The multilevel random coefficient model

Lecture 19

Multivariate Statistics

Psychology 613 – Spring 2022

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

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 β is a function of an "average" coefficient and random error (variation)

(Notation)

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

$\beta_{0j} = \gamma_{00} + u_{0j}$ ← "0": equation for β_0

$\beta_{1j} = \gamma_{10} + u_{1j}$ ← "1": equation for β_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 - γ_{00}
(2) "average" slope - γ_{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: γs (population parameters)

gs (sample estimates)

Similar to unstandardized regression params

Variance components: co/vars of error terms

L1: $\text{var}(e_{ij})$ = σ^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

L1: $MathAch_{ij} = \beta_{0j} + \beta_{1j}\ Homework_{ij} + e_{ij}$

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

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

The so-called *random effects regression* model:

- Normal intercept (b_0) and slope (b_1) but those terms vary randomly across groups

Example from NELS-88 data

Predicted Math = 44.77*** + 2.04 Homework

σ^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

```
Formula: mathscore ~ timeonmath + (timeonmath | Schoolid)

REML criterion at convergence: 1764

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.5110 -0.5357  0.0175  0.6121  2.5708

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Schoolid (Intercept) 69.30    8.325
          timeonmath  22.45    4.738    -0.81
 Residual             43.07    6.563
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
            Estimate Std. Error     df t value Pr(>|t|)
(Intercept)   44.771      2.744  9.227  16.318 4.06e-08 ***
timeonmath     2.040      1.554  8.785   1.313    0.222
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
           (Intr)
timeonmath -0.804
>
```

t_{00}

t_{11}

t_{01}

s^2

g_{00}

g_{10}

Example from NELS-88 data

Predicted Math = 44.77*** + 2.04 Homework

σ^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)

γ_{00} – an overall intercept

γ_{01} – the main effect of the level 2 (W_j) variable

γ_{10} – the main effect of the level 1 (X_{ij}) variable

γ_{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!*

Example from NELS-88

g_{00}

g_{01}

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	59.210218	7.407272	7.323	7.994	.000	41.850398	76.570038
[schooltype=1]	-15.965674	7.829124	7.402	-2.039	.079	-34.276521	2.345172
[schooltype=4]	0[b]	0	.	.	.	.	.
timeonmath	1.094640	5.243213	7.093	.209	.840	-11.270805	13.460085
[schooltype=1] * timeonmath	.951160	5.542392	7.172	.172	.868	-12.090995	13.993315
[schooltype=4] * timeonmath	0[b]	0	.	.	.	.	.

a. Dependent Variable: Math score.

b. This parameter is set to zero because it is redundant.

g_{10}

g_{11}

Example from NELS-88 (R)

g_{00}

```
Fixed effects:
                     Estimate Std. Error      df t value Pr(>|t|)
(Intercept)           43.2445     2.5353  8.1316  17.057 1.18e-07 ***
timeonmath             2.0458     1.7963  7.9063   1.139   0.2881
schooltype4           15.9657     7.8291  7.4024   2.039   0.0786 .
timeonmath:schooltype4 -0.9512    5.5424  7.1722  -0.172   0.8685
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

g_{10}

g_{01}

g_{11}

Example from NELS-88

	Private schools (Public = 0)	Public schools (Public = 1)
E(Math) for Homework = 0	59.21	59.21-15.97 = 43.26
Homework – Math slope	1.09	1.09 + 0.95 = 2.04

Example from NELS-88

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Schoolid (Intercept) 51.84    7.200
          timeonmath  27.27    5.222    -0.98
 Residual             42.96    6.554
Number of obs: 260, groups:  Schoolid, 10
```

