + \underline{u}_{1j}\text{Female}_{ij} + \sigma^2$$

$$\text{(c)} \quad \underline{\text{math}}_{ij} = \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{11}\overline{\text{SES}}_j + \gamma_{20}\text{Female}_{ij} + \underline{u}_{1j}\text{Female}_{ij} + \underline{u}_{2j}\text{Male}_{ij} + \sigma^2$$

Problem with One of Them

- Model (a): In the log file you will find the message

  "NOTE: Convergence criteria met but final hessian is not positive definite"

- The hessian is a matrix that is inverted in the estimation algorithm. If it is "not positive definite", it cannot be inverted. *In this case*, it indicates that some effects are (perfectly or highly) correlated.

  For males, the intercept is $\underline{u}_{0j} + \underline{u}_{2j}$
  For females, the intercept is $\underline{u}_{0j} + \underline{u}_{1j}$

- In Model (a), one of the random effects needs to be dropped: $\underline{u}_{0j}$, $\underline{u}_{1j}$ or $\underline{u}_{2j}$. These options lead to models (b) and (c).

HSB: Random Slopes for All

Level 1:

$$\text{math}_{ij} = \beta_{0j} + \beta_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \beta_{2j}(\text{gender})_{ij} + \beta_{3j}(\text{ethnic})_{ij} + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$.

Level 2:

$$\begin{aligned}
\beta_{0j} &= \gamma_{00} + U_{0j} \\
\beta_{1j} &= \gamma_{10} + U_{1j} \\
\beta_{2j} &= \gamma_{20} + U_{2j} \\
\beta_{3j} &= \gamma_{30} + U_{3j}
\end{aligned}$$

where. . .

HSB: Random Slopes for All

...where $\boldsymbol{U}_j \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$ $i.i.d.$

Note: $\boldsymbol{T}$ is (4×4).

$$\boldsymbol{T} = \begin{pmatrix} \tau_0^2 & \tau_{10} & \tau_{20} & \tau_{30} \\ \tau_{10} & \tau_1^2 & \tau_{12} & \tau_{31} \\ \tau_{20} & \tau_{12} & \tau_2^2 & \tau_{32} \\ \tau_{30} & \tau_{32} & \tau_{23} & \tau_3^2 \end{pmatrix}$$

Linear mixed model:

$$\begin{aligned} \text{math}_{ij} &= \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{20}(\text{gender})_{ij} + \gamma_{30}(\text{ethnic})_{ij} \\ & U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + U_{2j}(\text{gender})_{ij} + U_{3j}(\text{ethnic})_{ij} \\ & + R_{ij} \end{aligned}$$

Random Slopes & Discrete Variables

Linear Mixed Model:

$$\begin{aligned}\text{math}_{ij} &= \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{20}(\text{gender})_{ij} + \gamma_{30}(\text{ethnic})_{ij} \\ &\quad U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + U_{2j}(\text{gender})_{ij} + U_{3j}(\text{ethnic})_{ij} + R_{ij}\end{aligned}$$

Minority Female:

$$\begin{aligned}\text{math}_{ij} &= (\gamma_{00} + \gamma_{20} + \gamma_{30}) + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) \\ &\quad +(U_{0j} + U_{2j} + U_{3j}) + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}\end{aligned}$$

Minority Male:

$$\begin{aligned}\text{math}_{ij} &= (\gamma_{00} + \gamma_{30}) + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) \\ &\quad +(U_{0j} + U_{3j}) + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}\end{aligned}$$

Non-Minority Female:

$$\text{math}_{ij} = (\gamma_{00} + \gamma_{20}) + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + (U_{0j} + U_{2j}) + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}$$

Non-Minority Male:

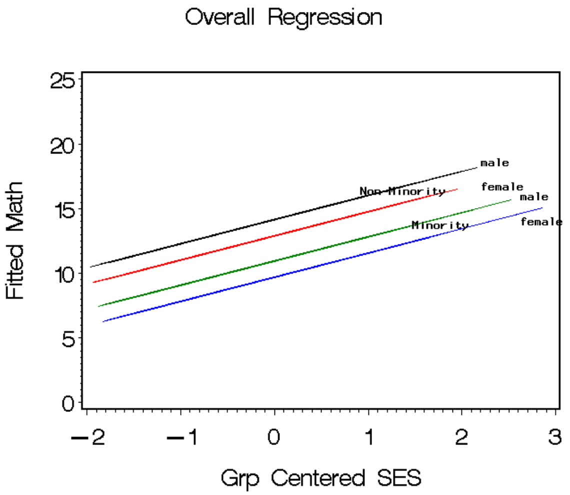
$$\text{math}_{ij} = \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}$$

Results: Random Slopes for All

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr $> \|t\|$
Intercept	14.1501	0.2358	159	60.02	$< .0001$
cses	1.8725	0.1197	159	15.64	$< .0001$
female	−1.2581	0.1832	122	−6.87	$< .0001$
minority	−3.2012	0.2503	135	−12.79	$< .0001$

Results: The Structure



Results: Random Slopes for All

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	6.1724	0.9763	6.32	$< .0001$
UN(2,1)	id	0.04062	0.3734	0.11	0.9134
UN(2,2)	id	0.3847	0.2445	1.57	0.0578
UN(3,1)	id	−0.9950	0.6121	−1.63	0.1041
UN(3,2)	id	−0.1339	0.2825	−0.47	0.6354
UN(3,3)	id	0.8502	0.5588	1.52	0.0640
UN(4,1)	id	0.9799	0.7496	1.31	0.1911
UN(4,2)	id	−0.5698	0.3675	−1.55	0.1210
UN(4,3)	id	0.1566	0.6062	0.26	0.7962
UN(4,4)	id	1.9131	0.9238	2.07	0.0192
Residual		35.3088	0.6116	57.73	$< .0001$

Results: Random Slopes for All

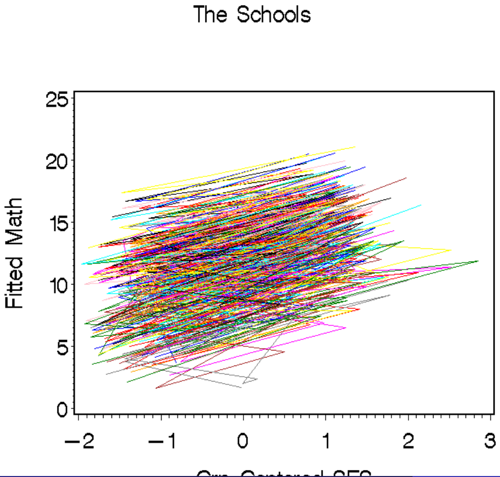
Covariance Parameter Estimates

$$\hat{\boldsymbol{T}} = \begin{pmatrix} \text{intercept} & \text{cses} & \text{female} & \text{minority} \\ 6.1724 & 0.04062 & -0.9950 & 0.9799 \\ & 0.3847 & -0.1339 & -0.5698 \\ & & 0.8502 & 0.1566 \\ & & & 1.9131 \end{pmatrix}$$

and $\hat{\sigma}^2 = 35.3088$.

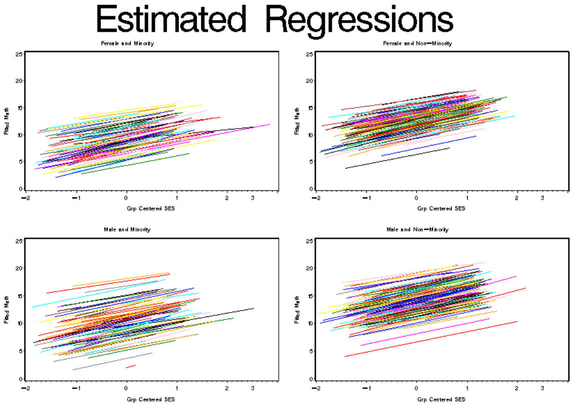
Results: The Stochastic Part

Regressions for each school:



The Stochastic Part (continued)

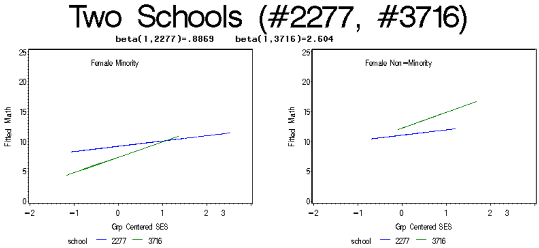
Regressions for each school and gender $\times$ ethnicity combination:



The Stochastic Part (continued)

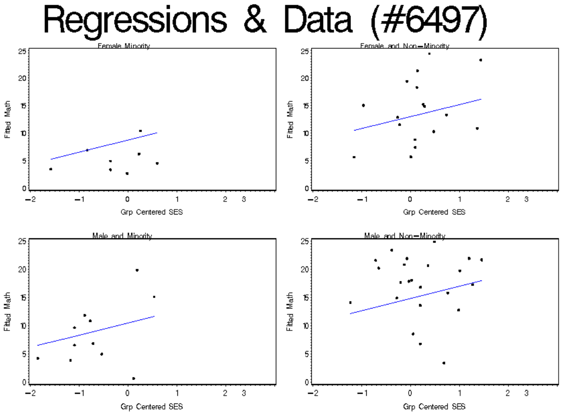
Regressions for two schools – the two most extreme ones in terms of

$$\hat{\beta}_{1j} = \hat{\gamma}_{10} + \hat{U}_{1j} = 1.8725 + \hat{U}_{1j}$$



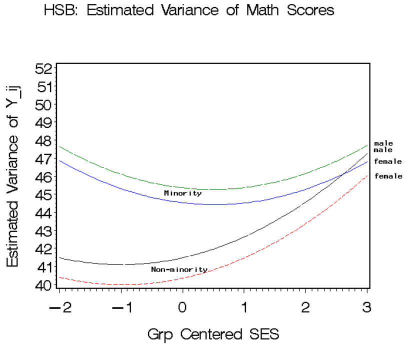
The Stochastic Part (continued)

Data and Regression lines for one school



The Variance of Math Scores

$$\widehat{\text{var}}(\text{math})_{ij} = \hat{\tau}_{00} + \sum_{k=1}^{3} \hat{\tau}_k^2 x_{kij}^2 \quad + 2\sum_{k=1}^{3} \hat{\tau}_{k0} x_{kij} + 2\sum_{k>l} \hat{\tau}_{kl} x_{kij} x_{lij} + \sigma^2$$



HSB: Summary & Comparison

Fixed Effects

		Null Model	Random Intercept	Random Intercept and Slopes				
		value	value	value	SE	value	SE	value
γ_{00}	intercept	12.64	12.65	12.65	.24	14.14	.24	14.15
γ_{10}	cses		2.19	2.19	.13	1.89	.12	1.87
γ_{20}	female					-1.22	.16	-1.26
γ_{20}	minority					-3.12	.21	-3.20

Fixed effects are basically the same for group centered SES.

