# Precursors of the multilevel random coefficient model

Lecture 18

Multilevel Modeling

Psychology 613 – Spring 2022

# Example data

Effect of homework on academic achievement

260 students nested within 10 schools

Subset of the National Education Longitudinal Study, 1988 (“NELS88”); now on Canvas

School is 1st var, homework is 5th, MathAch is 9th, School type is 6th

# Regression in one school

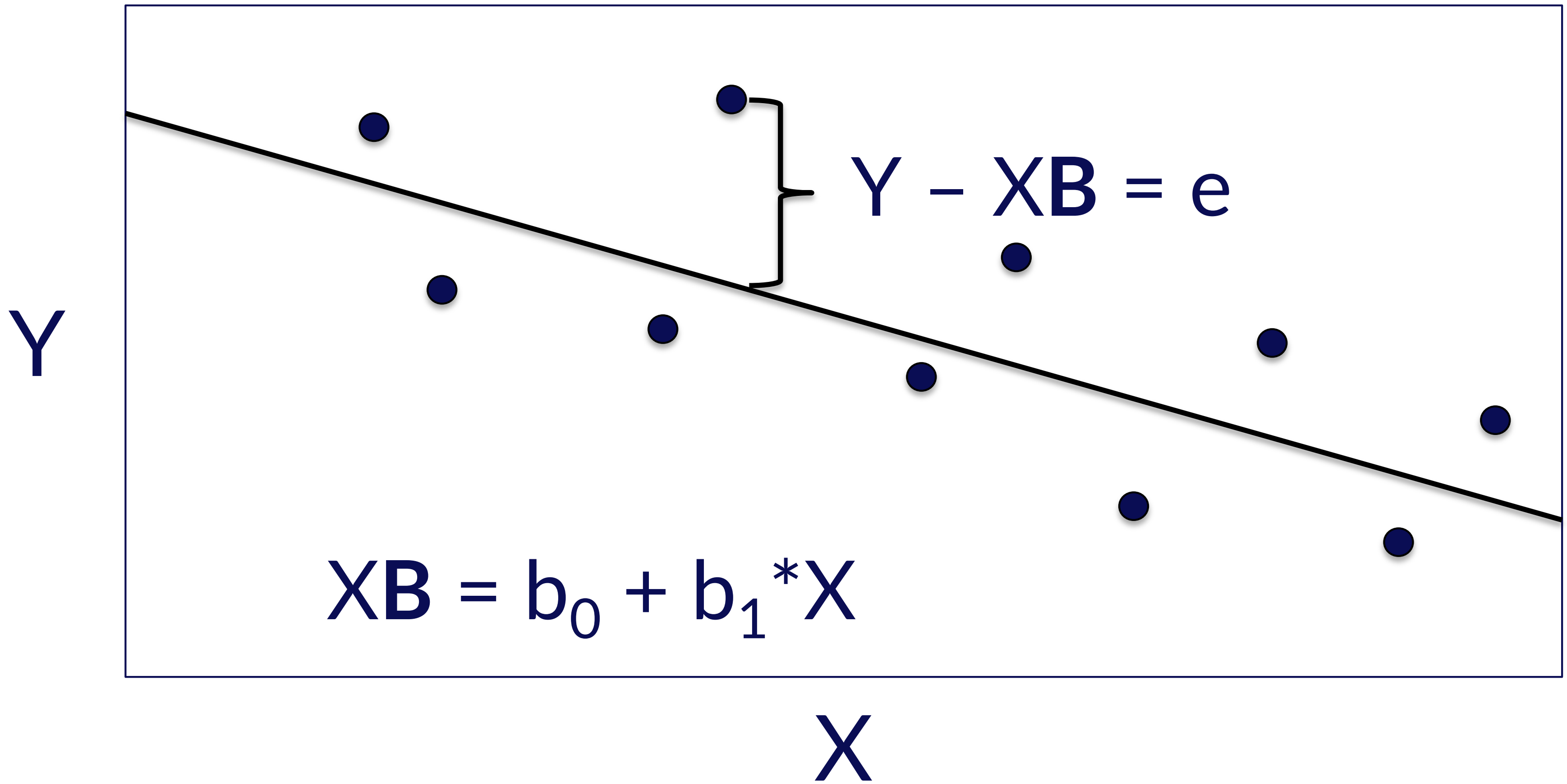
Run simple OLS regression with one school

With a single predictor X, the equation is:

$Y_i =$ $b_0 + b_1 X_i + e_i$ AKA

$MathAch_i =$ $b_0 + b_1 Homework_i + e_i$

# (Quick regression refresher)




OLS estimates of $b_0$ and $b_1$ define the best-fitting linear relationship between X and Y (in that they minimize the sum of squared errors).

