Predictor Variable Centering within Multi-level Models

Lecture 5 and 6

Lecture 5 and 6 Outline

- Review of predictor variable centering in non-clustered data

- Decomposition involves centering in clustered data
  - Centering can be done for level-1 or level-2 covariates
  - Choice of centering will depend on the goal of the analysis
  - We will review 5 common decomposition analyses and discuss ideal choice of centering for each
  - In class, we will walk through models/interpretations for each of these 5 common analyses within a hospital example

- Suggested readings:
  - Enders and Tofighi, 2007
  - Raudenbush and Bryk. Hierarchical Linear Models: Applications and Data Analysis Methods (2nd Edition), pages 134-149.
    - This lecture follows the structure of this section

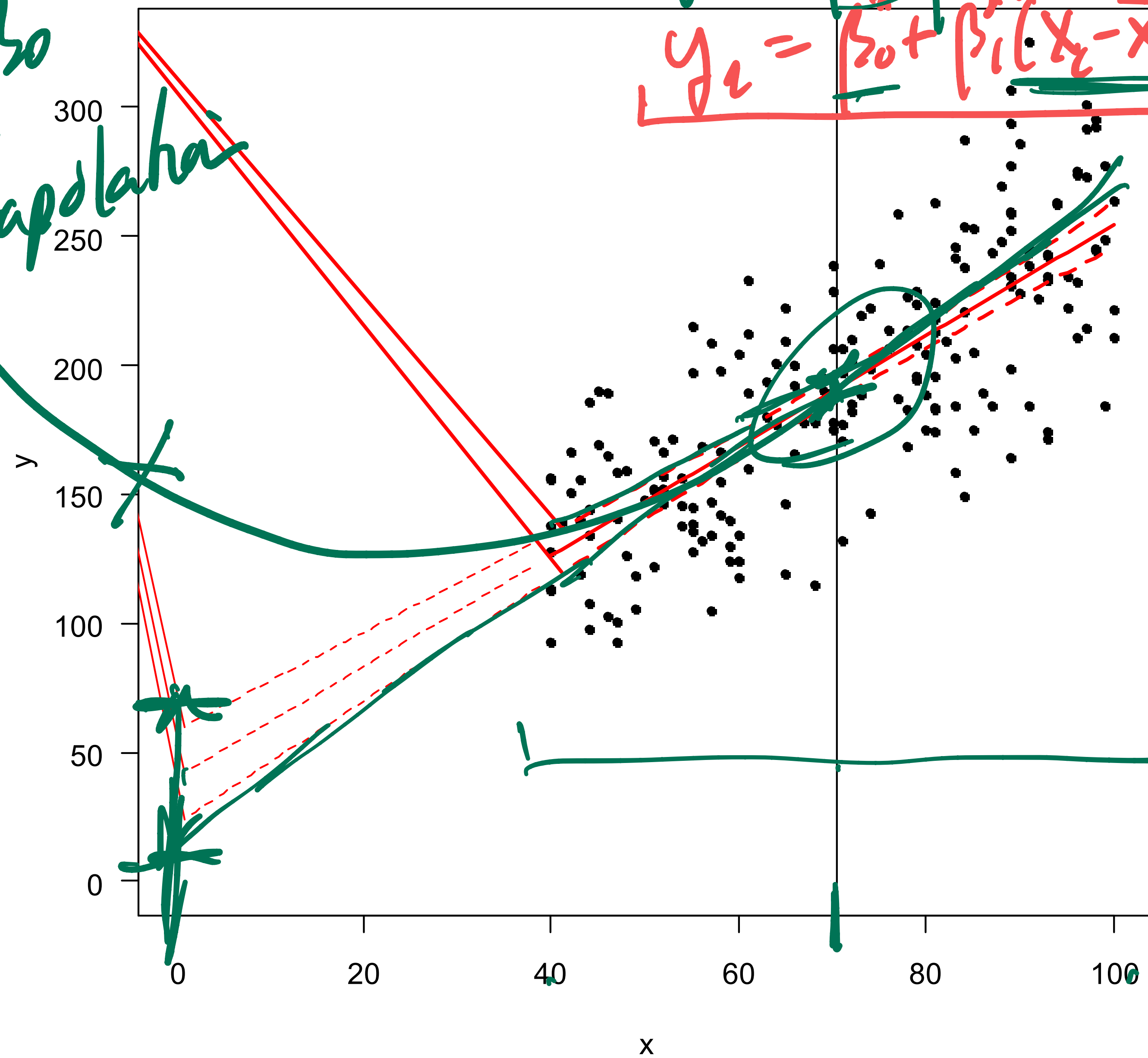
Centering in Non-Clustered Data

- In standard regression models, centering predictors will
  - Change the definition of the intercept
    - Particularly important when the range of the predictors do not naturally include 0
  - Defines the intercept so that it can be more precisely estimated
    - extrapolating to non-observed “0” ranges of the data produces imprecise and non-robust estimates of the intercept

interpretation of β_0 $\longrightarrow$ $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$y_i = \beta_0^* + \beta_1^*(x_i - \bar{x}_\cdot) + \varepsilon_i$

extrapolation



Centering in Multi-level Models

- Centering Level-2 predictors:
  - Only one option: Subtract level-2 mean

G $(\bar{I}_{i\cdot} - \bar{I}_{\cdot\cdot})$

$$Y_{ij} = \beta_{0i} + \beta_1 X_i + \varepsilon_{ij}$$

Interpretation

β_1 - $\Delta E(Y_{ij}|X_i)/\Delta X_i$

β_{0i} - $E(Y_{ij}|X_i = 0)$

$$Y_{ij} = \beta_{0i} + \beta_1(X_i - \bar{X}.) + \varepsilon_{ij}$$

like ordinary regression

Here
i denotes the level-2 cluster
j denotes the level-1 unit

Interpretation of :

β_1 : $\Delta E(Y_{ij}|X_i)/\Delta X_i$

β_{0i} - $E(Y_{ij}|X_i = \bar{X}_i)$ or $\beta_{0i} = E(\bar{Y}_{i\cdot}) - \beta_1(X_i - \bar{X}_\cdot)$

mean Y for average $\bar{X}_i$ mean Y less $(X_i - \bar{X}_\cdot)$ adjustment

Grand-mean Centering of X_i only:

(1) changes the labels on X axis for any multi-level plots

(2) changes interpretation of intercepts (β_{0i}).

* has little effect on level 2 coefficients or inferences

Centering in Multi-level Models

- Centering Level-1 predictors:
  - Two options:
    - Center on the grand mean (i.e. the average over all level-1 and level-2 observations)
    - Center on the group mean (i.e. center the level-1 values using the mean from THEIR level-2 group)

$$Y_{ij} = \beta_{0i} + \beta_1 X_{ij} + \varepsilon_{ij}$$

$$Y_{ij} = \beta_{0i} + \beta_1 (X_{ij} - \bar{X}_{..}) + \varepsilon_{ij}$$ (Center on Grand Mean - CGM)

$$Y_{ij} = \beta_{0i} + \beta_1 (X_{ij} - \bar{X}_{i.}) + \varepsilon_{ij}$$ (Center Within Cluster CWC)

The rest of the lecture will focus on interpretation of these models!