HSB: Summary & Comparison

		Null Model	Random Intercept	Random Intercept and Slope Models				
		value	value	value	SE	value	SE	value
τ_0^2	intercept	8.55	8.61	8.61	1.07	6.26	0.81	6.17
τ_{10}				.05	.40	-0.30	0.33	.04
τ_1^2	cses			.68	.28	0.48	0.25	.38
τ_{20}								-1.00
τ_{21}								$-.13$
τ_2^2	female							.85
τ_{30}								.98
τ_{31}								$-.57$
τ_{32}								$-.16$
τ_3^2	minority							1.91
σ^2		39.15	37.01	36.70	.63	35.66	0.61	35.31

Example 2: NELS ($N = 23$)

- homew = time spent doing homework.
- minority = white or not (“white” $= 1$ if white,
  $= 0$ otherwise).
- sector = public or not (“private” $= 1$ if private,
  $= 0$ public).

So far the models that were fit to the NELS88

- Null/baseline model.
- Random intercept with homew (time spent doing homework).
- Random intercept and slope for homew.

Example 2: NELS ($N = 23$)

Models that will be examined are random slope for "homew" and

1. Fixed slopes for minority and sector.
2. Random slope for minority but fixed slope for sector.
3. Random slope for sector but fixed slope for minority.

NELS ($N = 23$): Model 1

Random Slope for "homew" but fixed slopes for "white" and "private"

Level 1:

$$(\text{math})_{ij} = \beta_{0j} + \beta_{1j}(\text{homew})_{ij} + \beta_{2j}(\text{white})_{ij} + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ and independent.

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{private})_j + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \\ \beta_{2j} &= \gamma_{20} \end{aligned}$$

NELS ($N = 23$): Model 1

...where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \\ R_{ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} & 0 \\ \tau_{10} & \tau_1^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix} \right)$$

Independence within schools for R_{ij} and

Independence between schools for R_{ij}, U_{0j} and U_{1j}.

NELS ($N = 23$): Model 1

Linear mixed model:

$$\begin{aligned}(\text{math})_{ij} &= \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + \gamma_{20}(\text{white})_{ij} + \gamma_{01}(\text{private})_j \\ &\quad + U_{0j} + U_{1j}(\text{homew})_{ij} + R_{ij}\end{aligned}$$

Marginal Model:

$$(\text{math})_{ij} \sim \mathcal{N}(\mu(\text{math}), \text{var}(\text{math})_{ij}))$$

where

$$\begin{aligned}\mu(\text{math}_{ij}) &= \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + \gamma_{20}(\text{white})_{ij} + \gamma_{01}(\text{private})_j \\ \text{var}(\text{math}_{ij}) &= \tau_0^2 + \tau_1^2(\text{homew})_{ij}^2 + 2\tau_{10}(\text{homew})_{ij} + \sigma^2.\end{aligned}$$

Model 1: Fixed/Structural

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr $>$ $\|t\|$
Intercept	42.7038	1.8349	21	23.27	$<$ .0001
HOMEW	1.9058	0.8819	22	2.16	0.0418
WHITE	3.3588	0.9628	472	3.49	0.0005
PRIVATE	3.9088	1.7171	472	2.28	0.0233

Model 1: Random/Stochastic

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	SCHOOL	52.2813	18.0535	2.90	0.0019
UN(2,1)	SCHOOL	−25.3464	9.2785	−2.73	0.0063
UN(2,2)	SCHOOL	15.8384	5.5157	2.87	0.0020
Residual		52.6378	3.4288	15.35	< .0001

Model 1: Summary

	Null		Random Intercept		Random Intercept and Slope			
	value	SE	value	SE	value	SE	value	SE
Fixed Effects								
int	50.76	1.13	42.70	1.83	46.32	1.72	42.70	1.83
homew			2.40	.28	1.99	.91	1.91	.88
white							3.36	.96
private							3.91	1.72
Random Effects								
τ_0^2	24.85	8.41	20.23	7.07	59.24	19.96	52.28	18.05
τ_{10}					-26.12	9.85	-25.35	9.28
τ_1^2					16.77	5.83	15.84	5.52
σ^2	81.24	5.15	71.14	4.52	53.30	3.47	52.64	3.43

NELS88 ($N = 23$): Model 2

Random slopes for "homew" and "white"

Level 1: (same)

$$(\text{math})_{ij} = \beta_{0j} + \beta_{1j}(\text{homew})_{ij} + \beta_{2j}(\text{white})_{ij} + R_{ij}$$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{private})_j + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \\ \beta_{2j} &= \gamma_{20} + U_{2j} \end{aligned}$$

where ...

NELS88 ($N = 23$): Model 2

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ and independent, and $\boldsymbol{U}_j \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$ and independent over schools (i.e., j) and with respect to R_{ij}'s.

Note:

$$\boldsymbol{U}_j = \begin{pmatrix} U_{0j} \\ U_{1j} \\ U_{2j} \end{pmatrix} \quad \text{and} \quad \boldsymbol{T} = \begin{pmatrix} \tau_0^2 & \tau_{10} & \tau_{20} \\ \tau_{10} & \tau_1^2 & \tau_{12} \\ \tau_{20} & \tau_{12} & \tau_2^2 \end{pmatrix}$$

NELS88 ($N = 23$): Model 2

Linear mixed model:

$$\begin{aligned} (\text{math})_{ij} &= \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + \gamma_{20}(\text{white})_{ij} + \gamma_{01}(\text{private})_{j} \\ &\quad + U_{0j} + U_{1j}(\text{homew})_{ij} + U_{2j}(\text{white})_{ij} + R_{ij} \end{aligned}$$

Marginal Model:

$$(\text{math})_{ij} \sim \mathcal{N}(\mu(\text{math})_{ij}), \text{var}(\text{math})_{ij}))$$

where

$$\mu(\text{math}) = \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + \gamma_{20}(\text{white})_{ij} + \gamma_{01}(\text{private})_{ij}$$

$$\begin{aligned} \text{var}(\text{math})_{ij}) &= \tau_0^2 + \tau_1^2(\text{homew})_{ij}^2 + \tau_2^2(\text{white})_{ij}^2 + 2\tau_{10}(\text{homew})_{ij} \\ &\quad + 2\tau_{20}(\text{white})_{ij} + 2\tau_{12}(\text{homew})_{ij}(\text{white})_{ij} + \sigma^2. \end{aligned}$$

Model 2: Fixed Effects

Solution for Fixed Effects

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr> $\|t\|$
Intercept	43.2422	2.0478	15	21.12	$< .0001$
HOMEW	1.9467	0.8771	22	2.22	0.0371
WHITE	2.6761	1.5013	15	1.78	0.0949
PRIVATE	4.9319	1.5762	457	3.13	0.0019

Model 2: Fixed Effects

Covariance Parameter Estimates

	Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
τ_{00}	UN(1,1)	SCHOOL	64.4072	28.6721	2.25	.0123
τ_{10}	UN(2,1)	SCHOOL	−26.9855	11.3448	−2.38	.0174
τ_{11}	UN(2,2)	SCHOOL	15.6880	5.4592	2.87	.0020
τ_{20}	UN(3,1)	SCHOOL	−20.1742	19.3586	−1.04	.2974
τ_{21}	UN(3,2)	SCHOOL	2.7660	7.2729	0.38	.7037
τ_{22}	UN(3,3)	SCHOOL	24.0397	21.2476	1.13	.1289
τ_{sigma}	Residual		51.1534	3.3845	15.11	$< .0001$

Model 2: Summary Table

		Null value	Random intercept value	value	Random Intercept and Slopes value	SE	value	SE
Fixed Effects								
γ_{00}	int	50.76	46.35	46.32	46.62	2.12	43.24	2.05
γ_{10}	homew		2.40	1.99	1.91	.88	1.95	.88
γ_{20}	white				3.36	.96	2.68	1.50
γ_{01}	private				3.91	1.72	4.93	1.58
Random Effects								
τ_0^2	int	24.85	20.23	59.24	52.28	18.05	64.41	28.67
τ_{10}				-26.12	-25.35	9.28	-26.99	11.34
τ_1^2	homew			16.77	15.84	5.52	15.69	5.46
τ_{20}							-20.17	19.36
τ_{21}							2.77	7.27
τ_2^2	white						24.04	21.25
σ^2	res	81.24	71.14	53.30	52.64	3.43	51.15	3.38

NELS88 ($N = 23$): Model 3

Random slope for "homew" and "private" but fixed slope for "white"

Linear Mixed Model:

$$
\begin{aligned}
(\text{math})_{ij} &= \gamma_{00} + \gamma_{10}(\text{homew})_{ij} + \gamma_{20}(\text{white})_{ij} \\
&\quad +\gamma_{30}(\text{private})_{j} \\
&\quad +U_{0j} + U_{1j}(\text{homew})_{ij} + U_{2j}(\text{private})_{j} \\
&\quad +R_{ij}
\end{aligned}
$$

and the usual assumptions for the residuals.

Model 3: Results

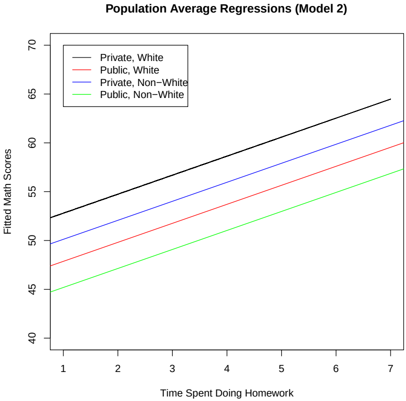
In the SAS ouput file:

Convergence criteria met but final hessian is not positive definite.

- What's wrong with the model we just fit?
- We'll go with Model 2 which has
  - Random intercept with "private" as an explanatory variable for the intercept (i.e., level 2).
  - Random slope for "homew".
  - Fixed slope for "white".