# OLS predictor in a single-group

OLS regression equation with a single predictor:

$$Y_i = b_0 + b_1X_i + e_i$$

Where:

$b_1$ = cov(X,Y) / var(X) = *ΔY for 1-unit Δ in X*

$b_0$ = value of Y when X=0

Assumes the errors to be independent and identically distributed ~ $N(0, \sigma^2)$

# Single-group regression example

First school, OLS regression:

$$Math_i = 44.07^{***} + 3.57^{**}\ Homework_i + e_i$$

$b_0$ = 44.07; expected value of Math for a Homework value of 0

$b_1$ = 3.57; expected change in Math for a one unit increase in Homework

Note: With additional predictors, bs are interpreted as partial regression coefficients

# Using ANCOVA to incorporate information about multiple groups

Traditional way to think about ANCOVA:

- Grouping variable = Categorical treatment
- Continuous variable = “Covariate”
- *Interested in effect of treatment after removing the effect of (controlling for) the covariate*
- (Removes variance due to the covariate for a more powerful test of the treatment effect)

But here, we are interested in the *effect of the continuous variable, controlling for the grouping*

# (ANCOVA refresher)

Create dummy variables to indicate group membership:

| | D1 | D2 | D3 | … | D9 |
|---|---|---|---|---|---|
| Group 1 | 1 | 0 | 0 | 0 | 0 |
| Group 2 | 0 | 1 | 0 | 0 | 0 |
| … | | | | | |
| Group 10 | 0 | 0 | 0 | 0 | 0 |

$$Y_i = b_0 + b_1 D_1 + b_2 D_2 + \ldots + b_9 D_9 + b_{10} X_i + e_i$$

(Math) (Homework)

Dummy code for group 1

Dummy code for group 2

| D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# ANCOVA in R

```
> dataset <- read.spss("filename.sav",
  as.data.frame = TRUE)
> model1 <- lm(mathscore ~
  D1+D2+D3+D4+D5+D6+D7+D8+D9+timeonmath,
  data = dataset)
> summary(model1)
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  44.4314     1.8990  23.398  < 2e-16 ***
D1           -1.6650     2.4585  -0.677  0.49888    
D2           -7.3025     2.5576  -2.855  0.00466 ** 
D3            4.9015     2.4349   2.013  0.04519 *  
D4           -4.3822     2.4830  -1.765  0.07881 .  
D5            3.5870     2.4990   1.435  0.15244    
D6           -0.4885     2.5473  -0.192  0.84807    
D7           11.3418     2.1489   5.278 2.84e-07 ***
D8            0.7585     2.5182   0.301  0.76350    
D9           -0.9469     2.5131  -0.377  0.70665    
timeonmath    2.1366     0.3836   5.570 6.60e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# ANCOVA example

$Math_i$ = 44.43*** - 1.66$D_1$ – 7.30$D_2$** + 4.90$D_3$* + … - 0.95$D_9$ + 2.13***$HW_i$ + $e_i$

$R^2$ change test to determine whether the nine dummy codes containing the grouping information contribute to model fit (compared to $Math_i = b_0 + b_1 HW_i + e_i$):

$F(9, 249) = 13.93$, $p < .0001$

# Model comparison in R

```
model0 = lm(mathscore ~ timeonmath)
model1 = lm(mathscore ~ D1+D2+D3+D4+D5+D6+D7+D8+D9+timeonmath)

summary(model0)
summary(model1)

anova(model0, model1)
```

```
> anova(model0, model1)
Analysis of Variance Table

Model 1: mathscore ~ timeonmath
Model 2: mathscore ~ D1 + D2 + D3 + D4 + D5 + D6 + D7 + D8 + D9 + timeonmath
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1    258 24183
2    249 16082  9    8100.4 13.935 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

# Interpreting the ANCOVA estimates

$Math_i$ = 44.43*** - 1.66$D_1$ – 7.30$D_2$** + 4.90$D_3$* + … - 0.95$D_9$ + 2.13***$HW_i$ + $e_i$

$b_0$: Expected Math score for HW=0 in school #10

$b_1$-$b_9$: Difference in expected Math score between indicated school and school #10