Centering in Multi-level Models

- Choice of centering approach will change interpretation of level-1 intercept and can have an impact on interpretation of level-1 slopes!
- Preferred approach depends on the question being asked: remember Question, Question, Question
- Review 5 common purposes:
  - Estimating level-1 effects $\boxed{X_{ij}} \xrightarrow{\beta} \boxed{Y_{ij}}$
  - Estimating the variances of level-1 coefficients
  - Decomposing person-level from contextual effects
  - Estimating level-2 effects while adjusting for level-1 covariates
  - Estimating an interaction between level-2 and level-1 covariates

Estimating Level-1 Coefficients

- Question: how are the level-1 outcomes related to the level-1 predictors
  - How much higher do we expect a child's GCSE (G) score to be relative to another child in her class if her LRT (L) is one unit higher.
- We will explore centering in this context using the Inner London School data

$$Y_{ij} \text{ gcse for student j in school i}$$

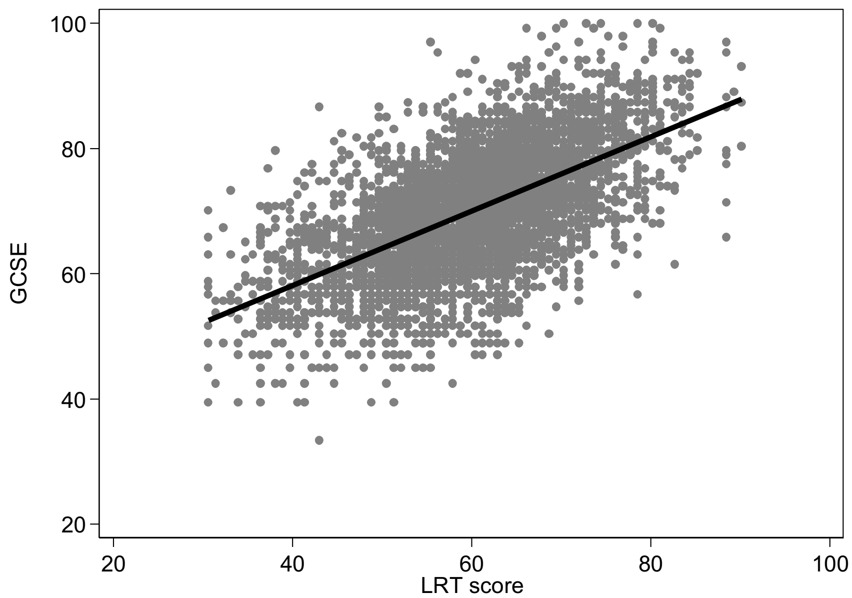
$$x_{ij} \text{ lrt for student j in school i}$$

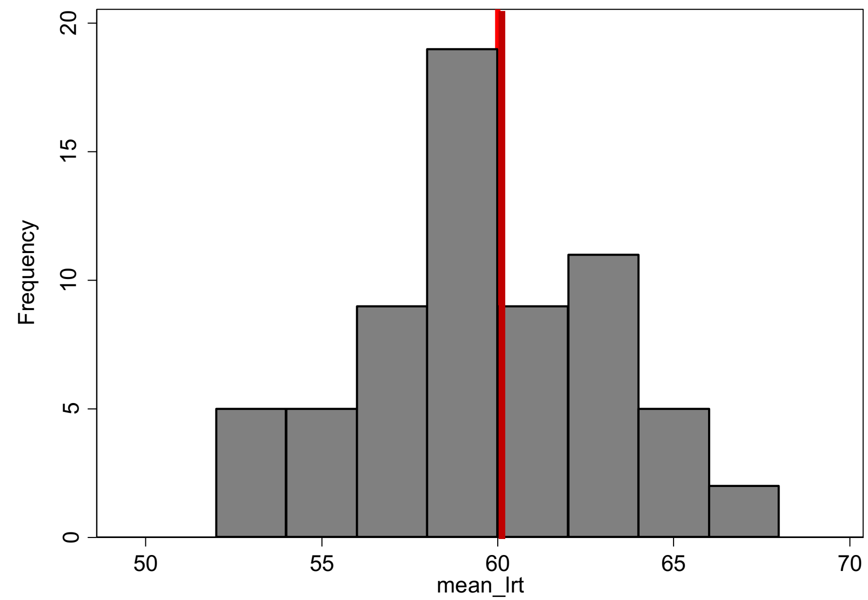
$$Y_{ij} = \beta_{0i} + \beta_1 x_{ij} + \varepsilon_{ij}$$

$$\beta_{0i} \sim N(\beta_0, \tau_1^2)$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Data Visualizations
Display, Look, Think (DLT)





1. Positive association for GCSE and LRT score (slope/correlation 0.59, ignoring clustering)
2. "Substantial" variation in mean LRT scores across schools (histogram)

?

"Substantial"? More than expected if $\boxed{\beta_{0i} = \beta_0 \text{ for all } i}$?

$$Y_{ij} = \beta_{0i} + \beta_1 X_{ij} + \varepsilon_{ij}$$

$\beta_{0i} \overset{iid}{\sim} G(0, \tau_1^2)$

$\varepsilon_{ij} \overset{iid}{\sim} G(0, \sigma^2)$ $\quad \beta_{0i} \perp \varepsilon_{ij}$

$\perp$ - independent of

$$\operatorname{Var} Y_{ij} = \tau_1^2 + \sigma^2$$

$$\operatorname{Var}(\bar{Y}_{i\cdot}) = \tau_1^2 + \sigma^2/n_i$$

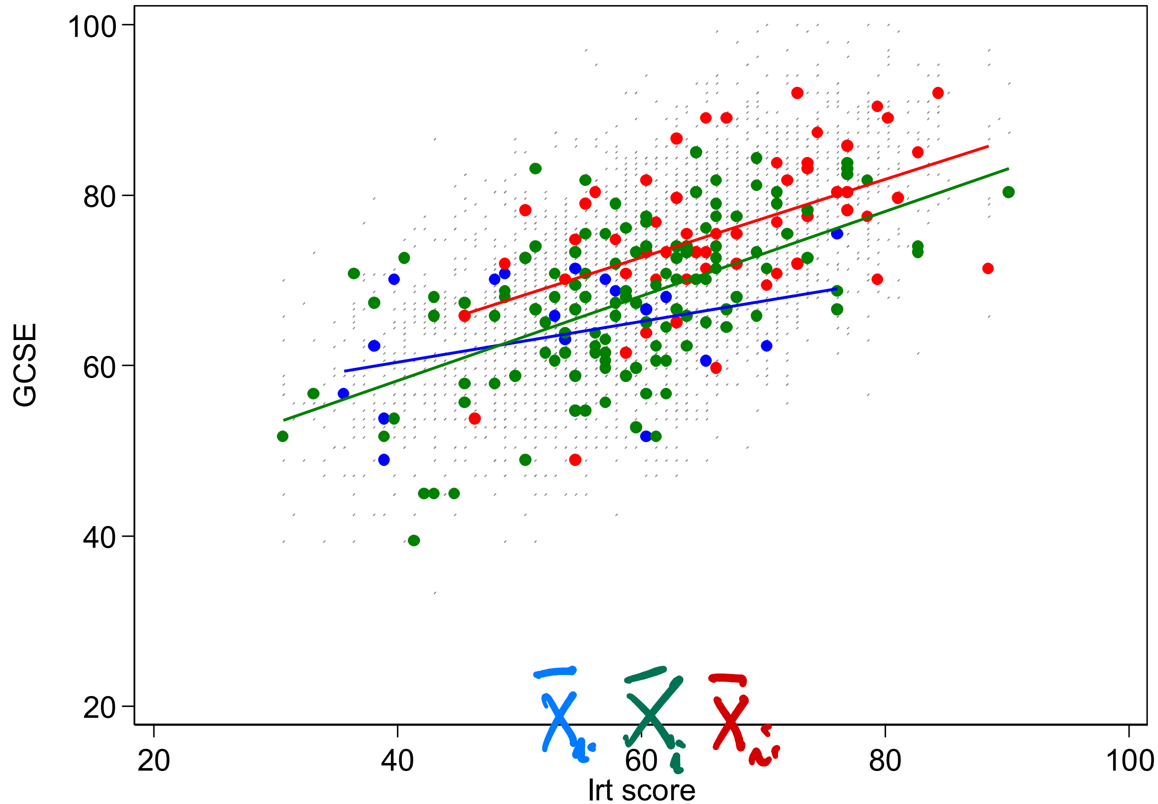
$\boxed{\text{Is } \tau_1^2 = 0 \text{ ?}}$

$$\tau^2 = \operatorname{Var}(\bar{Y}_{i\cdot}) - \sigma^2/n_i$$

$$= \frac{1}{m}\left[\operatorname{Var} \bar{Y}_{i\cdot} - \sigma^2/n_i\right]$$

$$\hat{\tau}^2 = S^2_{\bar{Y}_{i\cdot}} - \left[\overline{\hat{\sigma}^2/n_i}\right] = \frac{1}{m-1}\sum_{i=1}^{m}(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 - \frac{1}{m}\sum_{i=1}^{m}\hat{\sigma}^2/n_i$$