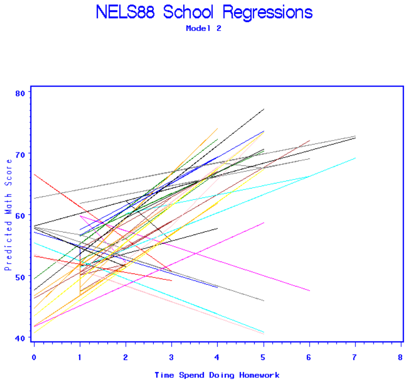
Model 2: For an average school

$$\widehat{\text{math}}_{ij} = 46.62 + 1.91(\text{homew})_{ij} + 3.35(\text{white})_{ij} + 3.90(\text{private})_j$$

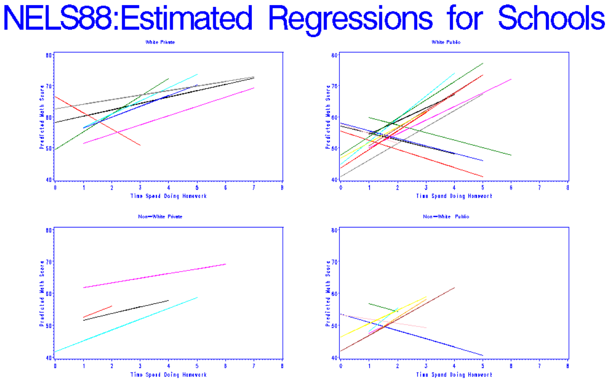


Model 2: For each of 23 schools

$$\begin{aligned}\widehat{\text{math}}_{ij} &= 46.62 + 1.91(\text{homew})_{ij} + 3.35(\text{white})_{ij} + 3.90(\text{private})_j \\ &\quad + \hat{U}_{0j} + \hat{U}_{1j}(\text{homew})_{ij}\end{aligned}$$

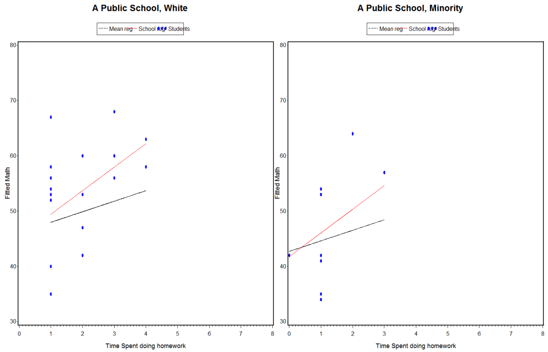


Model 2: A Better Look



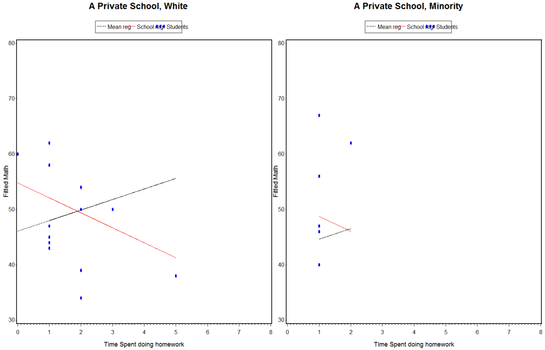
Model 2: A Public School

$$\widehat{\text{math}}_{ij} = 46.62 + 1.91(\text{homew})_{ij} + 3.35(\text{white})_{ij} + \hat{U}_{0j} + \hat{U}_{1j}(\text{homew})_{ij}$$



Model 2: A Private School

$$\widehat{\text{math}}_{ij} = 46.62 + 1.91(\text{homew})_{ij} + 3.35(\text{white})_{ij} + 3.90(\text{private})_j + \hat{U}_{0j} + \hat{U}_{1j}(\text{homew})_{ij}$$



Cross-Level Interactions

- General model.
- Example 1: High School & Beyond.
- Example 2: NELS88 ($N = 23$ schools.

Cross-Level: A Simple Model

Level 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$.

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + U_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}z_j + U_{1j}$$

where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} \\ \tau_{10} & \tau_1^2 \end{pmatrix} \right)$$

Cross-Level: A Simple Model

An aside:

We could have different marco variables in the level 2 regression for the intercepts and slopes, e.g.,

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_{1j} + U_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11} z_{2j} + U_{1j}$$

Linear Mixed Model

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{10}x_{ij} + \gamma_{01}z_j + \gamma_{11}x_{ij}z_j}_{\text{fixed}} + \underbrace{U_{0j} + U_{1j}x_{ij} + R_{ij}}_{\text{random}}$$

Structural Stochastic

and independent over j and with respect to R_{ij}.

Macro level variables in level 2 regression models for the slopes yield "cross-level interaction" terms in the fixed part of the model.

Marginal Model

$$Y_{ij} \sim \mathcal{N}(\mu_{ij}, \text{var}(Y_{ij}))$$

where

$$\mu_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} z_j + \gamma_{11} x_{ij} z_j$$

and

$$\text{var}(Y_{ij}) = \tau_0^2 + 2\tau_{10} x_{ij} + \tau_1^2 x_{ij}^2 + \sigma^2$$

Example: HSB

Level 1:

$$\text{math}_{ij} = \beta_{0j} + \beta_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ and independent.

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SES}})_j + U_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\overline{\text{SES}})_j + U_{1j}$$

where

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} \\ \tau_{10} & \tau_1^2 \end{pmatrix} \right)$$

HSB Example

Linear Mixed Model:

$$
\begin{aligned}
Y_{ij} &= \underbrace{\gamma_{00} + \gamma_{01}(\overline{\text{SES}}_j)}_{\text{Intercept}} + \underbrace{(\gamma_{10} + \overbrace{\gamma_{11}\overline{\text{SES}}_j}^{\text{Moderator}})(\text{SES}_{ij} - \overline{\text{SES}}_j)}_{\text{Slope}} \\
&\quad + U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij} \\
&= \gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{01}(\overline{\text{SES}})_j \\
&\qquad + \gamma_{11}(\text{SES}_{ij} - \overline{\text{SES}}_j)(\overline{\text{SES}})_j \\
&\qquad\qquad + U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}
\end{aligned}
$$

What's the marginal model?

HSB: SAS/MIXED Results

Dimensions

Covariance Parameters	4
Columns in X	4
Columns in Z Per Subject	2
Subjects	160
Max Obs Per Subject	67
Observations Used	7185
Observations Not Used	0
Total Observations	7185

Convergence criteria met.

HSB: Fixed & Random Effects

Solution for Fixed Effects

Effect	Estimate	Standard Error	t Value	Pr> $\|t\|$
Intercept	12.6589	0.1482	85.42	$< .0001$
cses	2.1960	0.1272	17.26	$< .0001$
meanses	5.8700	0.3589	16.36	$< .0001$
cses*meanses	0.2863	0.3169	0.90	0.3663

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	2.6443	0.3963	6.67	$< .0001$
UN(2,1)	id	−0.2575	0.2380	−1.08	0.2794
UN(2,2)	id	0.6496	0.2753	2.36	0.0091
Residual		36.7156	0.6261	58.64	$< .0001$

Summary of Models for HSB

		Null Model	Random Intercept	Random Intercept and Slope			
		value	value	value	SE	value	SE
Fixed Effects							
γ_{00}	intercept	12.64	12.65	12.65	.24	12.66	.15
γ_{01}	cses		2.19	2.19	.13	2.20	.13
γ_{10}	meanses					5.87	.36
γ_{11}	cses*meanses					.29	.32
Random Effects							
τ_0^2	intercept	8.55	8.61	8.61	1.07	2.64	.40
τ_{10}				.05	.40	$-.26$	.24
τ_1^2	cses			.68	.28	.65	.28
σ^2		39.15	37.01	36.70	.63	36.72	.63

Which/what model might be a good one?

Another Model for HSB Data

So far we have fit two random intercept & slopes models

- Simple Random intercept & Slope:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + U_{0j}\\ \beta_{1j} &= \gamma_{10} + U_{1j}\end{aligned}$$

- Random Intercept and Slope using $\overline{(\text{SES}_j)}$:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SES}})_j + U_{0j}\\ \beta_{1j} &= \gamma_{10} + \gamma_{11}(\overline{\text{SES}})_j + U_{1j}\end{aligned}$$

- A third possibility:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SES}})_j + U_{0j}\\ \beta_{1j} &= \gamma_{10} + U_{1j}\end{aligned}$$

Model 3 for the HSB Data

Linear Mixed Model:

$$\text{math}_{ij} = \overbrace{\gamma_{00} + \gamma_{10}(\text{SES}_{ij} - \overline{\text{SES}}_j) + \gamma_{01}(\overline{\text{SES}})_j}^{\text{fixed}} + \underbrace{U_{0j} + U_{1j}(\text{SES}_{ij} - \overline{\text{SES}}_j) + R_{ij}}_{\text{random}}$$

or in matrix terms . . .

Model 3 for the HSB Data

Linear Mixed Model in Matrix Terms

$$\begin{pmatrix} \text{math}_{1j} \\ \text{math}_{2j} \\ \vdots \\ \text{math}_{n_j j} \end{pmatrix} = \begin{pmatrix} 1 & (\text{SES}_{1j} - \overline{\text{SES}}_j) & (\overline{\text{SES}})_j \\ 1 & (\text{SES}_{2j} - \overline{\text{SES}}_j) & (\overline{\text{SES}})_j \\ \vdots & \vdots & \vdots \\ 1 & (\text{SES}_{n_j j} - \overline{\text{SES}}_j) & (\overline{\text{SES}})_j \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{01} \end{pmatrix}$$

$$+ \begin{pmatrix} 1 & (\text{SES}_{1j} - \overline{\text{SES}}_j) \\ 1 & (\text{SES}_{2j} - \overline{\text{SES}}_j) \\ \vdots & \vdots \\ 1 & (\text{SES}_{n_j j} - \overline{\text{SES}}_j) \end{pmatrix} \begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} + \begin{pmatrix} R_{1j} \\ R_{2j} \\ \vdots \\ R_{n_j j} \end{pmatrix}$$

$$\boldsymbol{Y}_j = \boldsymbol{X}_j \boldsymbol{\Gamma} + \boldsymbol{Z}_j \boldsymbol{U}_j + \boldsymbol{R}_j$$

More Linear Algebra

Fact: If $\boldsymbol{Y}$ is a vector of random variables,
$\boldsymbol{L}$ and $\boldsymbol{b}$ are vectors of weights (constants),
$\boldsymbol{b}+\boldsymbol{LY}$ is a transformed variable, and
$\boldsymbol{\Sigma}$ is the covariance matrix for $\boldsymbol{Y}$;
Then $\boldsymbol{\Sigma}_{(\boldsymbol{b}+\boldsymbol{LY})} = \boldsymbol{L\Sigma L}'$.

Another way to write $\boldsymbol{Y}_j = \boldsymbol{X}_j\boldsymbol{\Gamma} + \boldsymbol{Z}_j\boldsymbol{U}_j + \boldsymbol{R}_j$,

$$\boldsymbol{Y}_j = \boldsymbol{X}_j\boldsymbol{\Gamma} + \left(\begin{array}{c|c} \boldsymbol{Z}_j & \boldsymbol{I} \end{array}\right)\left(\begin{array}{c} \boldsymbol{U}_j \\ \hline \boldsymbol{R}_j \end{array}\right)$$

where $\boldsymbol{I}$ is an $(n_j \times n_j)$ identity matrix.