→ Unique intercept for each group

e.g., school #1: $b_0 + b_1 = 44.43 – 1.66 = 42.77$

→ $b_{10}$: Common homework-achievement slope across all groups

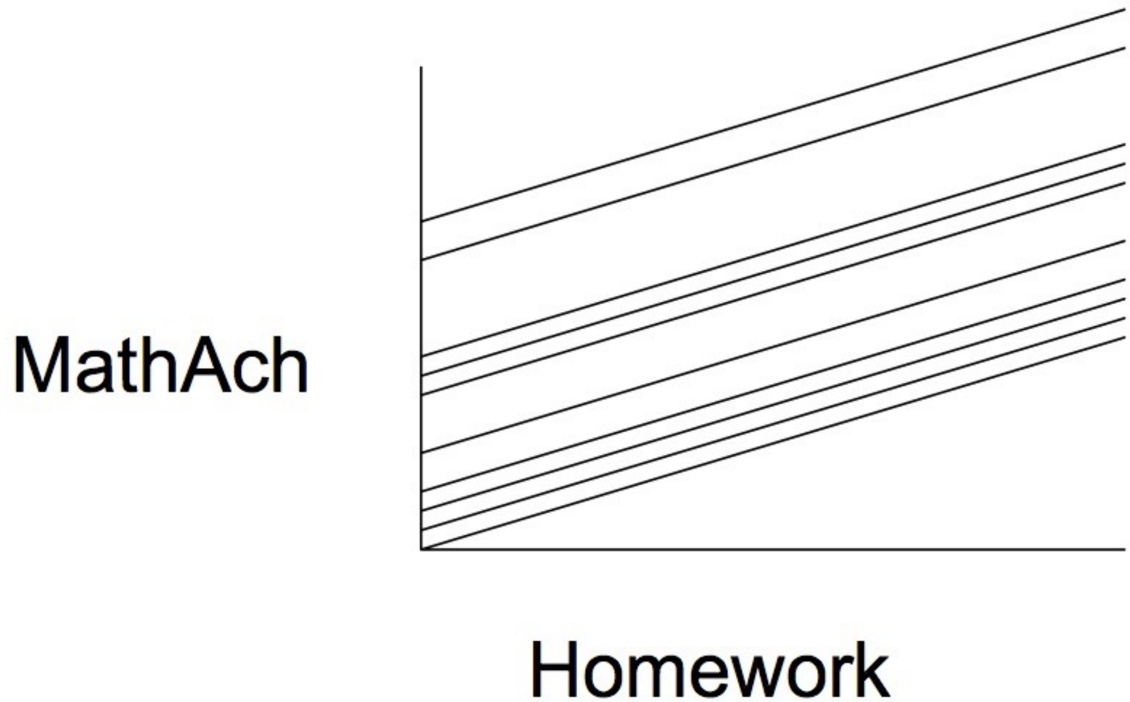
# Interpreting the ANCOVA estimates

**Coefficients**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 1 | (Constant) | 44.074 | .989 | | 44.580 | .000 |
| | Time spent on math homework | 3.572 | .388 | .497 | 9.200 | .000 |
| 2 | (Constant) | 44.431 | 1.899 | | 23.398 | .000 |
| | Time spent on math homework | 2.137 | .384 | .297 | 5.570 | .000 |
| | Dummy: Group 1 | -1.665 | 2.458 | -.043 | -.677 | .499 |
| | Dummy: Group 2 | -7.302 | 2.558 | -.175 | -2.855 | .005 |
| | Dummy: Group 3 | 4.901 | 2.435 | .128 | 2.013 | .045 |
| | Dummy: Group 4 | -4.382 | 2.483 | -.110 | -1.765 | .079 |
| | Dummy: Group 5 | 3.587 | 2.499 | .090 | 1.435 | .152 |
| | Dummy: Group 6 | -.489 | 2.547 | -.012 | -.192 | .848 |
| | Dummy: Group 7 | 11.342 | 2.149 | .446 | 5.278 | .000 |
| | Dummy: Group 8 | .759 | 2.518 | .019 | .301 | .763 |
| | Dummy: Group 9 | -.947 | 2.513 | -.023 | -.377 | .707 |

$Math_{ij} = 44.431 + [DUMMY] + \ldots +$ **2.137*homework$_{ij}$**

# Implied relationships in ANCOVA




- → ANCOVA allows different intercepts, but assumes a common covariate effect (i.e., slope across groups)
- → Expected value of MathAch for Homework=0 differs across groups

# Testing the common slope assumption

Create cross-products of dummy codes and covariate, e.g., $D_1$ * Homework

Estimate: $Math_i = b_0 + b_1D_1 + b_2D_2 + \ldots + b_9D_9 + b_{10}Homework_i + b_{11}D_1{*}HW_i + b_{12}D_2{*}HW_i + \ldots + b_{19}D_9{*}HW_i + e_i$

Test to see whether the set of interactions ($b_{11}$ through $b_{19}$) is significant ($R^2$ change test): $F(9,240) = 14.81$, $p<.0001$