Impact of Centering - 3 random schools



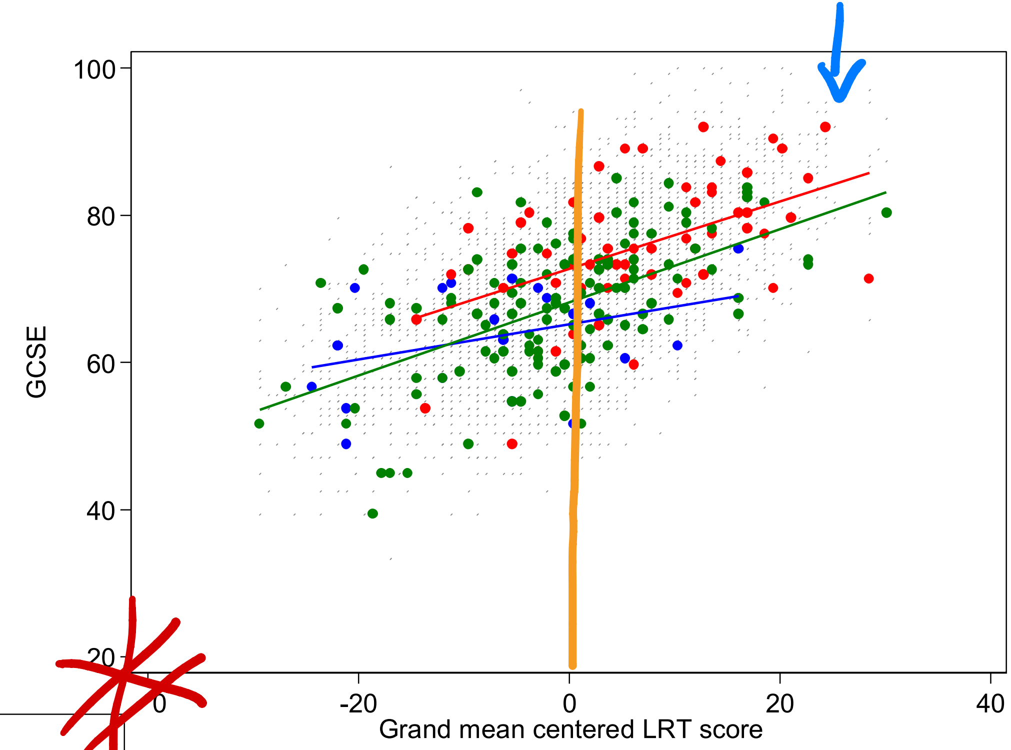
Observed data with data from schools highlighted
The selected schools have the smallest (blue), average (green) and largest (red) within school average LRT.

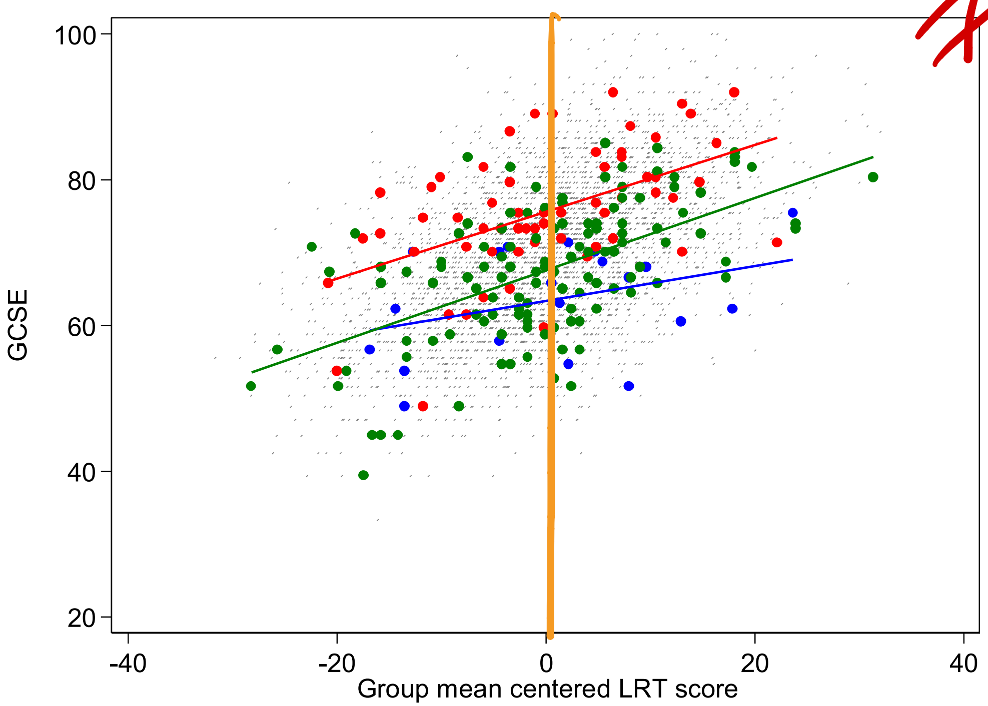
identical except x label

Impact of Centering

Grand mean centering:

- Rescales the x-axis so 0 is at the grand mean
- Y-intercepts are expected GCSE at the grand mean
- Preserves the natural clustering of data in the x-space; i.e. clusters still have different means





Group mean centering:

- Rescales the x-axis so 0 represents cluster mean
- Y-intercepts are expected GCSE at school mean LRT
- Removes the between cluster differences in the mean LRT scores

Which method is preferred given the question asked?

- Question is addressed by estimating β_1: the within cluster linear association between GCSE and LRT scores
  - Group-mean or cluster-mean center
- From lecture 4,

$$\begin{aligned} Model\ 1:\ \ E\left(Y_{ij} \middle| X_{ij}\right) \quad &= \ \alpha_{1i} + \beta_1 X_{ij} \\ Model\ 4:\ \ E\left(Y_{ij} \middle| X_{ij}, \bar{X}_{i.}\right) &= \alpha_{4i} + \beta_4\left(X_{ij} - \bar{X}_{i.}\right) \end{aligned}$$

- The grand-mean centered model is equivalent to model 1 with a different intercept interpretation

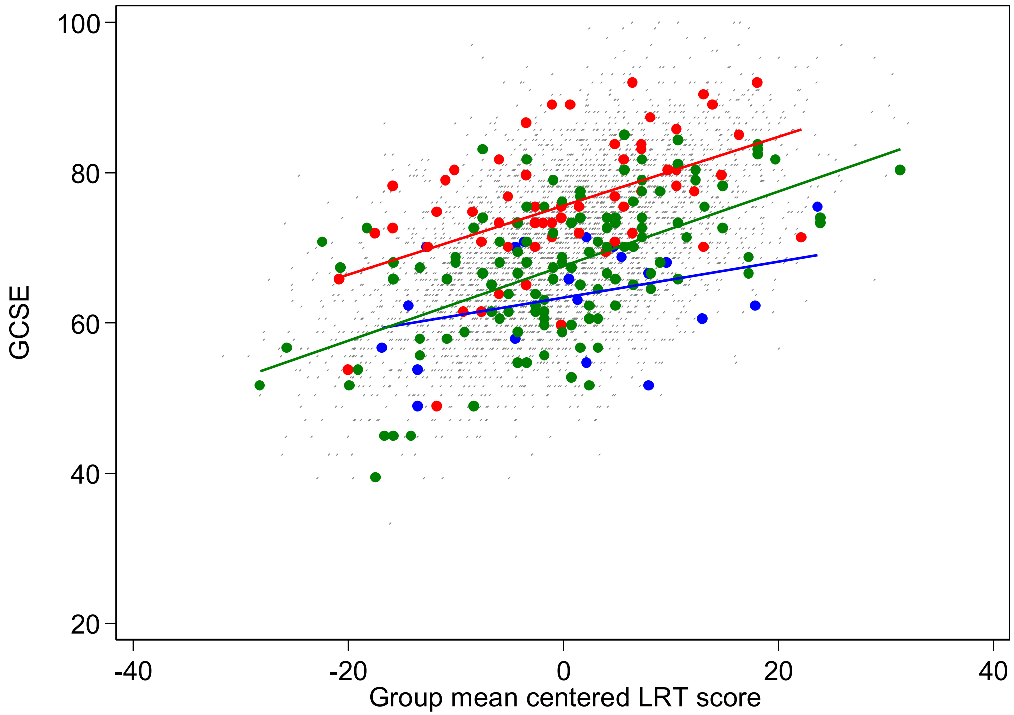
$$E\left(Y_{ij} \middle| X_{ij}\right) \quad = \ \alpha_{0i} + \beta_1(X_{ij} - \bar{X}_{..})$$

Which method is preferred given the question asked?