So for our linear mixed model,

$$\boldsymbol{\Sigma}_{y_j} = \left(\begin{array}{c|c} \boldsymbol{Z}_j & \boldsymbol{I} \end{array}\right)\left(\begin{array}{c|c} \boldsymbol{T} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma^2\boldsymbol{I} \end{array}\right)\left(\begin{array}{c} \boldsymbol{Z}_j' \\ \hline \boldsymbol{I} \end{array}\right) = \boldsymbol{Z}_j\boldsymbol{T}\boldsymbol{Z}_j' + \sigma^2\boldsymbol{I}$$

Model 3: SAS/MIXED Results

Dimensions

Covariance Parameters	4
Columns in X	3
Columns in Z Per Subject	2
Subjects	160
Max Obs Per Subject	67
Observations Used	7185
Observations Not Used	0
Total Observations	7185

Model 3: SAS/MIXED Results

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	2.6445	0.3963	6.67	$< .0001$
UN(2,1)	id	-0.2596	0.2396	-1.08	0.2787
UN(2,2)	id	0.6701	0.2766	2.42	0.0077
Residual		36.7133	0.6260	58.64	$< .0001$

Effect	Estimate	Standard Error	t Value	Pr $> \|t\|$
Intercept	12.6588	0.1482	85.41	$< .0001$
cses	2.1912	0.1276	17.17	$< .0001$
meanses	5.8959	0.3577	16.48	$< .0001$

Summary of Some of Models for HSB

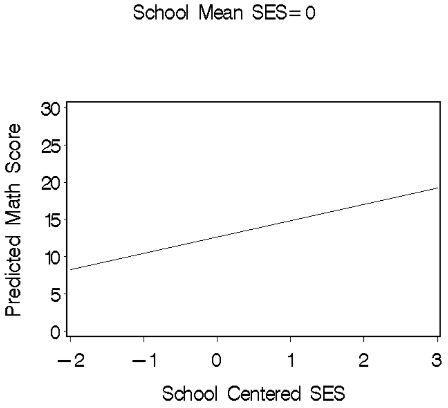
		Null Model	Random Intercept	Random Intercept and Slope					
		value	value	value	SE	value	SE	value	SE
Fixed Effects									
γ_{00}	intercept	12.64	12.65	12.65	.24	12.66	.15	12.66	.15
γ_{01}	cses		2.19	2.19	.13	2.20	.13	2.19	.13
γ_{10}	meanses					5.87	.36	5.90	.36
γ_{11}	cses*meanses					.29	.32		
Random Effects									
τ_0^2	intercept	8.55	8.61	8.61	1.07	2.64	.40	2.64	.40
τ_{10}				.05	.40	$-.26$	.24	$-.26$	.24
τ_1^2	cses			.68	.28	.65	.28	.67	.28
σ^2		39.15	37.01	36.70	.63	36.72	.63	36.71	.63

$$\widehat{\text{math}}_{ij} = 12.66 + 2.19(\text{SES}_{ij} - \overline{\text{SES}}_j) + 5.90(\overline{\text{SES}})_j$$

Model 3: For the Average School

when $\overline{\text{SES}}_j = 0$, so

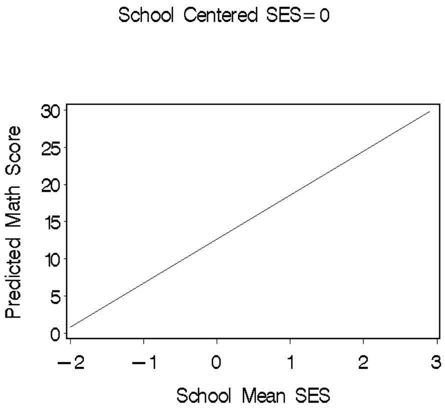
$$\widehat{\text{math}}_{ij} = 12.66 + 2.19(\text{SES}_{ij} - \overline{\text{SES}}_j)$$



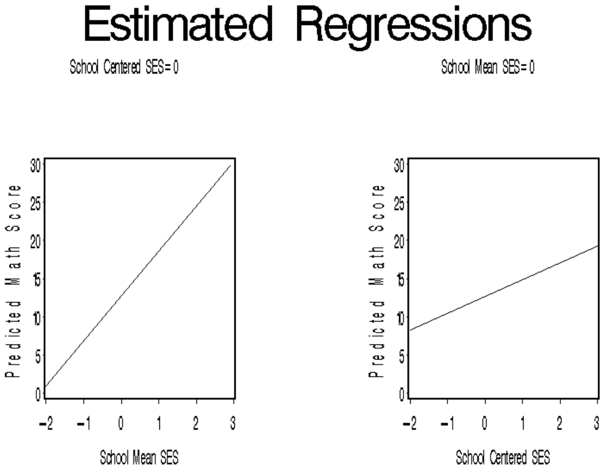
Model 3: For the Average Student

within a school when $(\text{SES}_{ij} - \overline{\text{SES}}_j) = 0$, so

$$\widehat{\text{math}}_{ij} = 12.66 + 5.90(\overline{\text{SES}})_j$$



Comparing Average Student & School



Lots of Micro Variables

The models can get very complex so here is where theory and substantive research questions are very important.

Level 1:

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^{p} \beta_{kj} x_{kij} + R_{ij}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ i.i.d.

Level 2:

$$\begin{aligned}
\beta_{0j} &= \gamma_{00} + \sum_{l=1}^{q} \gamma_{0l} z_{lj} + U_{0j} \\
\beta_{1j} &= \gamma_{10} + \sum_{l=1}^{q} \gamma_{1l} z_{lj} + U_{1j} \\
\vdots \quad & \quad \vdots \\
\beta_{pj} &= \gamma_{p0} + \sum_{l=1}^{q} \gamma_{pl} z_{lj} + U_{pj}
\end{aligned}$$

where $\boldsymbol{U}_j \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{T})$ i.i.d.

Lots of Micro & Macro Variables

Linear Mixed Model:

$$
\begin{aligned}
Y_{ij} &= \gamma_{00} + \sum_{k=1}^{p} \gamma_{k0} x_{kij} + \sum_{l=1}^{q} \gamma_{0l} z_{lj} + \sum_{k=1}^{p} \sum_{l=1}^{q} \gamma_{kl} x_{kij} z_{lj} \\
& \quad + U_{0j} + \sum_{k=1}^{p} U_{kj} x_{kij} + R_{ij}
\end{aligned}
$$

where

$$
\begin{pmatrix} U_{0j} \\ U_{1j} \\ \vdots \\ U_{pj} \\ R_{ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \tau_{10} & \ldots & \tau_{p0} & 0 \\ \tau_{10} & \tau_1^2 & \ldots & \tau_{p1} & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ \tau_{p0} & \tau p1 & \ldots & \tau_p^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 \end{pmatrix} \right) \qquad i.i.d.
$$

Lots of Micro & Macro Variables

Marginal Model

$$Y_{ij} \sim \mathcal{N}\left(\mu_{Y_{ij}}, \text{var}(Y_{ij})\right)$$

where

$$\mu_{Y_{ij}} = \gamma_{00} + \sum_{k=1}^{p} \gamma_{k0} x_{kij} + \sum_{l=1}^{q} \gamma_{0l} z_{lj} + \sum_{k=1}^{p} \sum_{l=1}^{q} \gamma_{kl} x_{kij} z_{lj}$$

and

$$\text{var}(Y_{ij}) = \tau_0^2 + 2\sum_{k=1}^{p} \tau_{k0} x_{kij} + 2\sum_{k \neq l} \tau_{kl} x_{kij} x_{lij} + \sum_{k=1}^{p} \tau_k^2 x_{kij}^2 + \sigma^2$$

Example: HSB (again)

with lots micro & macro-level explanatory variables.

Level 1

- SES (group mean centered)
- Gender ($=1$ for female, $=0$ for male)
- Minority status ($=1$ for minority, $=0$ for white)

Level 2

- Mean SES
- School enrollment (size)
- Sector ($=1$ private, $0=$ public)
- Proportion students in academic track
- Disciplinary climate
- Minority majority (himinty $=1$ if 40% minority, $0=$ if $< 40\%$ minority)

HSB: Model 1

Level 1:

$$\begin{aligned}(\text{math})_{ij} &= \beta_{0j} + \beta_{1j}(\text{cSES})_{ij} + \beta_{2j}(\text{minority})_{ij} \\ &\quad + \beta_{3j}(\text{female})_{ij} + R_{ij}\end{aligned}$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ and i.i.d.

Level 2: Everything into the intercept

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SES}})_j + \gamma_{02}(\text{size})_j + \gamma_{03}(\text{sector})_j \\ &\quad + \gamma_{04}(\text{pacad})_j + \gamma_{05}(\text{disclim})_j + \gamma_{06}(\text{himinty})_j \\ &\quad + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j} \\ \beta_{2j} &= \gamma_{20} + U_{2j} \\ \beta_{3j} &= \gamma_{30} + U_{3j}\end{aligned}$$

where $\boldsymbol{U}_j \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$.

HSB: “Kitchen Sink” for Intercept

and simple models for the slopes

Linear Mixed Model:

$$
\begin{aligned}
(\text{math})_{ij} \;=\; & \gamma_{00} + \gamma_{10}(\text{cSES})_{ij} + \gamma_{20}(\text{female})_{ij} \\
& +\gamma_{30}(\text{minority})_{ij} \\
& +\gamma_{01}(\overline{\text{SES}})_{j} + \gamma_{02}(\text{size})_{j} + \gamma_{03}(\text{sector})_{j} \\
& +\gamma_{04}(\text{pacad})_{j} + \gamma_{05}(\text{disclim})_{j} \\
& +\gamma_{06}(\text{himinty})_{j} \\
& +U_{0j} + U_{1j}(\text{cSES})_{ij} + U_{2j}(\text{minority})_{ij} \\
& +U_{3j}(\text{female})_{ij} + R_{ij}
\end{aligned}
$$

Model 1: SAS/MIXED (edited) Output

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr$>|t|$	
Intercept	11.4831	0.5008	153	22.93	$< .0001$	
cSES	1.8867	0.1203	159	15.69	$< .0001$	Micro
female	−1.2814	0.1737	122	−7.38	$< .0001$	
minority	−3.0752	0.2429	135	−12.66	$< .0001$	
meanses	3.1679	0.4236	6606	7.48	$< .0001$	
size	0.000694	0.000209	6606	3.31	.0009	Macro
sector	0.8131	0.3776	6606	2.15	.0313	
pracad	2.7217	0.7952	6606	3.42	.0006	
disclim	−0.4086	0.1753	6606	−2.33	.0198	
himinty	0.1856	0.3142	6606	0.59	.5548	

Model 1: SAS/MIXED (Edited) Output

Covariance Parameter Estimates

Cov Parm		Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	intercept	1.9323	0.4866	3.97	$<$ .0001
UN(2,1)		0.3281	0.2735	1.20	.2303
UN(2,2)	cSES	0.4029	0.2468	1.63	.0513
UN(3,1)		−0.8775	0.4183	−2.10	.0359
UN(3,2)		−0.08297	0.2623	−0.32	.7517
UN(3,3)	female	0.7350	0.5111	1.44	.0752
UN(4,1)		−0.2981	0.5291	−0.56	.5732
UN(4,2)		−0.6821	0.3271	−2.08	.0371
UN(4,3)		−0.01683	0.5077	−0.03	.9736
UN(4,4)	minority	1.6198	0.7871	2.06	.0198
Residual		35.3381	0.6111	57.83	$<$ .0001

Model 1: SAS/MIXED Output

The previous slide gave us $\hat{\boldsymbol{T}}$ (and $\hat{\sigma}^2$).