# Relaxing common slope *assumption* in R

```
model0 = lm(mathscore ~ timeonmath)
model1 = lm(mathscore ~ D1+D2+D3+D4+D5+D6+D7+D8+D9+timeonmath)
model2 = lm(mathscore ~ D1+D2+D3+D4+D5+D6+D7+D8+D9+timeonmath + D1*timeonmath + D2*timeonmath + D3*timeonmath +
            D4*timeonmath + D5*timeonmath + D6*timeonmath + D7*timeonmath + D8*timeonmath + D9*timeonmath)
```

```
Analysis of Variance Table

Model 1: mathscore ~ timeonmath
Model 2: mathscore ~ D1 + D2 + D3 + D4 + D5 + D6 + D7 + D8 + D9 + timeonmath
Model 3: mathscore ~ D1 + D2 + D3 + D4 + D5 + D6 + D7 + D8 + D9 + timeonmath +
    D1 * timeonmath + D2 * timeonmath + D3 * timeonmath + D4 *
    timeonmath + D5 * timeonmath + D6 * timeonmath + D7 * timeonmath +
    D8 * timeonmath + D9 * timeonmath
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1    258 24183
2    249 16082  9    8100.4 20.886 < 2.2e-16 ***
3    240 10342  9    5740.2 14.801 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Now, the equation for each school contains the common intercept + group intercept
AND
common slope + group slope

e.g., Group 2:

$$Y_{i2} = 37.71 + 11.29 + 6.33*hw_i - 9.25*hw_i$$

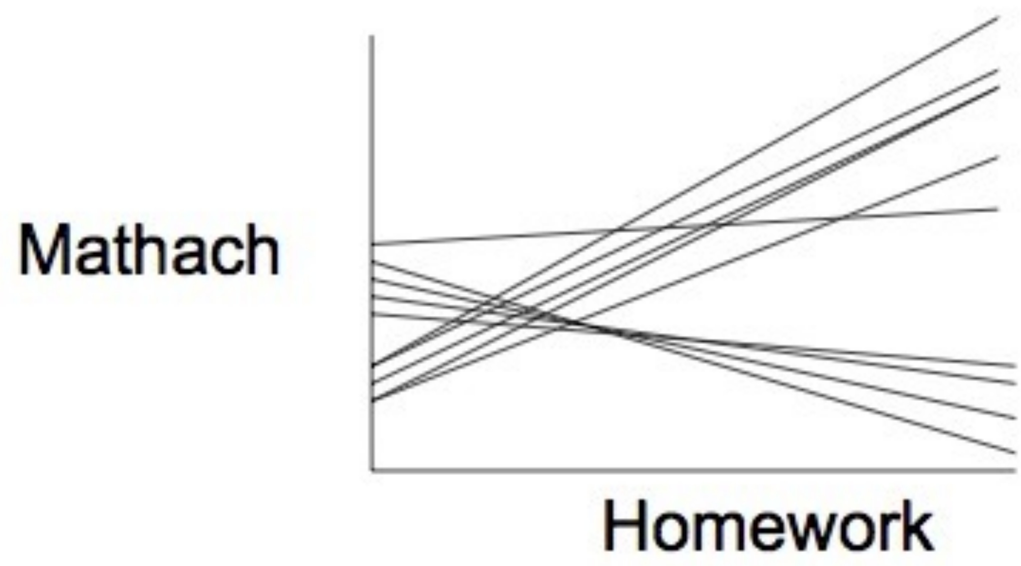
$$Y_{i2} = 49 – 2.92*hw_i$$

**Coefficients**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 1 | (Constant) | 44.074 | .989 | | 44.580 | .000 |
| | Time spent on math homework | 3.572 | .388 | .497 | 9.200 | .000 |
| 2 | (Constant) | 44.431 | 1.899 | | 23.398 | .000 |
| | Time spent on math homework | 2.137 | .384 | .297 | 5.570 | .000 |
| | Dummy: Group 1 | -1.665 | 2.458 | -.043 | -.677 | .499 |
| | Dummy: Group 2 | -7.302 | 2.558 | -.175 | -2.855 | .005 |
| | Dummy: Group 3 | 4.901 | 2.435 | .128 | 2.013 | .045 |
| | Dummy: Group 4 | -4.382 | 2.483 | -.110 | -1.765 | .079 |
| | Dummy: Group 5 | 3.587 | 2.499 | .090 | 1.435 | .152 |
| | Dummy: Group 6 | -.489 | 2.547 | -.012 | -.192 | .848 |
| | Dummy: Group 7 | 11.342 | 2.149 | .446 | 5.278 | .000 |
| | Dummy: Group 8 | .759 | 2.518 | .019 | .301 | .763 |
| | Dummy: Group 9 | -.947 | 2.513 | -.023 | -.377 | .707 |
| 3 | (Constant) | 37.714 | 2.236 | | 16.870 | .000 |
| | Time spent on math homework | 6.335 | 1.054 | .882 | 6.011 | .000 |
| | Dummy: Group 1 | 12.970 | 3.148 | .331 | 4.121 | .000 |
| | Dummy: Group 2 | 11.298 | 3.803 | .271 | 2.971 | .003 |
| | Dummy: Group 3 | 1.036 | 3.425 | .027 | .303 | .763 |
| | Dummy: Group 4 | -3.320 | 3.067 | -.083 | -1.082 | .280 |
| | Dummy: Group 5 | 16.225 | 3.084 | .406 | 5.261 | .000 |
| | Dummy: Group 6 | 11.545 | 3.420 | .277 | 3.376 | .001 |
| | Dummy: Group 7 | 21.496 | 2.834 | .846 | 7.585 | .000 |
| | Dummy: Group 8 | -1.659 | 3.792 | -.041 | -.437 | .662 |
| | Dummy: Group 9 | .806 | 3.413 | .020 | .236 | .813 |
| | D1*HW | -9.889 | 1.637 | -.455 | -6.042 | .000 |
| | D2*HW | -9.255 | 1.560 | -.599 | -5.932 | .000 |
| | D3*HW | 1.574 | 1.606 | .089 | .980 | .328 |
| | D4*HW | -.742 | 1.423 | -.042 | -.522 | .602 |
| | D5*HW | -11.053 | 2.129 | -.324 | -5.192 | .000 |
| | D6*HW | -8.821 | 2.132 | -.299 | -4.138 | .000 |
| | D7*HW | -5.240 | 1.153 | -.794 | -4.543 | .000 |
| | D8*HW | .161 | 1.667 | .009 | .097 | .923 |
| | D9*HW | -.475 | 1.923 | -.019 | -.247 | .805 |

a. Dependent Variable: Math score