Check the variance of the intercepts (t_{00}) and the slopes (t_{11})…

Assumptions of the multilevel random coefficient models

$Cov(X_{ij}, e_{ij}) = 0$

$Cov(W_j, e_{ij}) = 0$

$Cov(W_j, u_j) = 0$

$Cov(X_{ij}, u_j) = 0$

$Cov\ (e_{ij}, u_j) = 0$

e_{ij} ~ iid $N(0, \sigma^2)$ within each group j

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{iid}, N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} t_{00} & t_{10} \\ t_{01} & t_{11} \end{bmatrix} \right)$$

Full multilevel random coefficient model: Recap

Multi-equation form:

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

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

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

Single equation form:

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

- Predictors at both levels
- All L1 coefficients (intercept and slope(s)) treated as random
- **Simplifying the models gives several useful submodels**

Random effects ANOVA model

Null model / empty model <- ***USE THIS TO GET AN ICC***

No predictors, only a random intercept

L1: $Y_{ij} = b_{0j} + e_{ij}$

L2: $b_{0j} = g_{00} + u_{0j}$

Single equation: $Y_{ij} = g_{00} + (u_{0j} + e_{ij})$

g_{00} represents the overall level or “average” of the dependent variable

Variance components partition variance into within groups (σ^2) and between groups (τ_{00}) portions

g_{00}

$$Y_{ij} = 48.86^{***} + (u_{0j} + e_{ij})$$

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	48.861323	1.927392	9.631	25.351	.000	44.544428	53.178217

a. Dependent Variable: Math score.

Covariance Parameters

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		72.255843	6.455032	11.194	.000	60.649918	86.082671
Intercept [subject = Schoolid]	Variance	34.011005	16.920147	2.010	.044	12.827880	90.174558

a. Dependent Variable: Math score.

$\mathrm{var}(u_{0j})$ $\mathrm{Var}(e_{ij})$

Calculating the ICC

ICC = τ_{00} / (τ_{00}+σ^2)

NELS-88 Example: DV = math achievement

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	SE	Z-score	P(z)
Intercept	school	34.0116.92	2.01	0.0222	
Residual		72.266.46		11.19	<.0001

→ ICC = 34.01 / (34.01 + 72.26) = 0.32

One-way ANCOVA with random effects

Add non-random L1 variable X_{ij}

Allows intercepts to differ randomly by assumes a common slope

L1: $Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$ Multi-equation form

L2: $b_{0j} = g_{00} + u_{0j}$

$b_{1j} = g_{10}$ (No random error term)

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

Single equation form

$$Y_{ij} = 44.98 + 2.21 * math + (u_{0j} + e_{ij})$$

```
Formula: mathscore ~ timeonmath + (1 | Schoolid)

REML criterion at convergence: 1839.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.6060 -0.6872 -0.0244  0.5983  3.3770

Random effects:
 Groups   Name        Variance Std.Dev.
 Schoolid (Intercept) 25.22    5.022
 Residual             64.52    8.033
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)   44.982      1.803  12.670  24.949 3.76e-12 ***
timeonmath     2.207      0.379 257.150   5.823 1.72e-08 ***
---
```

$var(u_{0j})$

$var(e_{ij})$

g_{00}

g_{10}

Random coefficients regression (RCR)

Add random error to X_{ij} slope

L1: $Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$ Multi-equation form

L2: $b_{0j} = g_{00} + u_{0j}$

$b_{1j} = g_{10} + u_{1j}$ (random error term)

$Y_{ij} = g_{00} + g_{10}X_{ij} + (u_{0j} + u_{1j}X_{ij} + e_{ij})$ Single equation form

Allows both intercept and slope to vary randomly

Does not yet predict L2 variation in intercept or slope

$$Y_{ij} = 44.77 + 2.04{*}\text{math} + (u_{0j} + u_{1j}{*}\text{math} + e_{ij})$$

```
Formula: mathscore ~ timeonmath + (timeonmath | Schoolid)

REML criterion at convergence: 1764

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.5110 -0.5357  0.0175  0.6121  2.5708

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Schoolid (Intercept) 69.30    8.325
          timeonmath  22.45    4.738    -0.81
 Residual             43.07    6.563
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
            Estimate Std. Error     df t value Pr(>|t|)
(Intercept)   44.771      2.744  9.227  16.318 4.06e-08 ***
timeonmath     2.040      1.554  8.785   1.313    0.222
---
```