We can also get

$$\boldsymbol{Z}_j \hat{\boldsymbol{T}} \boldsymbol{Z}_j' \qquad \text{where} \quad j = 1224$$

		Estimated G Matrix:			
Effect	school	Col1	Col2	Col3	Col4
Intercept	1224	1.9539	0.3451	−0.8893	−0.3179
cSES	1224	0.3451	0.4037	−0.0795	−0.6716
female	1224	−0.8893	−0.0795	0.7468	−0.02782
minority	1224	−0.3179	−0.6716	−0.02782	1.6014

Model 1: Notes about R

The model fit above fails to converge in R ; that is,

```
model.5 ← lmer(mathach ∼ 1 + ses.centered + female +
minority + meanses + size + sector + pracad + disclim +
himinty +(1 + ses.centered + female + minority| id),
data=hsb, REML=FALSE)
```

There is a warning about differing scales. In this case, rescaling school size does the trick:

hsb\$xsize ← scale(hsb\$size,center=FALSE,scale=TRUE)

```
model.5 ← lmer(mathach ∼ 1 + ses.centered + female +
minority + meanses + xsize + sector + pracad + disclim +
himinty +(1 + ses.centered + female + minority| id),
data=hsb, REML=FALSE)
```

```
as.data.frame(VarCorr(model.5))[4]
```

Model 1: Notes about R

- After rescaling school size, the estimated fixed effects and std errors from R are essentially the same from SAS; however, the estimated variances of random effects are a bit different (not much). There are more noticeable differences are seen in the covariances.
- R and SAS don't always yield exactly the same results because of differences in estimation algorithms; however, they should be close.
- The next model needs more finessing to get it to converge. . . more rescaling and changing the optimizer.

HSB Model 2: With Cross-Level Interactions

- Drop percept minority from intercept model.
- Retain other macro-variables for intercept model.
- Add `sector` to model for a couple of slopes.
- Level 1:

$$\begin{aligned}(\text{math})_{ij} &= \beta_{0j} + \beta_{1j}(\text{cSES})_{ij} + \beta_{2j}(\text{minority})_{ij} \\ &\quad + \beta_{3j}(\text{female})_{ij} + R_{ij}\end{aligned}$$

  where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ and i.i.d.
- Level 2:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SES}})_j + \gamma_{02}(\text{size})_j + \gamma_{03}(\text{sector})_j \\ &\quad + \gamma_{04}(\text{pacad})_j + \gamma_{05}(\text{disclim})_j + U_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{sector})_j + U_{1j} \\ \beta_{2j} &= \gamma_{20} + \gamma_{21}(\text{sector})_j + U_{2j} \\ \beta_{3j} &= \gamma_{30}\end{aligned}$$

  where $\boldsymbol{U}_j \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$.

HSB Model 2: With Cross Level

Linear Mixed Model

$$
\begin{aligned}
(\text{math})_{ij} \quad = \quad & \gamma_{00} + \gamma_{10}(\text{cSES})_{ij} + \gamma_{20}(\text{female})_{ij} + \gamma_{30}(\text{minority})_{ij} \\
& + \gamma_{01}(\overline{\text{SES}})_j + \gamma_{02}(\text{size})_j + \gamma_{03}(\text{sector})_j + \gamma_{04}(\text{pacad})_j \\
& + \gamma_{05}(\text{disclim})_j \\
& + \gamma_{11}(\text{cSES})_{ij}(\text{sector})_j + \gamma_{21}(\text{minority})_{ij}(\text{sector})_j \\
& + U_{0j} + U_{1j}(\text{cSES})_{ij} + U_{2j}(\text{minority})_{ij} + R_{ij}
\end{aligned}
$$

where $\begin{pmatrix} \boldsymbol{U}_j \\ R_{ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{T} & \mathbf{0} \\ \mathbf{0} & \sigma^2 \end{pmatrix} \right). \qquad i.i.d.$

HSB Model 2: SAS/MIXED (edited) Output

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr $> \|t\|$	
Intercept	11.1964	0.4953	154	22.60	$< .0001$	
cSES	2.3174	0.1526	158	15.18	$< .0001$	Micro
female	-1.2433	0.1576	6728	-7.89	$< .0001$	
minority	-3.8520	0.3090	134	-12.47	$< .0001$	
meanses	2.8402	0.3925	6728	7.24	$< .0001$	Macro
size	0.0009	0.0002	6728	4.24	$< .0001$	
sector	0.4696	0.3902	6728	1.20	.2289	
pracad	3.2717	0.7913	6728	4.13	$< .0001$	
disclim	-0.3555	0.1801	6728	-1.97	.0484	
cSES*sector	-1.0110	0.2277	6728	-4.44	$< .0001$	Cross-level
minority*sector	1.7885	0.4232	6728	4.23	$< .0001$	

HSB Model 2: SAS/MIXED Output

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm		Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	intercept	1.2191	0.2729	4.47	$< .0001$
UN(2,1)		0.1401	0.1934	0.72	.4687
UN(2,2)	cSES	0.1634	0.2133	0.77	.2218
UN(3,1)		−0.0606	0.3554	−0.17	.8646
UN(3,2)		−0.1427	0.2889	−0.49	.6214
UN(3,3)	minority	0.7940	0.6684	1.19	.1174
Residual		35.4594	0.6081	58.31	$< .0001$

HSB Model 3: Refining Model 2

- Same Level 1 model.
- Note that τ_1^2 (from slope of cSES) and τ_2^2 (from slope of minority) are small relative to their standard errors —

  Later we'll talk about valid statistical tests of random part of the model.
- The cross-level interactions between sector and cSES and between sector and minority appear to be important.
- Let's make slopes for cSES and minority fixed, but still include explanatory variables for them.

HSB Model 3: Refining Model 2

The change in the Level 2 model:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{\text{SES}})_j + \gamma_{02}(\text{size})_j + \gamma_{03}(\text{sector})_j \\ &\quad + \gamma_{04}(\text{pacad})_j + \gamma_{05}(\text{disclim})_j + U_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{sector})_j \\ \beta_{2j} &= \gamma_{20} + \gamma_{21}(\text{sector})_j \\ \beta_{3j} &= \gamma_{30} \end{aligned}$$

where

$$U_{0j} \sim \mathcal{N}(\mathbf{0}, \tau_{00})$$

HSB Model 3: SAS/MIXED Output

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	1.3000	0.2423	5.37	$< .0001$
Residual	35.6316	0.6014	59.25	$< .0001$

HSB Model 3: SAS/MIXED Output

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr $>$ $\|t\|$	
Intercept	11.2320	0.4968	154	22.61	$< .0001$	
cses	2.3167	0.1457	7020	15.90	$< .0001$	Micro
female	-1.2371	0.1575	7020	-7.85	$< .0001$	
minority	-3.8231	0.2863	7020	-13.35	$< .0001$	
meanses	2.8394	0.3856	7020	7.36	$< .0001$	Macro
size	0.0009	0.0002	7020	4.32	$< .0001$	
sector	0.4775	0.3922	7020	1.22	0.2235	
pracad	3.1853	0.7880	7020	4.04	$< .0001$	
disclim	-0.3749	0.1793	7020	-2.09	0.0366	
cses*sector	-1.0075	0.2175	7020	-4.63	$< .0001$	Cross
minority*sector	1.7673	0.3862	7020	4.58	$< .0001$	-level

Centering Variables

General comments:

- The location or centering of explanatory variables determines the meaning of the intercept.
- In standard linear regression, centering variables leads to the same model (i.e., same fit, same slope parameters, etc), but what about HLM's?
- We'll investigate various forms of centering under different types of centering (variables in model) and for both Random Intercept Models and Random Intercept and Slopes Models.

Types of Centering

- None or "Raw Score" (RS): x_{ij}
- Grand Mean Centered (GMC): $(x_{ij} - \bar{x})$
- Group Mean Centered ($GpMC$): $(x_{ij} - \bar{x}_j)$

- Raw Score + Group Mean ($RS + GpM$):

$$x_{ij} \quad \text{and} \quad \bar{x}_j$$

- Group Mean Centered + Group Mean ($GpMC + GPM$):

$$(x_{ij} - \bar{x}_j) \quad \text{and} \quad \bar{x}_j$$

Comparisons in terms of . . .

- Statistical Equivalence: Whether models fit data equally well (i.e., same fit statistics, same predictions).
- Parameter Equivalence: Whether estimated parameters are the same or different.
  - fixed effects (i.e., γ's)
  - Variances and covariances (i.e., τ's & σ^2).
  - Estimated covariance matrix for parameter estimates.
  - Estimated correlation matrices for $\hat{\boldsymbol{T}}$ & $\hat{\boldsymbol{\Gamma}}$.

Comparisons in terms of . . .

- Parameter Stability: On replications of study and/or inclusion of other variables whether the estimated parameters are likely to be closer to those of the original.

  Problem of "Bouncing Beta's"

- Interpretation or the meaning of the intercept and coefficients (i.e., the slopes).

The Plan

We'll study random intercept models and then random intercept and slopes models. For each we'll

- Empirically compare models using the NELS88 ($N = 23$) data with Y_{ij} = math and x_{ij} = time spent doing homework.
- Determine why we get what we get.
- Look at other implications.
- Conclude with recommendations.

Random Intercept Model & Centering

- Two models are statistically equivalent, if they have the same "-2loglike."
- For empirical comparison, we'll use the NELS88 $(N = 23)$ with $x_{ij} =$ time spent doing homework.
- Empirical results. . .