- Group-mean centered model

$$Y_{ij} = \beta_{0i} + \beta_{1i}\left(X_{ij} - \bar{X}_{i\cdot}\right) + \varepsilon_{ij}$$

- β_1 is the within-school linear association, where we have assumed this is the same across all schools



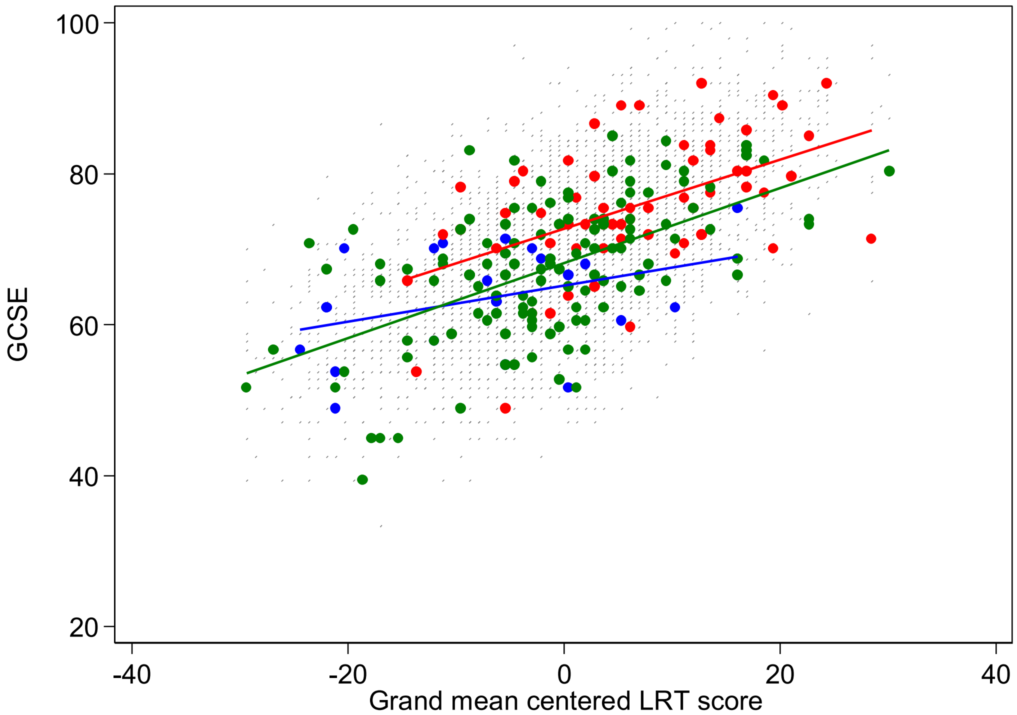
Which method is preferred given the question asked?

- In the grand-mean centered model

$$Y_{ij} = \alpha_{0i} + \beta_1\left(X_{ij} - \bar{X}_{..}\right) + \varepsilon_{ij}$$

β_1 is the total effect including variation in X within and among schools

- β_1 will be confounded by school-level variables.



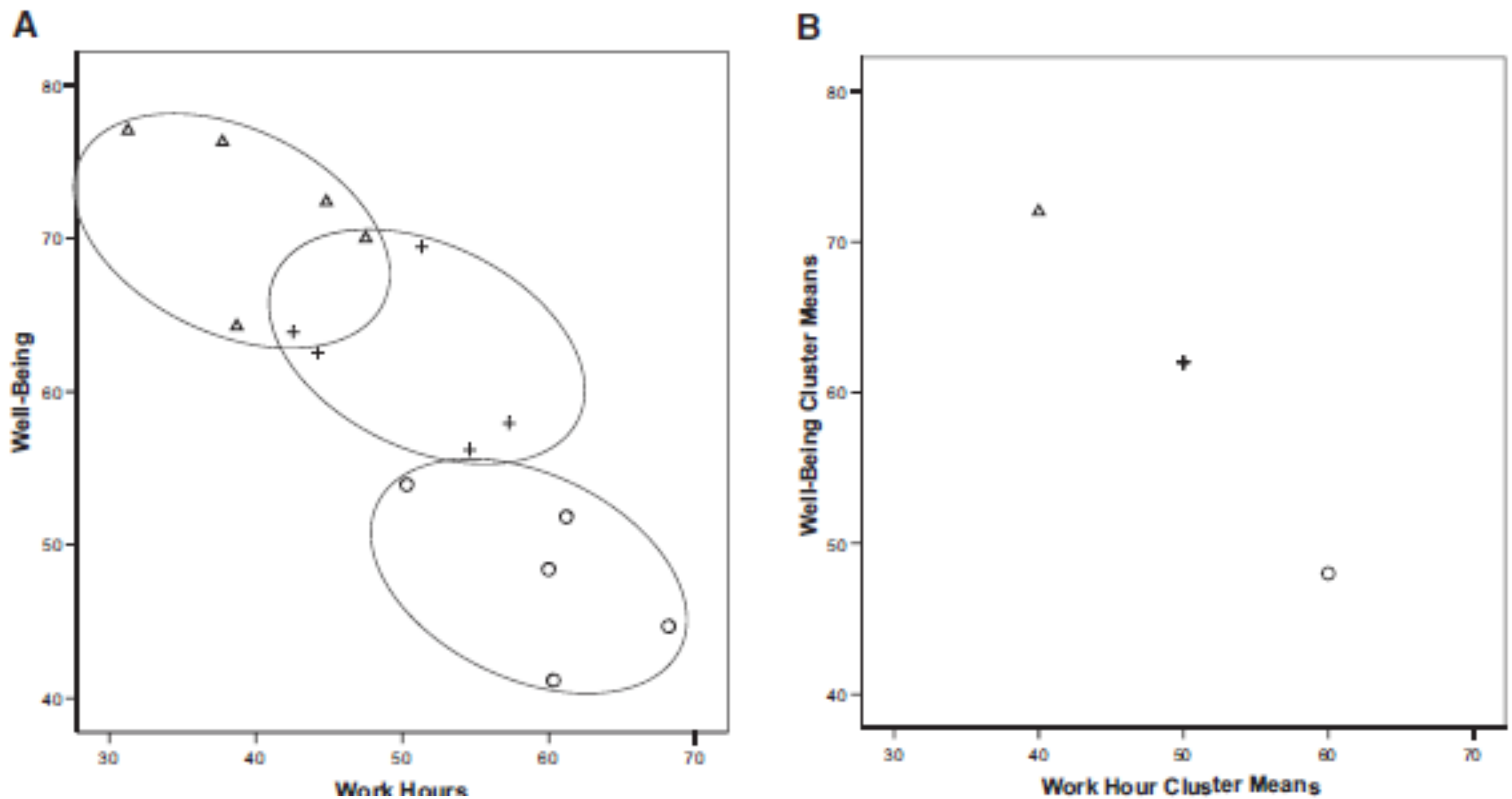
Comparison of Models

Estimate	Group-mean Centered	Grand-mean Centered
β_0	69.84 (0.54)	70.03 (0.40)
β_1	0.559 (0.013)	0.562 (0.012)
τ	4.18 (0.39)	3.03 (0.30)
σ	7.51 (0.08)	7.51 (0.08)

What is driving the difference?