$\text{var}(u_{0j})$

$\text{var}(u_{1j})$

$\text{var}(e_{ij})$

g_{00}

g_{10}

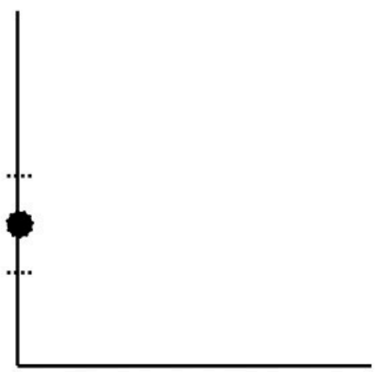
Testing for significant variance

- The significance tests of the variance components are suspect because variances are not normally distributed
- BUT, can create nested models that differ by only one variance term

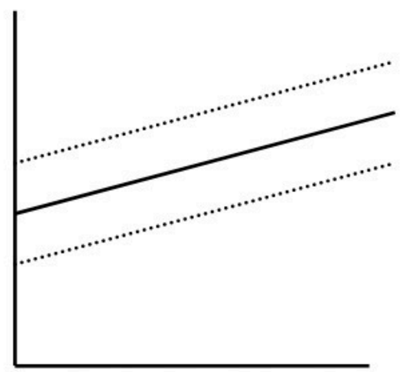
```
> anova(model1,model2)
refitting model(s) with ML (instead of REML)
Data: NULL
Models:
object: mathscore ~ timeonmath + (1 | Schoolid)
..1: mathscore ~ timeonmath + (timeonmath | Schoolid)
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
object  4 1850.7 1864.9 -921.33   1842.7
..1     6 1781.4 1802.8 -884.69   1769.4 73.272      2  < 2.2e-16 ***
---
```

Graphs

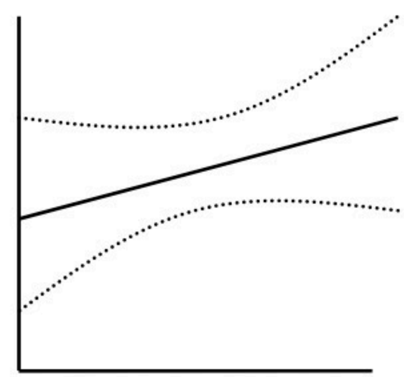
Random Coefficient Models:



ANOVA



ANCOVA



RCR

Add main effect of L2 variable

Add W_j to intercept equation (no cross-level interaction)

L1:	$Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$	Multi-equation form
L2:	$b_{0j} = g_{00} + g_{01}W_j + u_{0j}$	
	$b_{1j} = g_{10} + u_{1j}$	(random error term)

$Y_{ij} = g_{00} + g_{10}X_{ij} + g_{01}W_j + (u_{0j} + u_{1j}X_{ij} + e_{ij})$
Single equation form

"Main effects" of L1 and L2 variables (g_{10} and g_{01}, respectively)

$$Y_{ij} = 38.5 + 1.9*\text{math} + 4.9*\text{type} + (u_{0j} + u_{1j}X_{ij} + e_{ij})$$

```
Formula: mathscore ~ timeonmath + schooltype + (timeonmath | Schoolid)

REML criterion at convergence: 1746.2

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.63744 -0.56207 -0.05469  0.64216  2.58580

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Schoolid (Intercept) 45.84    6.771
          timeonmath  23.96    4.895    -0.97
 Residual             42.96    6.554
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  38.4930     2.4558 11.7740  15.675 3.01e-09 ***
timeonmath    1.9524     1.5990  8.7760   1.221 0.253869
schooltype    4.8869     0.7025  5.9730   6.957 0.000447 ***
---
```