Empirical Results: Statistically Equal

Type of Centering	Variable(s) in model	Random Intercept
None or "Raw Score"	x_{ij}	3730.5
Grand Mean Centered	$(x_{ij} - \bar{x})$	3730.5
Group mean centered	$(x_{ij} - \bar{x}_j)$	3734.4
Raw Score + Group Mean	x_{ij} & $\bar{x}_j$	3729.6
Group Mean Centered + group mean	$(x_{ij} - \bar{x}_j)$ & $\bar{x}_j$	3729.6

RS & GMC: Statistical Equivalence

Raw Score (RS)

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij}$
where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ i.i.d.

Level 2: $\beta_{0j} = \gamma_{00} + U_{0j}$ where $U_{0j} \sim \mathcal{N}(0, \tau_0^2)$ i.i.d.
$\beta_{1j} = \gamma_{10}$

Linear mixed model:

$$Y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + U_{0j} + R_{ij}$$

RS & GMC: Statistical Equivalence

Grand Mean Centered (GMC)

Level 1: $Y_{ij} = \acute{\beta}_{0j} + \acute{\beta}_{1j}(x_{ij} - \bar{x}) + R_{ij}$

Level 2: $\acute{\beta}_{0j} = \acute{\gamma}_{00} + U_{0j}$

$\acute{\beta}_{1j} = \acute{\gamma}_{10}$

Linear mixed model:

$$
\begin{aligned}
Y_{ij} &= \acute{\gamma}_{00} + \acute{\gamma}_{10}(x_{ij} - \bar{x}) + U_{0j} + R_{ij} \\
& \quad \underbrace{(\acute{\gamma}_{00} - \acute{\gamma}_{10}\bar{x})} + \underbrace{\acute{\gamma}_{10}} x_{ij} + U_{0j} + R_{ij} \\
&= \qquad \gamma_{00} \qquad + \gamma_{10} x_{ij} + U_{0j} + R_{ij}
\end{aligned}
$$

Same as RS

RS & GMC: Statistical Equivalence

$$\begin{aligned} Y_{ij} &= \acute{\gamma}_{00} + \acute{\gamma}_{10}(x_{ij} - \bar{x}) + U_{0j} + R_{ij} \\ &= \underbrace{(\acute{\gamma}_{00} - \acute{\gamma}_{10}\bar{x})} + \underbrace{\acute{\gamma}_{10}} x_{ij} + U_{0j} + R_{ij} \\ &= \qquad \gamma_{00} \qquad + \gamma_{10} x_{ij} + U_{0j} + R_{ij} \end{aligned}$$

That is,

$$\begin{array}{ccc} RS & & GMC \\ \gamma_{00} & = & (\acute{\gamma}_{00} - \acute{\gamma}_{10}\bar{x}) \\ \gamma_{10} & = & \acute{\gamma}_{10} \end{array}$$

U_{0j}, R_{ij}, $\hat{\tau}_0^2$ and $\hat{\sigma}^2$ are also the same.

RS & GMC: Parameter Equivalence

		RS: x_{ij}		GMC: $(x_{ij} - \bar{x})$	
Parameter		value	(SE)	value	(SE)
Fixed effects					
intercept	γ_{00}	46.3494	(1.14)	51.0840	(1.02)
homework	γ_{10}	2.4020	(.28)	2.4020	(.28)
Random effects					
intercept	τ_0^2	20.2251	(7.07)	20.2251	(7.07)
residual	σ^2	71.1440	(4.52)	71.1440	(4.52)

RS & GMC: Parameter Equivalence

- The equivalency

$$\begin{aligned} \hat{\gamma}_{00} &= (\hat{\acute{\gamma}}_{00} - \hat{\acute{\gamma}}_{10}\bar{x}) \\ 46.3493 &= 51.0840 - 2.4020(1.9710983) \\ 46.3493 &= 46.3493 \end{aligned}$$

- The SE for the intercept is larger in the raw score model than it is in the grand mean centered; i.e.,

$$SE(\gamma_{00}) = 1.14 \qquad \text{versus} \qquad SE(\acute{\gamma}_{00}) = 1.02$$

- The other SE's are equivalent.

RS & GMC: Parameter Equivalence

- The estimated covariance between the estimated intercept and slope parameters is larger in the raw score case; that is,

$$\widehat{\mathbf{\Sigma}}_{\mathbf{\Gamma}} = \begin{pmatrix} 1.3022 & -.1405 \\ -.1405 & .07660 \end{pmatrix} \quad \text{versus} \quad \widehat{\mathbf{\Sigma}}_{\acute{\mathbf{\Gamma}}} = \begin{pmatrix} 1.0459 & .01048 \\ .01048 & .07660 \end{pmatrix}$$

- And so is the estimated correlation

$$\widehat{\boldsymbol{R}}_{\mathbf{\Gamma}} = \begin{pmatrix} 1.0000 & -.4449 \\ -.4449 & 1.0000 \end{pmatrix} \quad \text{versus} \quad \widehat{\boldsymbol{R}}_{\acute{\mathbf{\Gamma}}} = \begin{pmatrix} 1.0000 & -.03702 \\ -.03702 & 1.0000 \end{pmatrix}$$

RS & GMC: Parameter Equivalence

A little result regarding linear transformations of random variables (vectors):

- The linear transformation,

$$\begin{aligned}\hat{\gamma}_{00} &= (1 \times \hat{\hat{\gamma}}_{00}) + (-1.9710983 \times \hat{\hat{\gamma}}_{10}) \\ &= (1, -1.9710983) \begin{pmatrix} \hat{\hat{\gamma}}_{00} \\ \hat{\hat{\gamma}}_{10} \end{pmatrix} \\ &= \boldsymbol{L}' \begin{pmatrix} \hat{\hat{\gamma}}_{00} \\ \hat{\hat{\gamma}}_{10} \end{pmatrix}\end{aligned}$$

RS & GMC: Parameter Equivalence

- The (estimated) variance for transformed (random) variable,

$$\widehat{\text{var}}(\hat{\gamma}_{00}) = \boldsymbol{L}'\widehat{\boldsymbol{\Sigma}}_{\acute{\boldsymbol{\Gamma}}}\boldsymbol{L}$$

- Applying this to our case:

$$\begin{aligned}\widehat{\text{var}}(\gamma_{00}) &= (1, -1.9710983)\begin{pmatrix} 1.0459 & .01048 \\ .01048 & .07660 \end{pmatrix}\begin{pmatrix} 1 \\ -1.9710983 \end{pmatrix} \\ &= 1.3022,\end{aligned}$$

  which is exactly what SAS gave us.
- Also, $SE(\gamma_{00}) = \sqrt{1.3022} = 1.14$

RS & GMC

- Parameter Stability is not an issue here since there is only one explanatory variable.
- Interpretation of the intercept:
  - Raw Score:
    γ_{00} equals the value of Y_{ij} when $x_{ij} = 0$.
  - Grand mean centering:
    $\acute{\gamma}_{00}$ equals the value of Y_{ij} when $x_{ij} = \bar{x}$.

What about the equivalence between $RS + GpM$ and $GpMC + GpM$?

$RS + GpM$ and $GpMC + GpM$

Raw Score plus group mean ($RS + GpM$):

Level 1: $Y_{ij} = \beta^*_{0j} + \beta^*_{1j} x_{ij} + R_{ij}$
where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ i.i.d.

Level 2: $\beta^*_{0j} = \gamma^*_{00} + \gamma^*_{01} \bar{x}_j + U_{0j}$
where $U_{0j} \sim \mathcal{N}(0, \tau_0^2)$ i.i.d.
$\beta^*_{1j} = \gamma^*_{10}$

Linear mixed model:

$$Y_{ij} = \gamma^*_{00} + \gamma^*_{10} x_{ij} + \gamma^*_{01} \bar{x}_j + U_{0j} + R_{ij}$$

$RS + GpM$ and $GpMC + GpM$

Group mean centered plus group mean $GpMC + GpM$

Level 1: $Y_{ij} = \tilde{\beta}_{0j} + \tilde{\beta}_{1j}(x_{ij} - \bar{x}_j) + R_{ij}$

Level 2:
$$\begin{aligned}\tilde{\beta}_{0j} &= \tilde{\gamma}_{00} + \tilde{\gamma}_{01}\bar{x}_j + U_{0j}\\ \tilde{\beta}_{1j} &= \tilde{\gamma}_{10}\end{aligned}$$

Linear mixed model:

$$\begin{aligned}Y_{ij} &= \tilde{\gamma}_{00} + \tilde{\gamma}_{10}(x_{ij} - \bar{x}_j) + \tilde{\gamma}_{01}\bar{x}_j + U_{0j} + R_{ij}\\ &= \underbrace{\tilde{\gamma}_{00}} + \underbrace{\tilde{\gamma}_{10}}\, x_{ij} + \underbrace{(\tilde{\gamma}_{01} - \tilde{\gamma}_{10})}\,\bar{x}_j + U_{0j} + R_{ij}\\ &= \gamma^*_{00} + \gamma^*_{10}x_{ij} + \gamma^*_{01}\bar{x}_j + U_{0j} + R_{ij}\end{aligned}$$

Parameter Equivalence

Linear mixed model for $GpMC + GpM$

$$\begin{aligned} Y_{ij} &= \tilde{\gamma}_{00} + \tilde{\gamma}_{10}(x_{ij} - \bar{x}_j) + \tilde{\gamma}_{01}\bar{x}_j + U_{0j} + R_{ij} \\ &= \underbrace{\tilde{\gamma}_{00}} + \underbrace{\tilde{\gamma}_{10}}\, x_{ij} + \underbrace{(\tilde{\gamma}_{01} - \tilde{\gamma}_{10})}\, \bar{x}_j + U_{0j} + R_{ij} \\ &= \gamma^*_{00} + \gamma^*_{10} x_{ij} + \qquad \gamma^*_{01}\bar{x}_j + U_{0j} + R_{ij} \end{aligned}$$

Same as linear mixed model for $RS + GpM$ where

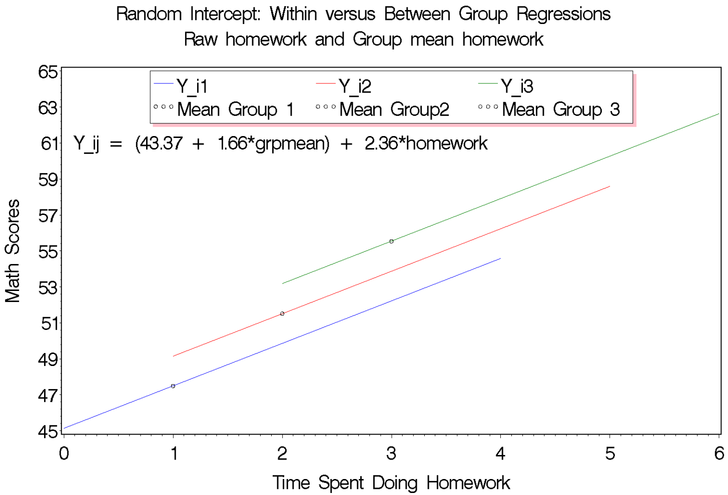
$$\begin{aligned} \gamma^*_{00} &= \tilde{\gamma}_{00} \\ \gamma^*_{10} &= \tilde{\gamma}_{10} \\ \gamma^*_{01} &= (\tilde{\gamma}_{01} - \tilde{\gamma}_{10}) \end{aligned} \qquad \text{or} \qquad \tilde{\gamma}_{01} = \gamma^*_{10} + \gamma^*_{01}$$

and U_{0j}, R_{ij}, τ_0^2, and σ^2 are the same.