# Implied relationships in ANCOVA + interactions model




- →Unique intercept for each school (e.g., $b_0 + b_1$ for school #1)
- →Unique slope for each group (e.g., $b_{10} + b_{11}$ for school #1)
- →Alternatively, estimate regression in each group

# Now we can predict variability in the slopes and intercepts!

10 OLS intercepts and 10 slopes can be used as group-level outcome variables

- I.e., one from each school

Create *separate* models at the group level to predict intercept and slope variability from group-level predictors

# “Slopes and intercepts as outcomes” model

A two-stage analysis that implicitly recognizes the grouped structure of the data:

(1) Run a regression within each group (or an ANCOVA with interactions)

(2) Use the coefficients (one intercept and one slope per group) as outcome variables in regression equations at the group level

# Regressions within each group

| School | $b_{0j}$ | $b_{1j}$ |
|---|---|---|
| School 1: Y = | 50.68 | - 3.55 X |
| School 2: Y = | 49.01 | - 2.92 X |
| School 3: Y = | 38.75 | + 7.91 X |
| School 4: Y = | 34.39 | + 5.59 X |
| School 5: Y = | 53.94 | - 4.72 X |
| School 6: Y = | 49.26 | - 2.49 X |
| School 7: Y = | 59.21 | + 1.09 X |
| School 8: Y = | 36.06 | + 6.50 X |
| School 9: Y = | 38.52 | + 5.86 X |
| School 10: Y = | 37.71 | + 6.34 X |

# Equations at the school level

Regress intercepts and slopes on school-level variable of *school type* (public=1; private=0)

$b_{0j} = g_{00} + g_{01}\ \text{Public}_j$

$b_{0j} = 59.21^{***} - 16.06^{+}\ \text{Public}_j$

$b_{1j} = g_{10} + g_{11}\ \text{Public}_j$

$B_{1j} = 1.09 + 0.96\ \text{Public}_j$

# Slopes and intercepts as outcomes: Results

Describes two lines with marginally different intercepts (-16.06, $p<.10$) and non-significantly different slopes (0.96, ns).

| | Private schools (Public = 0) | Public schools (Public = 1) |
|---|---|---|
| E(Math) for Homework = 0 | 59.21 | 59.21-16.06 = 42.15 |
| Homework – Math slope | 1.09 | 1.09 + 0.96 = 2.05 |

# Slopes and intercepts as outcomes: Advantages

Two-stage estimation procedure acknowledges multilevel data structure:

- First analysis (lower level): associations between L1 outcomes and L1 predictors

- Second analysis (higher level): associations between L2 predictors and group intercepts/slopes

# Slopes and intercepts as outcomes: Disadvantages

Requires large samples in each group
(e.g., big drawback for dyad studies)

Small and large groups/ponds weigh equally

Error terms inconsistently defined at various levels of the model → questionable *p*-values

Better alternative: Random coefficient model!

# Slopes and intercepts as outcomes: When?

- In some cases we might be interested in only the particular groups we have in our analysis, in the values of these particular intercepts and slopes
  - (Fixed effects analysis)

- But in most cases, we want to generalize beyond the groups in our data to the population of similar groups

- So we are not interested in the particular values we have, but in the pattern of values

# The logic: Moving from *sets* to *distributions* of coefficients

Our groups → A random sample of possible groups

Our lines → A random sample of possible lines (i.e., slopes and intercepts)

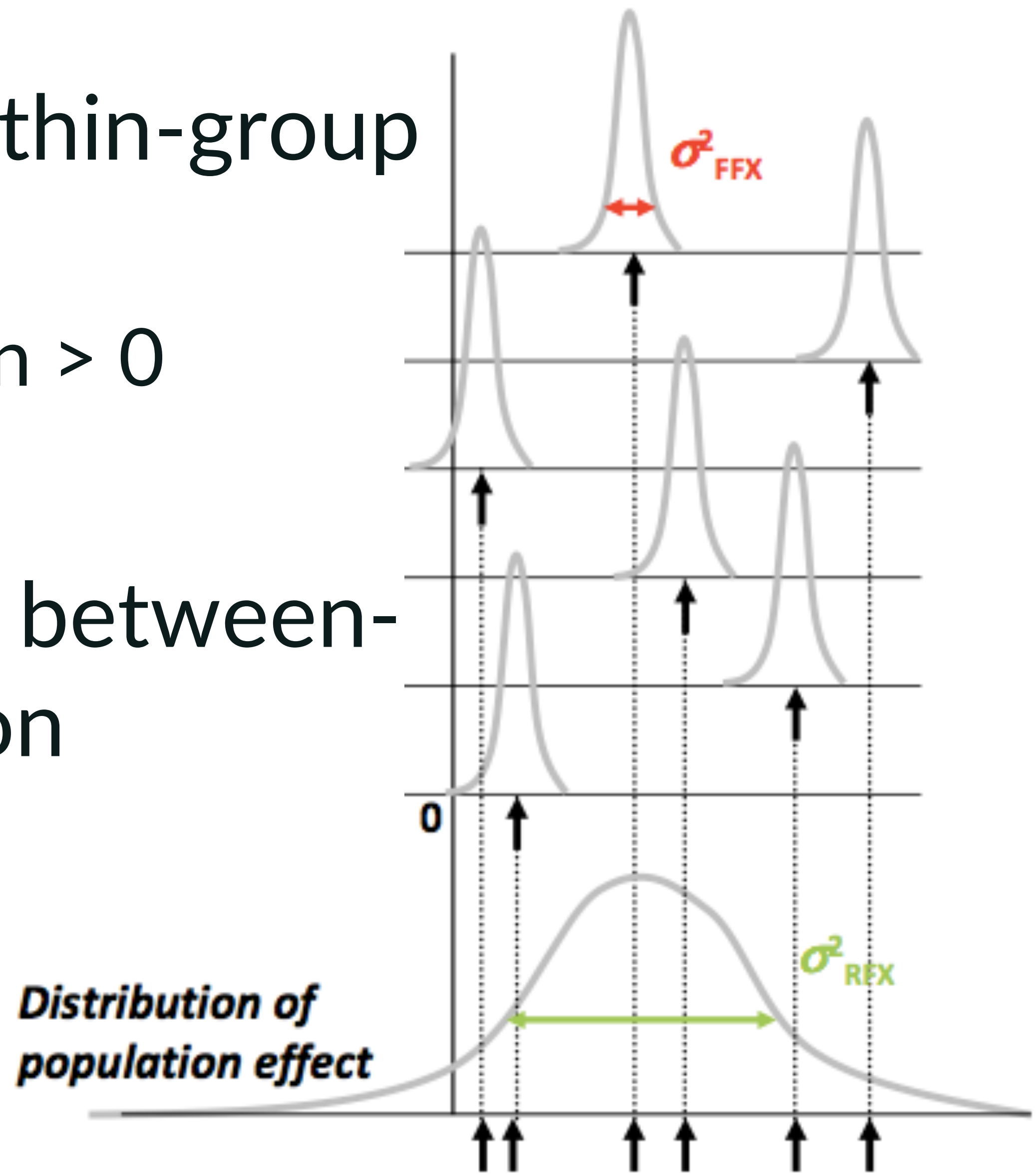