$var(u_{0j})$

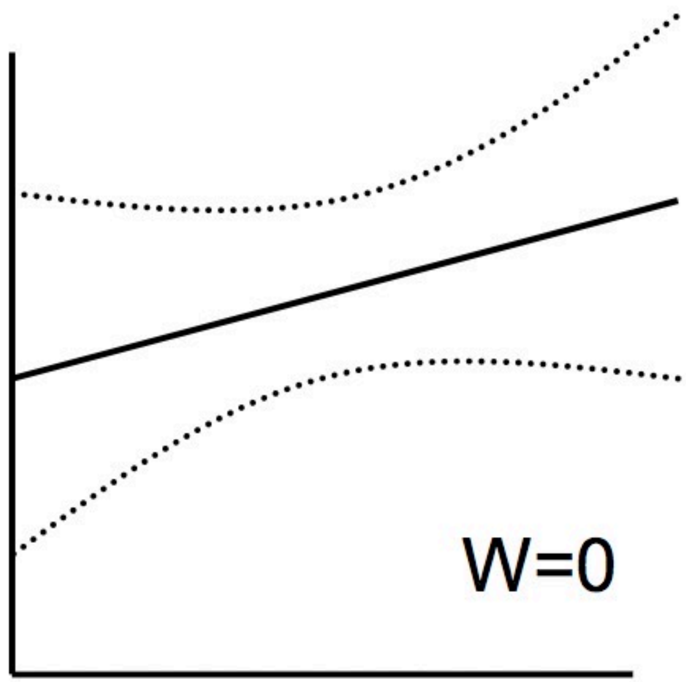
$var(u_{1j})$

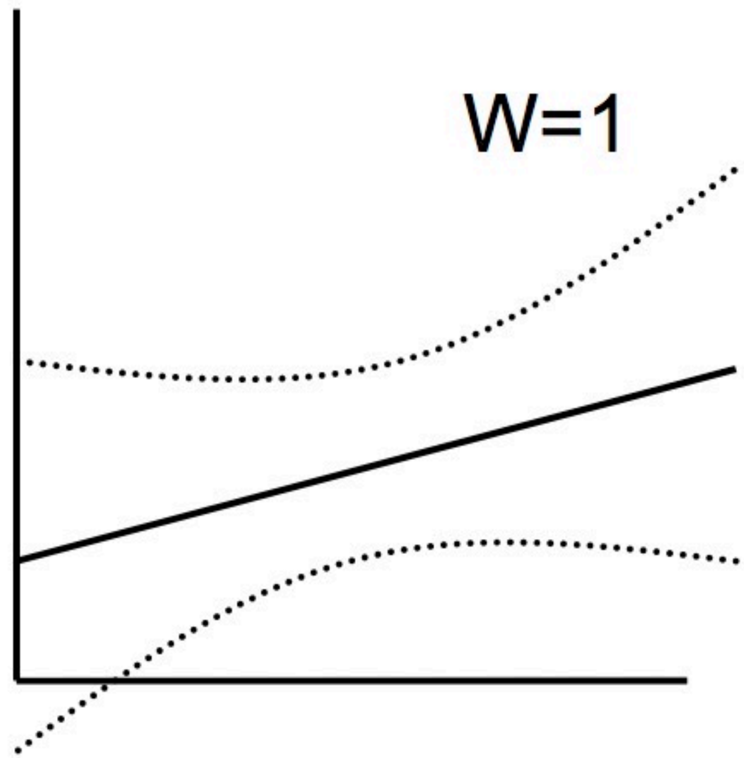
$var(e_{ij})$

g_{10} g_{01} g_{00}

Graphs

Random coefficient models with dichotomous L2 predictor W_j (e.g., public=0, private=1).