- Confounding by group-mean LRT

More Extreme Example: Enders and Tofighi



More Extreme Example: Enders and Tofighi

- Grand mean centered: CGM
- Group mean centered: CWC

Table 1

Cluster Means and Correlations for the Artificial Data Under Different Forms of Centering

	Cluster M			
Cluster	*WELLBEING*	$HOURS_{raw}$	$HOURS_{cgm}$	$HOURS_{cwc}$
1	72	40	−10	0
2	62	50	0	0
3	48	60	10	0

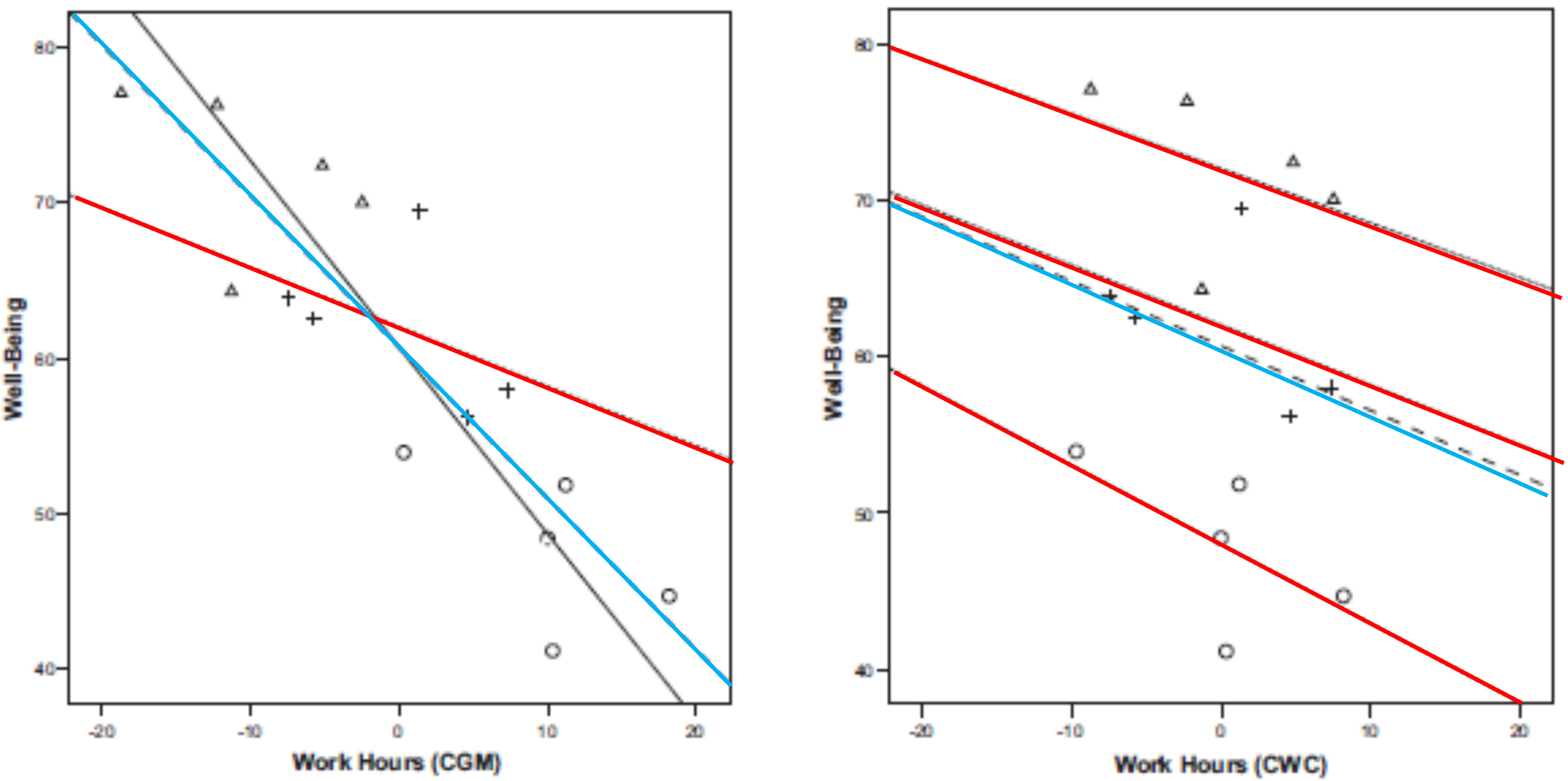
Data	1	2	3	4	5	6	7
1. *WELLBEING*	—						
2. $HOURS_{raw}$	−.86	—					
3. $HOURS_{cgm}$	−.86	1.00	—				
4. $HOURS_{cwc}$	−.22	.57	.57	—			
5. $\bar{x}_{HOURS_j}$	−.90	.82	.82	0	—		
6. $\bar{y}_{WELLBEING_j}$	.90	−.82	−.82	0	−.99	—	
7. *SIZE*	−.44	.33	.33	0	.48	−.48	—

Note. RAW = original metric; CGM = grand mean centered; CWC = centered within cluster; WELLBEING = psychological well-being; *HOURS* = workload measured in hours per week; *SIZE* = workgroup size.

BIG Picture:

- Cluster means: r = -0.99
- Correlation between wellbeing and grand mean centered work hours: r = -0.86
- Correlation between wellbeing and group mean centered work hours: r = -0.22
- The grand mean centered correlation includes information about both the between and within effects

More Extreme Example: Enders and Tofighi



RED: Within cluster effects
BLUE: Estimated average slope from the grand mean (left) or group mean (right) centered model
BLACK: Between cluster association

Comparison of Models - revisited

Estimate	Group-mean Centered	Grand-mean Centered
β_0	69.84 (0.54)	70.03 (0.40)
β_1	0.559 (0.013)	0.562 (0.012)
τ	4.18 (0.39)	3.03 (0.30)
σ	7.51 (0.08)	7.51 (0.08)

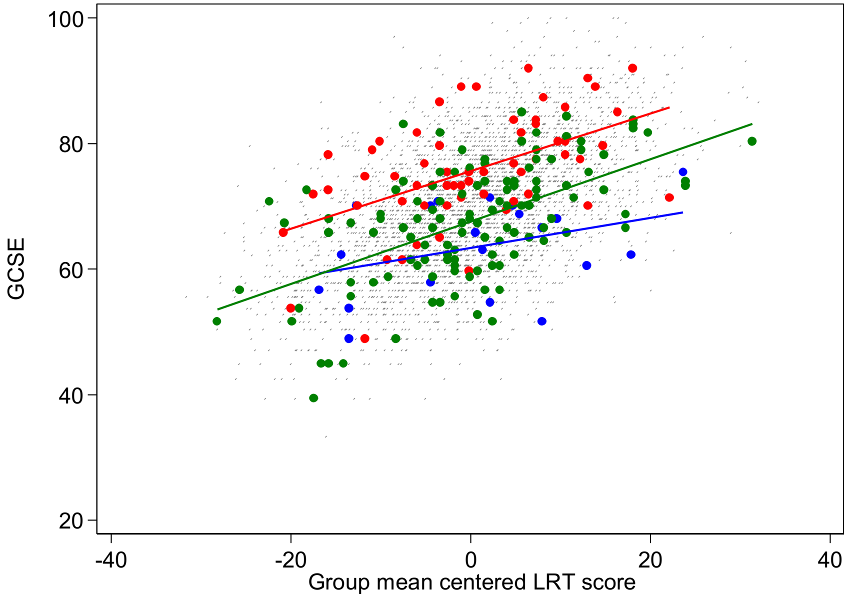
Notice anything else?

- Random intercept standard deviation estimates are substantially different!
- Why?

Centering in Multi-level Models

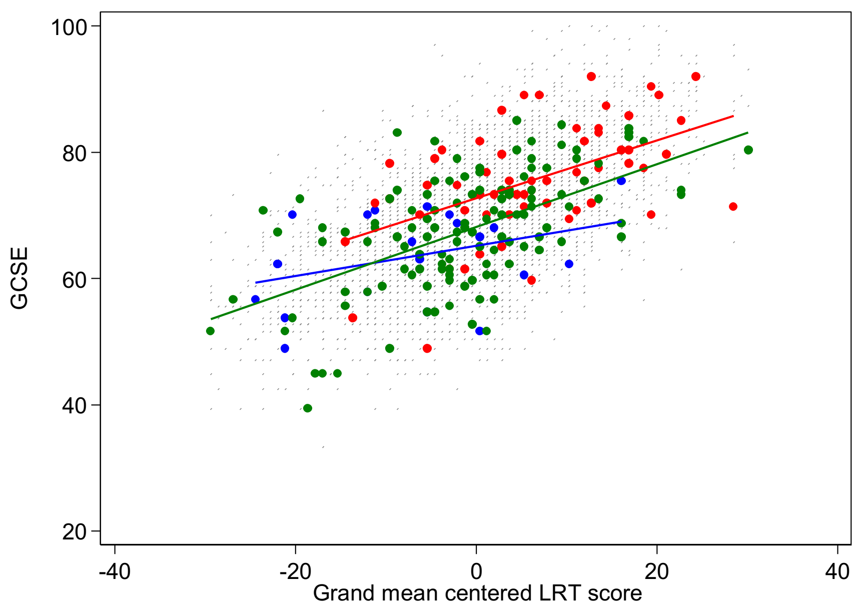
- Choice of centering approach will change interpretation of level-1 intercept and can have an impact on interpretation of level-1 slopes!
- Preferred approach depends on the goal of the analysis
- Review 5 common purposes:
  - Estimating level-1 coefficient
  - Estimating the variances of level-1 coefficients
  - Disentangling person-level vs. contextual effects
  - Estimating level-2 effects while adjusting for level-1 covariates
  - Estimating an interaction between level-2 and level-1 covariates

Effects on Intercept Interpretation!