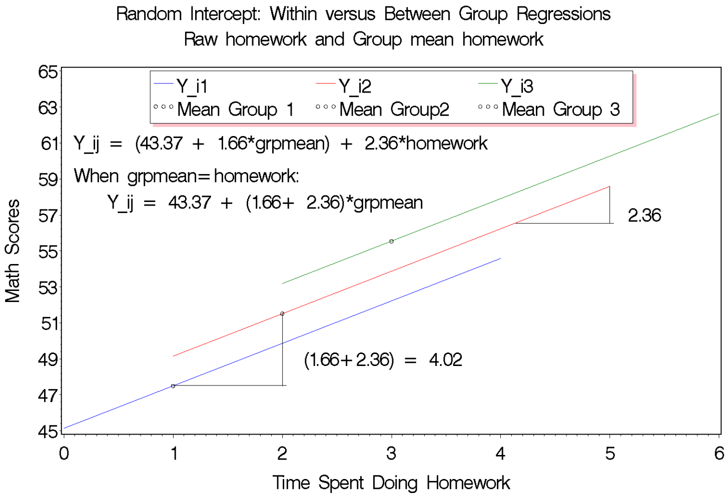
Parameter Equivalence

Parameter		$RS + GpM$: x_{ij} & $\bar{x}_j$ value	(SE)	$GpMC + GpM$: $(x_{ij} - \bar{x}_j)$ & $\bar{x}_j$ value	(SE)
Fixed effects					
intercept	γ_{00}	43.3712	(3.28)	43.3712	(3.28)
homework	γ_{10}	2.3610	(.28)	2.3610	(.28)
group mean	γ_{01}	1.6623	(1.72)	4.0233	(1.70)
Random effects					
intercept	τ_0^2	18.9094	(6.81)	18.9094	(6.81)
residual	σ^2	71.2016	(4.54)	71.2016	(4.54)

Interpretation of $RS + GpM$



Interpretation of $RS + GpM$



Parameter Equivalence

between $RS + GpM$ and $GpMC + GpM$

$$
\begin{aligned}
\hat{\gamma}^{*}_{01} &= \hat{\hat{\gamma}}_{01} - \hat{\hat{\gamma}}_{10} \\
1.6623 &= 4.0233 - 2.3610 \\
1.6623 &= 1.6623
\end{aligned}
$$

or

$$
\begin{aligned}
\hat{\hat{\gamma}}_{01} &= \hat{\gamma}^{*}_{10} + \hat{\gamma}^{*}_{01} \\
4.0233 &= 2.3610 + 1.6623 \\
4.0233 &= 4.0233
\end{aligned}
$$

Parameter Differences

between $RS + GpM$ and $GpMC + GpM$

- SE for the coefficient for the group mean is larger in the raw score model.
- Estimated covariance matrices for fixed

$$\hat{\mathbf{\Sigma}}_{\mathbf{\Gamma}^*} = \begin{matrix} \gamma^*_{00} \\ \gamma^*_{10} \\ \gamma^*_{01} \end{matrix} \begin{pmatrix} 10.7376 & .0000 & -.53139 \\ .0000 & .07863 & -.07863 \\ -.53139 & -.07863 & 2.9746 \end{pmatrix}$$

$$\hat{\mathbf{\Sigma}}_{\tilde{\mathbf{\Gamma}}} = \begin{matrix} \tilde{\gamma}_{00} \\ \tilde{\gamma}_{10} \\ \tilde{\gamma}_{01} \end{matrix} \begin{pmatrix} 10.7376 & .0000 & -.53139 \\ .0000 & .07863 & .0000 \\ -.53139 & .0000 & 2.8960 \end{pmatrix}$$

- The transformation leads to this difference.

Parameter Differences

between $RS + GpM$ and $GpMC + GpM$

- Estimated correlation matrices for fixed

$$\hat{\boldsymbol{R}}_{\boldsymbol{\Gamma}^*} = \begin{matrix} \gamma^*_{00} \\ \gamma^*_{10} \\ \gamma^*_{01} \end{matrix} \begin{pmatrix} 1.0000 & .0000 & -.9403 \\ .0000 & 1.0000 & -.1626 \\ -.9403 & -.1626 & 1.0000 \end{pmatrix}$$

versus

$$\hat{\boldsymbol{R}}_{\tilde{\boldsymbol{\Gamma}}} = \begin{matrix} \tilde{\gamma}_{00} \\ \tilde{\gamma}_{10} \\ \tilde{\gamma}_{01} \end{matrix} \begin{pmatrix} 1.000 & .0000 & -.9529 \\ .0000 & 1.0000 & .0000 \\ -.9529 & .0000 & 1.0000 \end{pmatrix}$$

Parameter Stability

for $RS + GpM$ and $GpMC + GpM$

- The correlation matrix for the explanatory variables: x_{ij}, $(x_{ij} - \bar{x}_j)$, and $\bar{x}_j$:

$$\begin{array}{lr} \text{Raw Score} & x_{ij} \\ \text{Grp Mean Center} & (x_{ij} - \bar{x}_j) \\ \text{Grp Mean} & \bar{x}_j \end{array} \left(\begin{array}{rrr} 1.00 & .89 & .45 \\ .89 & 1.00 & .00 \\ .45 & .00 & 1.00 \end{array} \right)$$

- Which model is better in terms of parameter stability?

Interpretation

for $RS + GpM$ and $GpMC + GpM$

- Raw Score + Group Mean Centered ($RS + GpM$):

$$Y_{ij} = \gamma^*_{00} + \gamma^*_{10} x_{ij} + \gamma^*_{01} \bar{x}_j + U_{0j} + R_{ij}$$

- Group Mean Centered + Group Mean ($GpMC + GpM$):

$$\begin{aligned} Y_{ij} &= \tilde{\gamma}_{00} + \tilde{\gamma}_{10}(x_{ij} - \bar{x}_j) + \tilde{\gamma}_{01}\bar{x}_j + U_{0j} + R_{ij} \\ &= \underbrace{\tilde{\gamma}_{00}} + \underbrace{\tilde{\gamma}_{10}} x_{ij} + \underbrace{(\tilde{\gamma}_{01} - \tilde{\gamma}_{10})} \bar{x}_j + U_{0j} + R_{ij} \\ &= \gamma^*_{00} + \gamma^*_{10} x_{ij} + \gamma^*_{01} \bar{x}_j + U_{0j} + R_{ij} \end{aligned}$$

Summary of Models

- Raw score:

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + U_{0j} + R_{ij}$$

- Grand mean centered:

$$Y_{ij} = \acute{\gamma}_{00} + \acute{\gamma}_{10}(x_{ij} - \bar{x}) + U_{0j} + R_{ij}$$

- Group mean centered:

$$Y_{ij} = \breve{\gamma}_{00} + \breve{\gamma}_{10}(x_{ij} - \bar{x}_j) + U_{0j} + R_{ij}$$

- Raw score plus group mean:

$$Y_{ij} = \gamma^*_{00} + \gamma^*_{10} x_{ij} + \gamma^*_{01} \bar{x}_j + U_{0j} + R_{ij}$$

- Group mean centered plus group mean:

$$Y_{ij} = \tilde{\gamma}_{00} + \tilde{\gamma}_{10}(x_{ij} - \bar{x}_j) + \tilde{\gamma}_{01} \bar{x}_j + U_{0j} + R_{ij}$$

Random Intercept & Slopes

Type of Centering	Variable(s) in model	Random Intercept
None or Raw score	x_{ij}	3639.04
Grand mean centered	$(x_{ij} - \bar{x})$	3639.04
Group mean centered	$(x_{ij} - \bar{x}_j)$	3645.76
Raw score + group mean	x_{ij} & $\bar{x}_j$	3633.89
Group mean centered + group mean	$(x_{ij} - \bar{x}_j)$ & $\bar{x}_j$	3640.86

Random Intercept & Slopes

Statistical Equivalence:

- Only raw score (RS) and grand mean centered (GMC).
- Why only these two and why none of the others?
- We need to look at the models. . .

Raw Score (RS)

Level 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + R_{ij},$$

where $R_{ij} \sim \mathcal{N}(0, \sigma^2)$ i.i.d.

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + U_{0j} \\ \beta_{1j} &= \gamma_{10} + U_{1j}, \end{aligned}$$

where $\boldsymbol{U} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{T})$ i.i.d.

Linear mixed model:

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + U_{0j} + U_{1j} x_{ij} + R_{ij}$$

Grand Mean Centered (GMC)

Level 1:

$$Y_{ij} = \acute{\beta}_{0j} + \acute{\beta}_{1j}(x_{ij} - \bar{x}) + R_{ij}$$

Level 2:

$$\begin{aligned} \acute{\beta}_{0j} &= \acute{\gamma}_{00} + U_{0j} \\ \acute{\beta}_{1j} &= \acute{\gamma}_{10} + U_{1j} \end{aligned}$$

Linear mixed model:

$$\begin{aligned} Y_{ij} &= \acute{\gamma}_{00} + \acute{\gamma}_{10}(x_{ij} - \bar{x}) + \acute{U}_{0j} + \acute{U}_{1j}(x_{ij} - \bar{x}) + R_{ij} \\ &= \underbrace{(\acute{\gamma}_{00} - \acute{\gamma}_{10}\bar{x})} + \acute{\gamma}_{10}x_{ij} + \underbrace{(\acute{U}_{0j} - \acute{U}_{1j}\bar{x})} + \acute{U}_{1j}x_{ij} + R_{ij} \\ &= \quad \gamma_{00} \quad + \quad \gamma_{10}x_{ij} + \quad U_{0j} \quad + U_{1j}x_{ij} + R_{ij} \end{aligned}$$

Same as RS

Parameter Equivalence

		RS		GMC	
		x_{ij}		$(x_{ij} - \bar{x})$	
Parameter		value	(SE)	value	(SE)
Fixed effects					
intercept	γ_{00}	46.32	(1.72)	50.24	(1.04)
homework	γ_{10}	1.99	(.91)	1.99	(.91)
Random effects					
intercept	τ_0^2	59.24	(19.96)	21.41	(7.16)
	τ_{10}	−26.12	(9.85)	6.93	(4.96)
homework	τ_1^2	16.77	(5.83)	16.77	(5.83)
residual	σ^2	53.30	(3.47)	53.30	(3.47)

Parameter Equivalence

- Relationship between parameters

$$\begin{aligned} \hat{\gamma}_{00} &= \hat{\acute{\gamma}}_{00} - \hat{\acute{\gamma}}\bar{x} \\ 46.32 &= 50.24 - 1.99(1.9710983) \\ 46.32 &= 46.32 \end{aligned}$$

- We have statistical equivalence and know how parameters are related, but grand mean centering has an effect on the intercept variance τ_0^2 (& therefore the covariance, τ_{10}).