Intercepts can vary within- and between-groups
Slopes only vary within-groups

Add cross-level interaction

Add W_j to slope equation

L1: $Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$ Multi-equation form

L2: $b_{0j} = g_{00} + g_{01}W_j + u_{0j}$

$b_{1j} = g_{10} + g_{11}W_j + u_{1j}$

$Y_{ij} = g_{00} + g_{10}X_{ij} + g_{01}W_j + g_{11}X_{ij}W_j + (u_{0j} + u_{1j}X_{ij} + e_{ij})$
Single equation form

g_{11} is the effect of W_j on the slope between X_{ij} and Y_{ij}

→*This is the full multilevel model*

$$Y_{ij} = 38 + 2.4{*}math + 5.3{*}type - 0.3{*}math{*}type + (u_{0j} + u_{1j}X_{ij} + e_{ij})$$

```
Formula: mathscore ~ timeonmath + schooltype + timeonmath * schooltype +
    (timeonmath | Schoolid)

REML criterion at convergence: 1743.1

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.64023 -0.56436 -0.05373  0.65209  2.58438

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Schoolid (Intercept) 51.84    7.200
          timeonmath  27.27    5.222    -0.98
 Residual             42.96    6.554
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
                      Estimate Std. Error      df t value Pr(>|t|)
(Intercept)            37.9227     4.1861  7.8360   9.059 2.02e-05 ***
timeonmath              2.3629     2.9650  7.6090   0.797   0.4496
schooltype              5.3219     2.6097  7.4020   2.039   0.0786 .
timeonmath:schooltype  -0.3171     1.8475  7.1720  -0.172   0.8685
---
```