In group mean centered model:

- Intercepts are Empirical Bayes school-specific means ($\bar{Y}_{EB,i}.$) at the school-specific average LRT
- NOTE: generally there will be little shrinkage here unless there are very few observations within a cluster!
- i.e. $\bar{Y}_{EB,i}. = \bar{Y}_{i}.$



In grand mean centered model:

- intercepts are the average GCSE score for students at the grand mean LRT
- Represent "adjusted" group means!

$$\bar{Y}_{EB,i}. - \beta_B(\bar{X}_i. - \bar{X}..)$$

$$\bar{Y}_i. - \beta_B(\bar{X}_i. - \bar{X}..)$$

Choice of Centering approach effects estimation of the random intercept variance

- Tend to see smaller random intercept variance within the grand-mean approach relative to group-mean approach
  - In group-mean approach, there will be little shrinkage to the average intercept unless a school’s intercept is estimated with little precision (small n for that school)
  - In grand-mean approach, there will be considerable shrinkage to the average intercept since the intercept represents an extrapolation for many schools
  - Therefore, we tend to see less variation in the estimates of the random intercepts within the grand-mean model relative to the group-mean model

Choice of Centering approach effects estimation of the random **SLOPE** variance

- In the grand-mean centering approach
  - As the cluster intercepts will be shrunk more towards the overall average intercept
  - This will affect calculation of the cluster slopes
    - Schools with relatively flat slopes but who have average LRT far from the grand mean will see substantial increases in cluster slope estimate
    - Overall result: homogenization of the slope variability and underestimate of random slope variance

$$Y_{ij} = \beta_{0i} + \beta_{1i} X_{ij} + \varepsilon_{ij} \qquad \begin{pmatrix} \beta_{0i} \\ \beta_{1i} \end{pmatrix} \sim G\left(\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, D_{2\times 2} \right)$$

$$D = \begin{pmatrix} \mathrm{Var}\,\beta_{0i} & \mathrm{Cov}(\beta_{0i}, \beta_{1i}) \\ & \mathrm{Var}(\beta_{1i}) \end{pmatrix}$$

GRAND-MEAN CENTER

$$\begin{aligned} Y_{ij} &= \beta_{0i} + \beta_{1i}(X_{ij} - \overline{X}_{\cdot\cdot}) + \varepsilon_{ij} \\ &= (\beta_{0i} - \beta_{1i}\overline{X}_{\cdot\cdot}) + \beta_{1i} X_{ij} + \varepsilon_{ij} \\ &= \beta_{0i}^{*} + \beta_{1i}^{*} X_{ij} + \varepsilon_{ij} \end{aligned} \qquad \begin{pmatrix} \beta_{0i}^{*} \\ \beta_{1i}^{*} \end{pmatrix} \sim G\left(\begin{pmatrix} \beta_0 + \beta_1 \overline{X}_{\cdot\cdot} \\ \beta_1 \end{pmatrix}, D^{*} \right)$$

$$D^* = \left(\begin{array}{c|c} Var\,\beta_{0_i} + Var\,\beta_{1_i} \cdot \bar{X}_{..}^2 - 2\,Cov(\beta_{0_i}, \beta_{1_i})\,\bar{X}_{..} & Cov(\beta_{0_i}^*, \beta_{1_i}^*) \\ & Var\,\beta_{1_i} \end{array} \right)$$

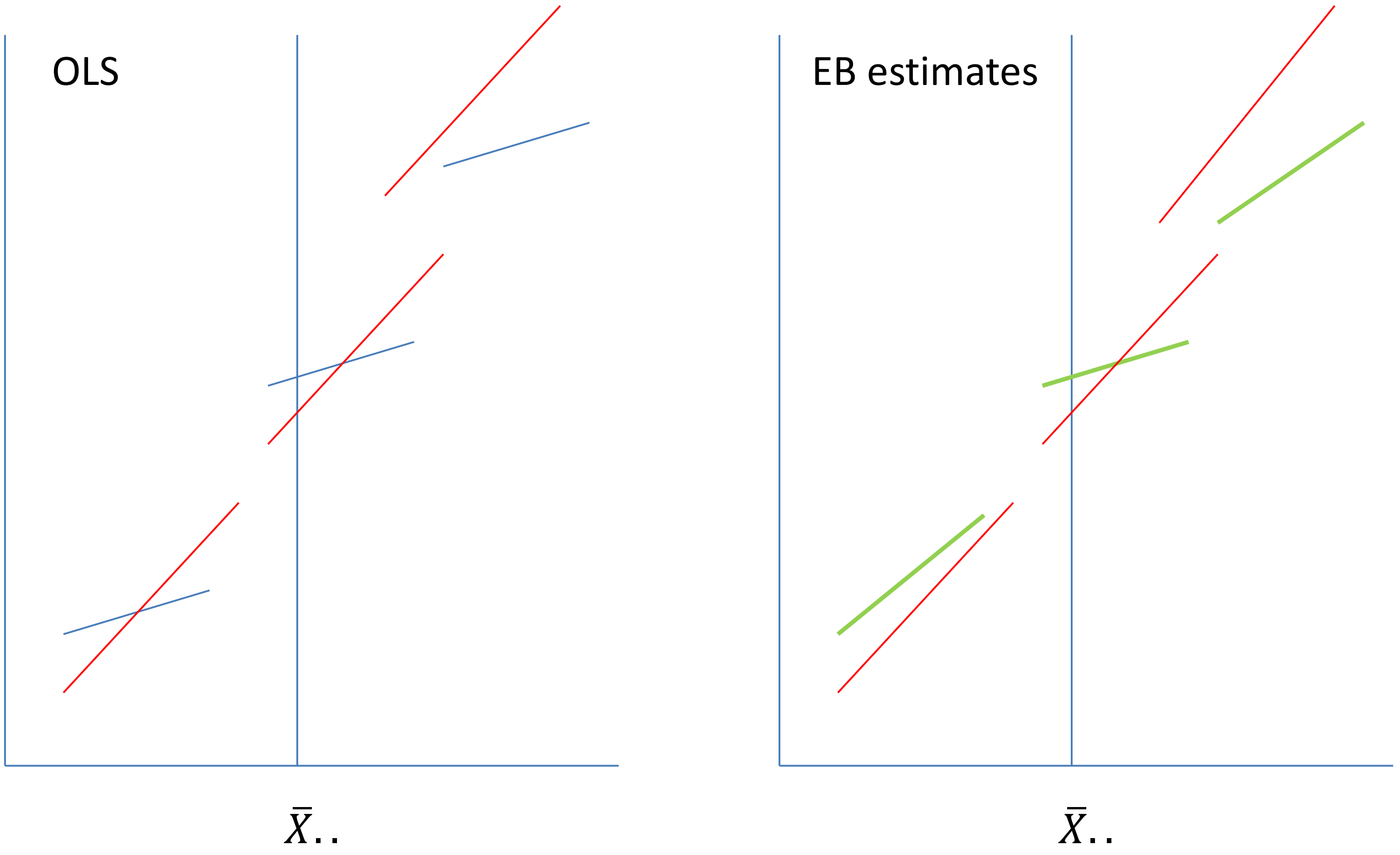
$$= Cov(\beta_{0_i} - \beta_{1_i}\bar{X}_{..}, \beta_{1_i}) = Cov(\beta_{0_i}, \beta_{1_i}) - \bar{X}_{..} Var\,\beta_{1_i}$$

$$= D_{12} - \bar{X}_{..} D_{22}$$

$$= D_{11} + D_{22} \cdot \bar{X}_{..}^2 - 2 D_{12} \bar{X}_{..}$$

$$\Rightarrow D^* = \left(\begin{array}{c|c} D_{11} + D_{22}\bar{X}_{..}^2 - 2 D_{12}\bar{X}_{..} & D_{12} - \bar{X}_{..} D_{22} \\ \hline & D_{22} \end{array} \right)$$