- Intuition regarding this

$$U_{0j} = \acute{U}_{0j} - \acute{U}_{1j}\bar{x}$$

More explicitly,

$$\text{var}(U_{0j}) = \text{var}(\acute{U}_{0j}) + \text{var}(\acute{U}_{1j})\bar{x}^2 - 2\text{cov}(\acute{U}_{0j}, \acute{U}_{1j})\bar{x}$$

Comparing Intercept & Slope Models

- Raw Score:
  $Y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + +U_{0j} + U_{1j}x_{ij} + R_{ij}$
- Grand Mean Centering:
  $Y_{ij} = \acute{\gamma}_{00} + \acute{\gamma}_{10}(x_{ij} - \bar{x}) + \acute{U}_{0j} + \acute{U}_{1j}(x_{ij} - \bar{x}) + R_{ij}$
- Group Mean Centering:
  $Y_{ij} = \breve{\gamma}_{00} + \breve{\gamma}_{10}(x_{ij} - \bar{x}_j) + \breve{U}_{0j} + \breve{U}_{1j}(x_{ij} - \bar{x}_j) + \breve{R}_{ij}$
- Raw Score plus group mean:
  $Y_{ij} = \gamma^*_{00} + \gamma^*_{10}x_{ij} + \gamma^*_{01}\bar{x}_j + U^*_{0j} + U^*_{1j}x_{ij} + R^*_{ij}$
- Group Mean Centering plus group mean:
  $Y_{ij} = \tilde{\gamma}_{00} + \tilde{\gamma}_{10}x_{ij} + (\tilde{\gamma}_{01} + \tilde{\gamma}_{10})\bar{x}_j + \tilde{U}_{0j} + \tilde{U}_{1j}x_{ij} - \tilde{U}_{1j}\bar{x}_j + \tilde{R}_{ij}$

Empirical Comparison

Estimated Parameters

Parameter		RS	GMC	$GpMC$	$RS + GpM$	
Fixed Effects						
intercept	γ_{00}	46.323	50.239	50.772	39.272	
x_{ij}	γ_{10}	1.987	1.987	1.968	1.892	
$\overline{x}_j$	γ_{01}	—	—	—	3.922	
Random Effects						
intercept	τ_0^2	59.244	21.405	25.921	67.58	
	τ_{10}	-26.1242	6.927	1.047	-30.936	-2.878 ° h
τ_1^2	16.768	16.768	16.411	17.760	16.621	
residual	σ^2	53.300	53.300	53.408	53.450	
-2loglike		3639.037	3639.037	3645.77	3633.88	

Empirical Comparison

Estimated Standard Errors

Parameter		RS	GMC	$GpMC$	$RS + GpM$	$GpMC$ $+GpM$
Fixed Effects						
intercept	γ_{00}	1.719	1.044	1.119	3.248	3.273
homework	γ_{10}	.906	.906	.899	.930	.904
$\overline{\text{homework}}_j$	γ_{01}	n.a.	n.a.	n.a.	1.509	1.701
Random Effects						
intercept	τ_0^2	19.960	7.164	8.392	23.32	6.916
	τ_{10}	9.847	4.964	4.819	11.275	4.814
homework	τ_1^2	5.830	5.830	5.758	6.142	5.837
residual	σ^2	3.467	3.467	3.480	3.459	3.485
-2loglike		3639.037	3639.037	3645.76	3633.886	3640.855

Centering & Random Slopes

- The differences here may not look large, but with more variables and different relationships in the data can get vastly different results.
- e.g., Kreft & de Leeuw give an example with more variables where the sign of some of the γ's changed.
- Also noteworthy are the estimated correlation matrices for the estimated parameters (the γ's) for the models...

Centering & Random Slopes

Raw Score plus Group Mean $(RS + GpM)$

$$\hat{\boldsymbol{R}}_{\hat{\boldsymbol{\Gamma}}^*} = \begin{array}{cc} \text{intercept} & \gamma_{00} \\ x_{ij} & \gamma_{10} \\ \bar{x}_j & \gamma_{01} \end{array} \begin{pmatrix} 1.0000 & -.4450 & -.8280 \\ -.4450 & 1.0000 & -.0584 \\ -.8280 & -.0584 & 1.0000 \end{pmatrix}$$

and Group Mean Centering plus Group Mean,

$$\hat{\boldsymbol{R}}_{\hat{\hat{\boldsymbol{\Gamma}}}} = \begin{array}{cc} \text{intercept} & \gamma_{00} \\ (x_{ij} - \bar{x}_j) & \gamma_{10} \\ \bar{x}_j & \gamma_{01} \end{array} \begin{pmatrix} 1.0000 & -.0355 & -.9521 \\ -.0355 & 1.0000 & -.0073 \\ -.9521 & -.0073 & 1.0000 \end{pmatrix}$$

Implications for parameter stability.

In Sum, With Random Slopes

The location of x_{ij} effects

- Meaning of the intercept.
- Intercept variance.
- Covariance between random intercept and random slope.

Recommendation 1

Snijders & Bosker (p 88):

> *Generally one should be reluctant to use group-mean centered random slopes models unless there is clear theory (or an empirical clue) that not in the first place the absolute score* X_{ij} *but rather the relative score* $(X_{ij} - \bar{X}_j)$ *is related to* Y_{ij}

Recommendation 2

Kreft & de Leeuw (p 113–114):

This question cannot be answered on the basis of technical considerations alone.... The choice should be made based on the researcher's knowledge of the data and the goals of the analysis.... If one decides to center, the only advice we can give ... is to add the subtracted mean to the model. If this is not done... between effect is not corrected for the mean effect of the centered first-level explanatory variables. This may be exactly what one wants, as in growth curve models where the explanatory variable is time.

Recommendation 3

Bryk & Raudenbush (page 28)

> *Again, the general caveat — be conscious of the choice of location for each Level-1 predictor because it has implications for interpretation of β_{0j}, var(β_{0j}), and by implication, all of the covariances involving β_{0j}.*

To Center or Not

Depends on

- Goals of the analysis, e.g.,
- Fitting as much variance as possible.
- Substantive theory, e.g.,

  Segerstrom, S.C., & Sephton, S.E. (2010). Optimistic Expectancies and Cell-Mediated Immunity : The Role of Positive Affect. *Psychological Science, 2*, 448–455. DOI: 10.1177/0956797610362061

- More specific recommendations. . .

When to Use Raw Scores

or Grand mean centering

- If you are more interested in the effects in individuals' performance than in group effects.

  e.g., Clustering is a nuisance factor

- If group mean centering is a different model that does not meet requirements of analysis.

When to use Group Mean Centering

- Theory says that individual and group effects are separate. i.e., Two separate models to be tested: one for individual-level and one for group-level.
- Group Mean Centering for technical reasons
  - Smaller correlations between random intercept and random slope.
  - Smaller correlations between Level 1 and Level 2 variables and cross-level interactions $\Longrightarrow$ stabilizes model, (coefficients are more or less independent estimates).

Summary: Topics/Concepts Covered

- Within group dependency accounted for by random effects.
- Between group differences accounted for by random effects.
- Central role of variance between and within groups.
- Regression models for intercept.
- Effect of adding Level 1 and Level 2 predictors.
- ICC (for random intercept models)
- Regression models for intercept and slope.
- Cross-level interactions occur when models for slopes include Level 2 measures.
- HLMs with random slopes (effects) predict heteroscedasticity.
- Graphing data in various ways can provide insight into possible effects to consider in developing models for data. . . . much more on this later.

Reading SAS/MIXED Documentation

Correspondence between SAS/MIXED manual & class notation

SAS Manual	Class Notation
$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$	$\boldsymbol{Y} = \boldsymbol{X\Gamma} + \boldsymbol{R}$
$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\epsilon}$	$\boldsymbol{Y} = \boldsymbol{X\Gamma} + \boldsymbol{ZU} + \boldsymbol{R}$
$\gamma \sim \mathcal{N}(\mathbf{0}, \boldsymbol{G})$	$\boldsymbol{U} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$
$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R})$	$\boldsymbol{R} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$
	where $\boldsymbol{\Sigma} = \sigma^2 I$
Covariance matrix for $\boldsymbol{y}$ or $\boldsymbol{Y}$	
$\boldsymbol{V} = \boldsymbol{ZGZ}' + \boldsymbol{R}$	$\boldsymbol{V} = \boldsymbol{ZTZ}' + \boldsymbol{\Sigma}$

SAS/MIXED: Random Slopes

```
proc mixed data=hsbcent noclprint covtest method=ML;
title 'HSB: SAS code for random int and slope';
class id;
model mathach = cSES meanSES female /solution covb corrb;
random intercept cses / subject=id type=un;
```

SAS/MIXED: Random Intercept and Slopes

- Write down the linear mixed model.
- The fixed part $\rightarrow$ `MODEL` statement.

  The `COVB` and the `CORRB` options (i.e., these are not necessary)
- The random part $\rightarrow$ `RANDOM` statement.
- Need to specify `TYPE` of covariance matrix as an option to the `RANDOM` statement.

  `TYPE=un` where "un" for unstructured.

R Random Intercept and Slopes

- Write down the linear mixed model.
- The fixed part goes after the "$\sim$" and before the random specification.
- For example,
  model.slopes y $\sim$ 1 + gender + cses + meanses + (1 + cses | id),
  data=hsb
- Output is a bit different from SAS. In R, you get

$$\sqrt{\hat{\tau}_j^2} \quad \text{and} \quad r(\hat{\tau}_j^2, \hat{\tau}_k^2);$$

whereas, in SAS you get

$$se(\hat{\tau}_k \quad \text{and} \quad \hat{\tau}_j^2$$

SAS R Examples

- Go through examples of
    - Random slopes.
    - Cross-level interactions.
- Note: CPU time and complex and/or poorly specified models.
- Computer Lab Session 2: Continue with the USA TIMSS data set but add some random slopes and a bit of graphics.