$var(u_{0j})$

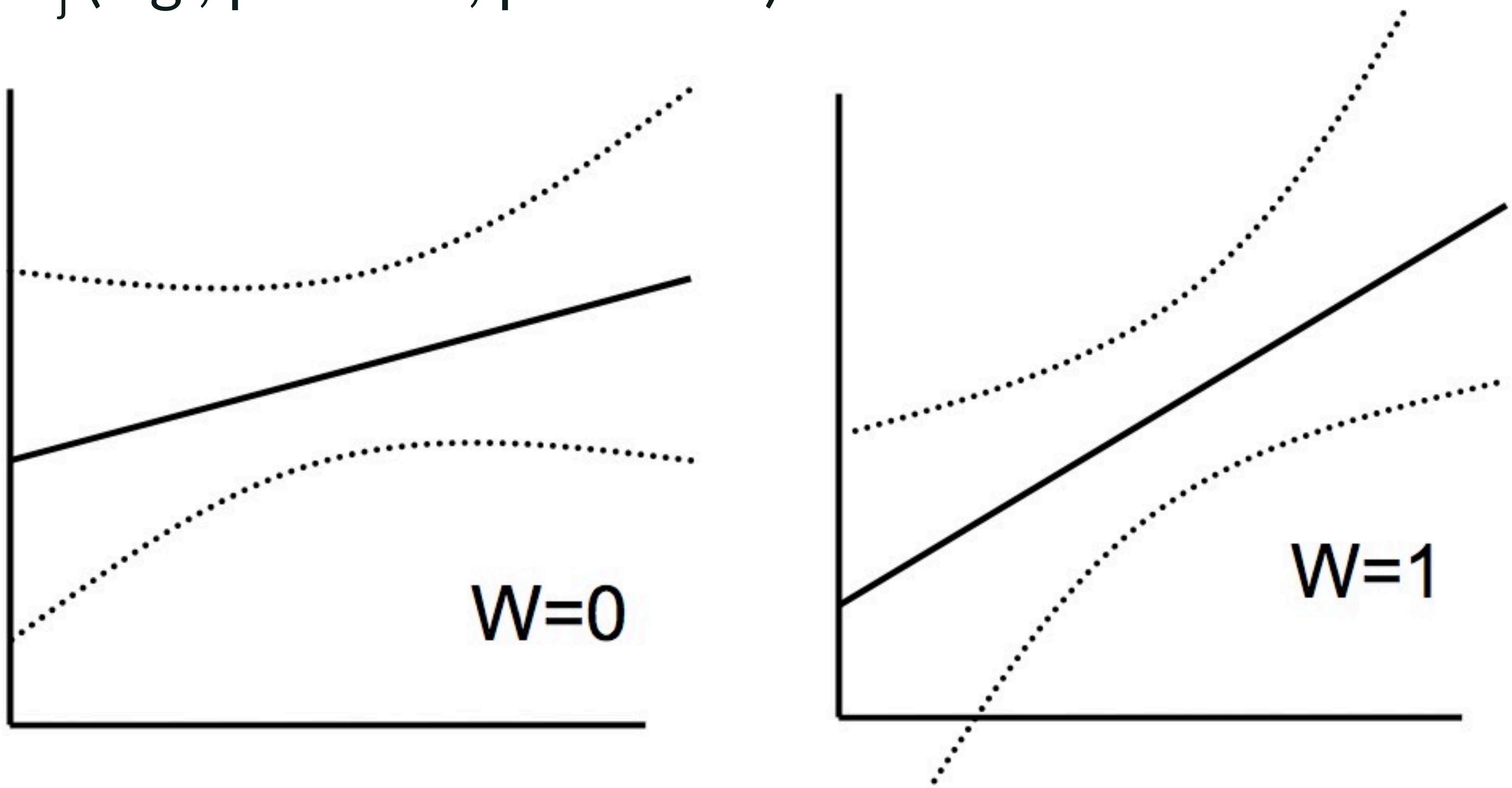
$var(u_{1j})$

$var(e_{ij})$

g_{10} g_{11} g_{01} g_{00}

Graphs

Random coefficient models with dichotomous L2 predictor W_j (e.g., public=0, private=1)



Intercepts can vary within- and between-groups
Slopes can vary within- and between groups

Non-randomly varying slopes

No residual variability in slope after accounting for W_j

L1: $Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$ Multi-equation form

L2: $b_{0j} = g_{00} + g_{01}W_j + u_{0j}$

$b_{1j} = g_{10} + g_{11}W_j$ (No random error)

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

Single equation form

Slope of predictor (X) is entirely determined by overall mean and group effect

$$Y_{ij} = 37.2 + 3.4*\text{math} + 5.5*\text{type} - 0.6*\text{math}*\text{type} + (u_{0j} + e_{ij})$$

```
Formula: mathscore ~ timeonmath + schooltype + timeonmath * schooltype +
    (1 | Schoolid)

REML criterion at convergence: 1826.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.9172 -0.6792 -0.0361  0.6133  3.5261

Random effects:
 Groups   Name        Variance Std.Dev.
 Schoolid (Intercept) 12.13    3.483
 Residual             63.30    7.956
Number of obs: 260, groups:  Schoolid, 10

Fixed effects:
                      Estimate Std. Error       df t value Pr(>|t|)
(Intercept)            37.1984     2.4397  12.7000  15.247 1.54e-09 ***
timeonmath              3.4401     0.6971 255.8200   4.935 1.45e-06 ***
schooltype              5.5030     1.4491  10.0200   3.797  0.00349 **
timeonmath:schooltype  -0.5864     0.2530 253.0900  -2.318  0.02127 *
---
```