Graphical Representation



RED: Within cluster association (β_1)
BLUE: OLS school-specific slopes
GREEN: EB estimates

Center Within Cluster (by $\bar{X}_{i\cdot}$).

$$Y_{ij} = \beta_{0i} + \beta_{1i} X_{ij} + \varepsilon_{ij} \quad \begin{pmatrix} \beta_{0i} \\ \beta_{1i} \end{pmatrix} \sim G\left(\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, D \right).$$

as before

Now center by $\bar{X}_{i\cdot}$

$$Y_{ij} = \beta_{0i} + \beta_{1i}(X_{ij} - \bar{X}_{i\cdot}) + \varepsilon_{ij}$$

$$= (\beta_{0i} - \beta_{1i}\bar{X}_{i\cdot}) + \beta_{1i} X_{ij} + \varepsilon_{ij}$$

$$= \beta_{0i}^{*} + \beta_{1i}^{*} X_{ij} + \varepsilon_{ij}$$

where $\begin{pmatrix} \beta^*_{0i} \\ \beta^*_{1i} \end{pmatrix} \sim G\left(\begin{pmatrix} \beta_0 + \beta_1 \bar{X}_{i\cdot} \\ \beta_1 \end{pmatrix}, D^*_i \right)$

$$D^*_{11} = Var(\beta^*_{0i}) = Var(\beta_{0i} - \beta_{1i}\bar{X}_{i\cdot}) = D_{11} + D_{22}\bar{X}_{i\cdot} - 2 D_{12}\bar{X}_{i\cdot}$$

$$D^*_{22} = D_{22}$$

$$D^*_{12} = Cov(\beta^*_{0i}, \beta^*_{1i}) = Cov(\beta_{0i} - \beta_{1i}\bar{X}_{i\cdot}, \beta_{1i})$$
$$= D_{12} - D_{22}\bar{X}_{i\cdot}$$

Centering in Multi-level Models

- Choice of centering approach will change interpretation of level-1 intercept and can have an impact on interpretation of level-1 slopes!
- Preferred approach depends on the goal of the analysis
- Review 5 common purposes:
  - Estimating level-1 coefficient
  - Estimating the variances of level-1 coefficients
  - Disentangling person-level vs. contextual effects
  - Estimating level-2 effects while adjusting for level-1 covariates
  - Estimating an interaction between level-2 and level-1 covariates

Estimation of the Contextual Effect

$$Y_{ij} = \beta_{oi} + \beta_1\left(X_{ij} - \bar{X}_i.\right) + \varepsilon_{ij}$$
$$\beta_{oi} = \beta_0 + \beta_2\bar{X}_i. + b_{0i}$$

$$\beta_1 = \beta_w$$
$$\beta_2 = \beta_b$$
$$\beta_c = \beta_b - \beta_w = \beta_2 - \beta_1$$

$$Y_{ij} = \alpha_{oi} + \alpha_1\left(X_{ij} - \bar{X}..\right) + \epsilon_{ij}$$
$$\alpha_{oi} = \alpha_0 + \alpha_2\bar{X}_i. + a_{0i}$$

$$\alpha_1 = \beta_w$$
$$\alpha_2 = \beta_c$$
$$\beta_b = \beta_c + \beta_w = \alpha_2 + \alpha_1$$

We discussed this case in Lecture 4.

- These two models will each provide estimates of the within cluster, between cluster and contextual effects.

- And, now we know that the random intercept variance within the grand-mean centered model is likely smaller compared to the group-mean centered model.

So what do I do?

- Recommendation from Enders and Tofighi paper is to group-mean center
  - Easy to estimate the contextual effect
  - If goal is also to look at the random effect variances, then this approach will provide best estimate

Centering in Multi-level Models

- Choice of centering approach will change interpretation of level-1 intercept and can have an impact on interpretation of level-1 slopes!
- Preferred approach depends on the goal of the analysis
- Review 5 common purposes:
  - Estimating level-1 coefficient
  - Estimating the variances of level-1 coefficients
  - Disentangling person-level vs. contextual effects
  - Estimating level-2 effects while adjusting for level-1 covariates
  - Estimating an interaction between level-2 and level-1 covariates

Estimating level-2 effects while adjusting for level-1 covariates

- We considered this case in Lecture 4.
- Group-mean centered level-1 covariates are independent of level-2 covariates
- Therefore, in order to adjust for level-1 covariates the recommendation is
  – Don't center
  – Grand-mean center
  – OR include only the calculated level-2 covariates, this is really what you are accounting for in the model.

Centering in Multi-level Models

- Choice of centering approach will change interpretation of level-1 intercept and can have an impact on interpretation of level-1 slopes!
- Preferred approach depends on the goal of the analysis
- Review 5 common purposes:
  - Estimating level-1 coefficient
  - Estimating the variances of level-1 coefficients
  - Disentangling person-level vs. contextual effects
  - Estimating level-2 effects while adjusting for level-1 covariates
  - Estimating an interaction between level-2 and level-1 covariates

Estimating an interaction between level-2 and level-1 covariates

- Similar ideas here as we discussed in “estimating leve1-1 coefficient”
- The recommended approach is to group-mean center
- The grand-mean centering approach suffers from the same issues as estimation of level-1 covariate effect
  - The coefficients for the main effect and interaction term combine both within-cluster and between-cluster effects.

Application to HSB data

- Model with grand mean centered level-1 covariate:

$$Y_{ij} = \beta_{0i} + \beta_1\left(SES_{ij} - \overline{SES}_{..}\right) + \beta_2 C_i + \beta_3 C_i\left(SES_{ij} - \overline{SES}_{..}\right) + \varepsilon_{ij}$$

- Model with group-mean centered level-1 covariate:

$$Y_{ij} = \beta_{0i} + \beta_1\left(SES_{ij} - \overline{SES}_{i.}\right) + \beta_2 C_i + \beta_3 C_i\left(SES_{ij} - \overline{SES}_{i.}\right) + \varepsilon_{ij}$$

Results

Estimate (se)	Group-mean centered	Grand-mean centered	
Intercept: β_0	11.57 (0.29)	11.96 (0.23)	
SES: β_1	2.87 (0.15)	3.03 (0.14)	
Catholic: β_2	2.62 (0.44)	2.00 (0.34)	
Interaction: β_3	-1.44 (0.22)	-1.38 (0.21)	
RI SD: τ_1	2.54 (0.17)	1.88 (0.14)	
Residual SD: σ	6.07 (0.05)	6.07 (0.05)	

- Within Non-Catholic schools, the average math achievement score increases by 2.87 points per unit increase in SES.

- Among Catholic schools, the average math achievement score increases by 2.87-1.44 = 1.43 points per unit increase in SES.

Decomposing Total Effects into Within- and Between-Cluster Effects

- The grand-mean centered model can be modified so that you separately estimate the within- and between-cluster effects

$$Y_{ij} = \beta_{0i} + \beta_1\left(SES_{ij} - \overline{SES}_{..}\right) + \beta_2 C_i + \beta_3 C_i\left(SES_{ij} - \overline{SES}_{..}\right) + \beta_4 \overline{SES}_{i.} + \beta_5 \overline{SES}_{i.}\left(SES_{ij} - \overline{SES}_{..}\right) + \varepsilon_{ij}$$

- Idea: Add in the cluster mean as a main effect and interaction term with the grand-mean centered level-1 covariate.
- This main effect and interaction effect will now represent the within-effects!