So, rather than characterize the particular lines, our goal is to characterize the *distribution* of lines

- "Average" intercept and some measure of the variability around that intercept
- "Average" slope and the variability around that slope

# Fixed vs. random effects

- Fixed effects (ffx): within-group variation
  - Each group on it's own > 0

- Random effects (rfx): between-group effects variation
  - Population ~= 0



# Fixed vs. random effects

- Different sources of error:
  - FX: only source of error is measurement
    - True group response is fixed
  - RX: sources of error = measurement and sampling
    - True *population* response is fixed
    - But groups vary based on sampling error

# Fixed vs. Random Effects
# "In Other Words"

- Fixed effect inference: *"I can see this effect in this cohort / sample"*
- Random effect inference: *"If I were to sample a new cohort from the same population I would get the same result"*
- Fixed isn't 'wrong', it just is not usually of interest to many people

# Random coefficient model: Multiple equation form

Within-Group (L1) Model: $Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}$

Coefficients of the within-group model then serve as criteria in between-group (L2) models:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Each $\beta$ is a function of an "average" coefficient and random error (variation)

# (Notation)

$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}$

$\beta_{0j} = \gamma_{00} + u_{0j}$ $\leftarrow$ "0": equation for $\beta_0$
$\beta_{1j} = \gamma_{10} + u_{1j}$ $\leftarrow$ "1": equation for $\beta_1$

- ij subscripts indicate individuals within groups
- Double subscripts on gamma coefficients are positional: the 1st indicates equation, the 2nd indicates the position within the equation

# Random coefficient model: Single equation form

Start with multiple equation form:

L1: $Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}$

L2: $\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

Combine: $Y_{ij} = \gamma_{00} + u_{0j} + (\gamma_{10} + u_{1j}) X_{ij} + e_{ij}$

$Y_{ij} = \gamma_{00} + u_{0j} + \gamma_{10} X_{ij} + u_{1j} X_{ij} + e_{ij}$

Rearrange: $Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + (u_{0j} + u_{1j} X_{ij} + e_{ij})$

→ *Looks like a standard regression with a complex error term*

# Estimates

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + (u_{0j} + u_{1j} X_{ij} + e_{ij})$$

(1) “average” intercept - $\gamma_{00}$
(2) “average” slope - $\gamma_{10}$
(3) variability of intercepts - $Var(u_{0j})$
(4) variability of slopes - $Var(u_{1j})$
(5) within-group variability – $Var(e_{ij})$

# Types of parameters / estimates

Fixed effects: $\gamma$s (population parameters)

gs (sample estimates)

*Similar to unstandardized regression params*

Variance components: co/vars of error terms

L1: var($e_{ij}$) = $\sigma^2$ (population parameter)

$s^2$ (sample estimate)

*Similar to regression SS-error*

# Level 2 variance components

$$\begin{bmatrix} Var(u_{oj}) & Cov(u_{1j},u_{0j}) \\ Cov(u_{oj},u_{1j}) & Var(u_{1j}) \end{bmatrix} = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} = \begin{bmatrix} t_{00} & t_{10} \\ t_{01} & t_{11} \end{bmatrix}$$

"True" population values (unknowable)

Sample estimates