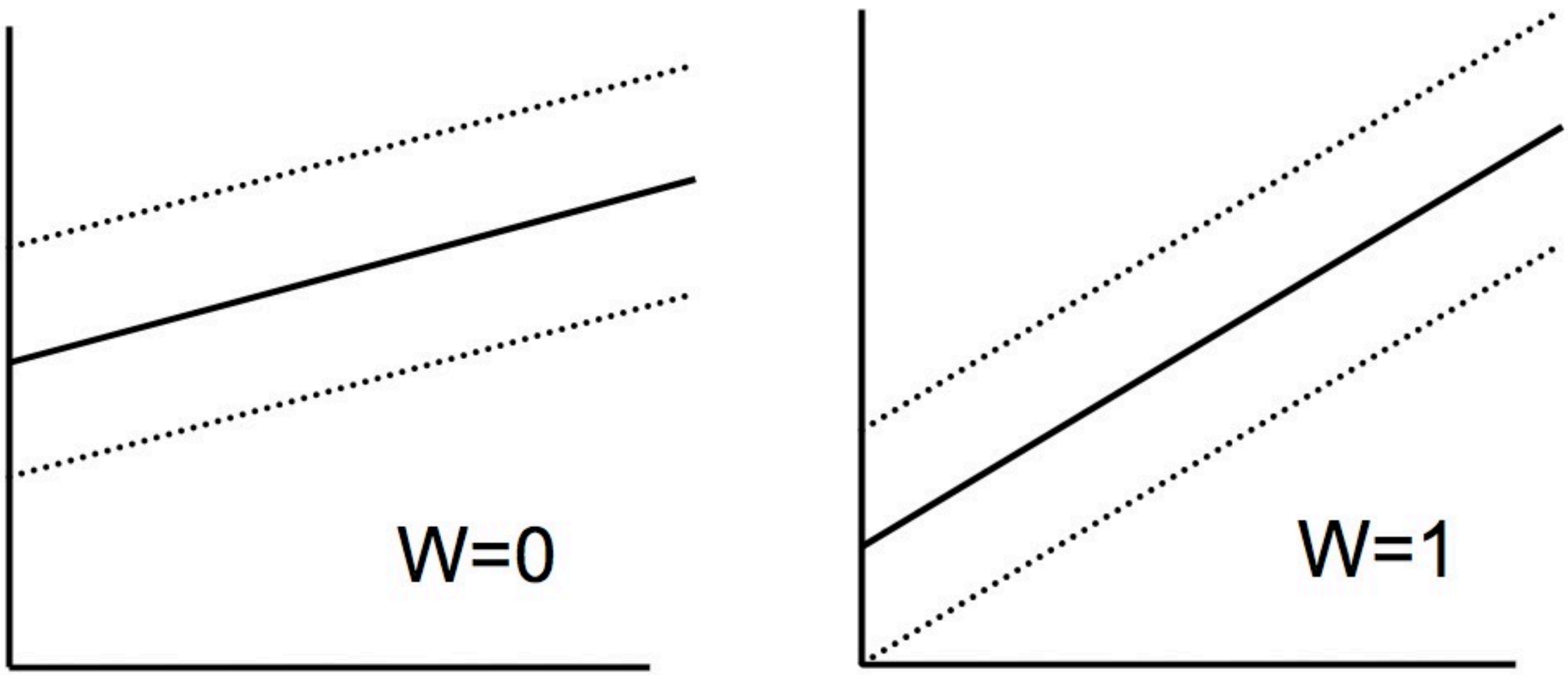
$var(u_{0j})$

$var(e_{ij})$

g_{10} g_{11} g_{01} g_{00}

Graphs

Random coefficient models with dichotomous L2 predictor W_j (e.g., public=0, private=1).



Intercepts can vary within- and between-groups
Slopes only vary between groups (and not within groups)

Non-randomly varying slopes and intercepts

In the extreme, no residual variability in intercepts *or* slopes

L1:	$Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$	Multi-equation form
L2:	$b_{0j} = g_{00} + g_{01}W_j$	(No random error)
	$b_{1j} = g_{10} + g_{11}W_j$	(No random error)

$Y_{ij} = g_{00} + g_{10}X_{ij} + g_{01}W_j + g_{11}X_{ij}W_j + (e_{ij})$
Single equation form

This is equivalent to a single-level disaggregation model, and could be estimated with standard OLS regression