Results

Estimate	Group-mean centered	Grand-mean centered	Grand-mean centered + between-cluster effects
Intercept: β_0	11.57 (0.29)	11.96 (0.23)	12.27 (0.20)
SES: β_1	2.87 (0.15)	3.03 (0.14)	2.87 (0.15)
Catholic: β_2	2.62 (0.44)	2.00 (0.34)	1.20 (0.30)
Interaction: β_3	-1.44 (0.22)	-1.38 (0.21)	-1.44 (0.22)
$\overline{SES}_{i.}$:β_4			2.88 (0.55)
$\overline{SES}_{i.}$ Interaction:β_5			0.43 (0.78)
RI SD: τ_1	2.54 (0.17)	1.88 (0.14)	1.52 (0.12)
Residual SD: σ	6.07 (0.05)	6.07 (0.05)	6.07 (0.05

Summary of Centering within MLMs

1. Estimating fixed level-1 covariates
   – Group-mean center
2. Estimating the variances of level-1 covariates
   – Group-mean center
3. Disentangling person-level vs. contextual effects
   – Either can be used to estimate the contextual effect, but group-mean centering will benefit from more accurate estimates of random effect variances
4. Estimating level-2 effects while adjusting for level-1 covariates
   – Grand-mean center
5. Estimating an interaction between a level-1 and level-2 covariate
   – Group-mean center
   – Grand-mean center but add in adjustment for between-cluster effects

In-Class Exercise

- Consider a survey conducted on a random sample of Type II diabetes patients from a random sample of JHH physicians
- Variables:
  - Patient satisfaction with care (continuous score, outcome)
  - Patient age
  - Patient gender
  - Physician years of experience
  - Physician gender

What is the physician-specific association between patient satisfaction and age, adjusting for gender?

I. Centering Level 1 Predictors to Estimate Level 1 effect (β_1)

Models

I.A: $Y_{ij} = \beta_{0i} + \beta_1 X_{ij} + \varepsilon_{ij}$

I.B. $Y_{ij} = \beta_{0i} + \beta_1 (X_{ij} - \bar{X}_{..}) + \varepsilon_{ij}$

I.C. $Y_{ij} = \beta_{0i} + \beta_1 (X_{ij} - \bar{X}_{i.}) + \varepsilon_{ij}$

(I.A and I.B: same; I.B and I.C: difference)

$$\hat\beta_1(\text{I.A}) \equiv \hat\beta_1(\text{I.B}) = \hat\beta_1(\text{I.C}) + \underbrace{\text{deviation due to relationship of } \bar{Y}_{i.} \text{ with } \bar{X}_{i.}}_{\hat\beta_1(\text{level 2 info})}$$

β

Type II Diabetics at JHHS

$Y_{ij} \equiv$ Patient$_j$ satisfaction with care by physician i

Age_{ij} - patient age

G_{ij} " gender

E_i - physician's experience

PG_i " , gender

I.A: $Y_{ij} = \beta_{0i} + \beta_1 (Age_{ij} * G_{ij}) + \varepsilon_{ij}$

I.A eg. $Y_{ij} = \beta_{0i} + \beta_1 Age_{ij} + \varepsilon_{ij}$

I.C: $Y_{ij} = \beta_{0i} + \beta_1^* (Age_{ij} - \overline{Age}_{i.}) + \varepsilon_{ij}$

$$\hat\beta_1 \neq \hat\beta_1^* \; ; \quad \hat\beta_1 = \lambda_i \hat\beta_1^* + (1-\lambda_i)\, \hat\beta_{1\,\text{across physicians}}$$

$$\lambda = \frac{1/\mathrm{Var}(\hat\beta_1^*)}{\left(1/\mathrm{Var}\,\hat\beta_1^* + 1/\mathrm{Var}\,\hat\beta_{1\,\text{across}}\right)}$$

$$\widehat{\mathrm{Var}}(b_{0i})_{\text{I.C}} \;\boxed{\langle \text{or} \rangle \; ?}\; \widehat{\mathrm{Var}}(b_{0i})_{\text{I.A (B)}}$$

What is the contextual effect of patient age?

I. Centering Level 1 Predictors to Estimate Level 1 effect (β_1)

Models

I.A: $Y_{ij} = \beta_{0i} + \beta_1 X_{ij} + \varepsilon_{ij}$

I.B. $Y_{ij} = \beta_{0i} + \beta_1 (X_{ij} - \bar{X}_{..}) + \varepsilon_{ij}$

I.C. $Y_{ij} = \beta_{0i} + \beta_1 (X_{ij} - \bar{X}_{i.}) + \varepsilon_{ij}$

(I.A and I.B: same difference)

$$\hat{\beta}_1(I.A) \equiv \hat{\beta}_1(I.B) = \hat{\beta}_1(I.C) + \underbrace{\text{deviation due to relationship of } \bar{Y}_{i.} \text{ with } \bar{X}_{i.}}_{\hat{\beta}_1(\text{level 2 info})}$$

β

Type II Diabetics at JHHS

$Y_{ij} \equiv$ Patient j satisfaction with care by physician i

Age_{ij} - patient age

G_{ij} " gender

E_i - physician's experience

PG_i " ' gender

$$I.A: \quad Y_{ij} = \beta_{0i} + \beta_1 (Age_{ij} * G_{ij}) + \varepsilon_{ij}$$

$$\boxed{\begin{aligned} &I.A \text{ eg.} \quad Y_{ij} = \beta_{0i} + \beta_1 Age_{ij} + \varepsilon_{ij} \\ &I.C: \quad Y_{ij} = \beta_{0i} + \beta_1^{*} (Age_{ij} - \overline{Age}_{i}) + \varepsilon_{ij} \end{aligned}}$$

$$\hat{\beta}_1 \neq \hat{\beta}_1^{*} \; ; \quad \hat{\beta}_1 = \lambda_i \hat{\beta}_1^{*} + (1-\lambda_i)\, \hat{\beta}_1 \text{ across physicians}$$

$$\lambda = \frac{1/\mathrm{var}(\hat{\beta}_1^{*})}{\left(1/\mathrm{Var}\,\hat{\beta}_1^{*} + 1/\mathrm{Var}\,\hat{\beta}_1 \text{ across}\right)}$$

$$\widehat{\mathrm{Var}}(b_{0i})_{I.C} \; \boxed{\begin{matrix} < \text{or} > \\ ? \end{matrix}} \; \widehat{\mathrm{Var}}(b_{0i})_{I.A\,(B)}$$

Describe the the association between patient satisfaction and physician gender

IV. What is the association between patient satisfaction and physician gender?

IV.A $Y_{ij} = \beta_{0i} + \beta_1 PG_i + \varepsilon_{ij}$

IV.B $Y_{ij} = \beta^*_{0i} + \beta_1^* (PG_i - \overline{PG}_{\cdot}) + \varepsilon_{ij}$

same except for β_0

$\beta_0^* = \beta_0 - \beta_1 \overline{PG}_{\cdot}$; $\beta_1^* = \beta_1$

What is the association between patient satisfaction and physician gender adjusting for patient age and gender

V. What is the association between patient satisfaction and physician gender among patients of the same age and (patient) gender

V. A. $Y_{ij} = \beta_{0i} + \beta_2 PG_i + \beta_2 Age_{ij} + \beta_3 G_{ij} + \varepsilon_{ij}$

V. B: $Y_{ij} = \beta^*_{0i} + \beta^*_1(PG_i - \overline{PG}_{\cdot}) + \beta^*_2(Age_{ij} - \overline{Age}_{i}) + \beta^*_3(G_{ij} - \overline{G}_{i\cdot})$

A. Same age in years
B. Same age difference from physician's mean patient a

Write out the model to estimate the association between patient satisfaction and patient age, separately for male and female physicians

VI. Is the relationship between patient satisfaction and patient age the same for male and female doctors?

VI. A. $Y_{ij} = \beta_{0i} + \beta_1 Age_{ij} + \beta_2 PG_i + \beta_3 (Age_{ij} * PG_i) +$

VI. B $Y_{ij} = \beta_{0i}^* + \beta_1^* (Age_{ij} - \overline{Age_{i\cdot}}) + \beta_2^* PG_i + \beta_3^* (Age_{ij} - \overline{Age_{i}}$ $+ \varepsilon_{ij}$

$\beta_1 \neq \beta_1^*$ $\beta_3 \neq \beta_3^*$