*These represent the variance of the slopes and intercepts around the gammas*

NOTE: $t_{10} = t_{01}$ and $\tau_{10} = \tau_{01}$, so there are only **3** variance components to estimate at the second level

# βs

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + (u_{0j} + u_{1j} X_{ij} + e_{ij})$$

Notice that single equation form contains no β's!

Need not calculate βs in order to estimate fixed effects and variance components

Allows estimation in presence of rank deficiency

If desired, can obtain post hoc estimates of β's

# Example from NELS-88 data

Predicted Math = 44.77*** + 2.04 Homework

$\sigma^2 = 43.07\ (3.93)^{***}$

$t_{00} = 69.24\ (34.97)\ ^{*}$

$t_{01} = -31.72\ (18.14)^{+}$     $t_{11} = 22.43\ (11.49)^{+}$

→ *Significant variability to be explained in the intercepts across schools*

# Example from NELS-88 data

Predicted Math = 44.77*** + 2.04 Homework

$\sigma^2 = 43.07\ (3.93)^{***}$

$t_{00} = 69.24\ (34.97)\ ^{*}$

$t_{01} = -31.72\ (18.14)^{+}$ $t_{11} = 22.43\ (11.49)^{+}$

*ICC = between / (between+within)*

*= 69.24 / (69.24+43.07) = 61.6%*

# Adding L2 predictors

Multiple equation form:

L1: $Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}$

L2: $\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j + u_{1j}$

In the combined (single) equation, this is:

$Y_{ij} = \gamma_{00} + \gamma_{01} W_j + \gamma_{10} X_{ij} + \gamma_{11} W_j X_{ij} \ldots$
$+ (e_{ij} + u_{0j} + u_{1j} X_{ij})$

# Fixed coefficients (w/ L2 predictors)

$\gamma_{00}$ – an overall intercept

$\gamma_{01}$ – the main effect of the level 2 ($W_j$) variable

$\gamma_{10}$ – the main effect of the level 1 ($X_{ij}$) variable

$\gamma_{11}$ – the *cross-level interaction* (the effect of the $W_j$ variable on the relationship between $X_{ij}$ and $Y_{ij}$)

# Variance components (with L2 predictors)

With a level 2 ($W_j$) predictor now in the model, our distribution of intercepts/slopes is conditional, i.e., expected value depends on the value of $W_j$ for a particular group

L2 error terms (u's) now represent *residuals*, after controlling for $W_j$ (no longer total variability in intercepts/slopes, but variability remaining after adjusting for $W_j$)

# Example from NELS-88

Predicted Math = 59.21*** + 1.09 Homework …
-15.97⁺ public + .95 HW*Public

$\sigma^2$ = 42.96 (3.91)***
$t_{00}$ = 51.81 (28.64)$^+$
$t_{01}$ = -36.70 (20.07)$^+$ $t_{11}$ = 27.26 (14.59)$^+$

→ *There is no longer significant L2 variability in the intercepts to be explained!*

# Interaction in MLM: Example

- Revisit the NELS-88 data (National Education Longitudinal Study 1988)

- Predict math achievement score from:
  - L1: Hours of homework + SES + interaction
  - L2: Random intercept and HW-slope

# Interaction in MLM: Example

L1: $Y_{ij} = \beta_{0j} + \beta_{1j}\ HW_{ij} + \beta_{2j}\ SES_{ij} + \beta_{3j}\ HW*SES + e_{ij}$

L2: $\beta_{0j} = \gamma_{00} + u_{0j}$ ← Intercept equation

$\beta_{1j} = \gamma_{10} + u_{1j}$ ← HW equation

$\beta_{2j} = \gamma_{20}$ ← SES equation

$\beta_{3j} = \gamma_{30}$ ← HW*SES intx equation

With two 2nd level random error terms ($u_0$, $u_1$), we'll get a 2x2 var/cov matrix with 3 unique var. params.

# Interaction in MLM: Example

First, estimate the "interaction null" model:

```
no_int = lmer(mathscore ~ c_timeonmath
              + ses + (c_timeonmath | schoolid))
```

I.e., fixed effects for "time on math" and "ses" but only random effects for intercept and "time on math"

# Interaction in MLM: Example

*Key question: Does the interaction add anything above and beyond this?*

```
Linear mixed model fit by REML ['lmerMod']
Formula: mathscore ~ timeonmath + ses + (timeonmath | Schoolid)

REML criterion at convergence: 1745.3

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.46818 -0.67278 -0.00633  0.64581  2.63591

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Schoolid (Intercept) 55.17    7.428
          timeonmath  20.10    4.483    -0.88
 Residual             41.27    6.425
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
            Estimate Std. Error t value
(Intercept)  46.1123     2.4840  18.564
timeonmath    1.8198     1.4742   1.234
ses           2.7601     0.6099   4.525
```

Chi-squared value of overall model fit

$g_{00}$

$g_{10}$

$g_{20}$

L1: $Y_{ij} = \beta_{0j} + \beta_{1j}\, HW_{ij} + \beta_{2j}\, SES_{ij} + e_{ij}$

L2: $\beta_{0j} = 46.1 + u_{0j}$

$\beta_{1j} = 1.8 \quad + u_{1j}$

$\beta_{2j} = 2.8$

# Interaction in MLM: Example

Second, estimate the interaction model:

```
int_model = lmer(mathscore ~ c_timeonmath +
                 ses + c_timeonmath*ses +
                 (c_timeonmath | schoolid))
```

I.e., fixed effects for "time on math," "ses," and their interaction, but still only random effects for intercept and "time on math"

# Interaction in MLM: Example

```
Fixed effects:
                    Estimate Std. Error t value
(Intercept)          49.8105     1.4782   33.70
c_timeonmath          1.7402     1.5027    1.16
ses                   2.6858     0.6218    4.32
c_timeonmath:ses     -0.3084     0.4652   -0.66
```

$g_{00}$, $g_{10}$, $g_{20}$, $g_{30}$

L1: $Y_{ij} = \beta_{0j} + \beta_{1j}\ HW_{ij} + \beta_{2j}\ SES_{ij} + \beta_{3j}\ HW{*}SES_{ij} \quad + e_{ij}$

L2: $\beta_{0j} = 49.8 + u_{0j}$ ← Intercept equation

$\beta_{1j} = 1.7 \quad + u_{1j}$ ← HW equation

$\beta_{2j} = 2.7$ ← SES equation

$\beta_{3j} = -0.3$ ← HW*SES interaction equation

```
REML criterion at convergence: 1744.6
```

*Use chi-squared change test from the deviance score (with 1 df)*

# Interaction in MLM: Example

Let's plot it!

First, decide which variable will go on x-axis and which will be the "low-med-high" variable.

Second, find range of x-axis and the "low" and "high" scores of that variable

Note: Mean= 0 if centered

# Interaction in MLM: Example

## X-axis: HW-hours. Range: [-2, 5]

## L-M-H: SES. Centered, SD = 1

L1: $Y_{ij} = \beta_{0j} + \beta_{1j}\, HW_{ij} + \beta_{2j}\, SES_{ij} + \beta_{3j}\, HW{*}SES_{ij} \quad + e_{ij}$

L2: $\beta_{0j} = 49.8 + u_{0j}$ ← Intercept equation

$\beta_{1j} = 1.7 \quad + u_{1j}$ ← HW equation

$\beta_{2j} = 2.7$ ← SES equation

$\beta_{3j} = -0.3$ ← HW*SES interaction equation

Combined: $Y_{ij} = 49.8 + 1.7\, HW_{ij} + 2.7\, SES_{ij} - 0.3\, HW{*}SES_{ij} \quad + (u_{0j} + u_{1j}\, HW + e_{ij})$

Rearrange: $Y_{ij} = (49.8 + 2.7\, SES_{ij}) + HW_{ij}\,(1.7 - 0.3\, SES_{ij})$

# Interaction in MLM: Example