# Fixed Effects: Why Centering Matters

- $y_i$ = Student achievement (GPA as percentage out of 100)
  $x_i$ = Parent **attitudes** about education (measured on 1–5 scale)
  $z_i$ = Parent **education** level (measured in years of education)

$$GPA_i = \beta_0 + \beta_1(Att_i) + \beta_2(Ed_i) + \beta_3(Att_i)(Ed_i) + e_i$$
$$GPA_i = 30 + 1(Att_i) + 2(Ed_i) + 0.5(Att_i)(Ed_i) + e_i$$

- **Interpret $\beta_0$:**
- **Interpret $\beta_1$:**
- **Interpret $\beta_2$:**
- **Interpret $\beta_3$: Attitude as Moderator:**

  **Education as Moderator:**

- **Predicted GPA** for **attitude = 3** and **Ed = 12**?

  **75 = 30 + 1*(3) + 2*(12) + 0.5*(3)*(12)**

# How Centering Changes Fixed Effects

- $y_i$ = Student achievement (GPA as percentage out of 100)
  $x_i$ = Parent **attitudes** about education (now centered at **3**)
  $z_i$ = Parent years of **education** (now centered at **12**)

$$GPA_i = \beta_0 + \beta_1(Att_i - 3) + \beta_2(Ed_i - 12) + \beta_3(Att_i - 3)(Ed_i - 12) + e_i$$
$$GPA_i = 75 + 7(Att_i - 3) + 3.5(Ed_i - 12) + 0.5(Att_i - 3)(Ed_i - 12) + e_i$$

- **Interpret $\beta_0$:**
- **Interpret $\beta_1$:**
- **Interpret $\beta_2$:**
- **Interpret $\beta_3$: Attitude as Moderator:**

  **Education as Moderator:**

- But how did I know what the new fixed effects would be???

# Model-Implied Predictor Simple Slopes

- Example equation for predicted GPA using centered predictors:
  $\widehat{GPA}_i = \beta_0 + \beta_1(Att_i - 3) + \beta_2(Ed_i - 12) + \beta_3(Att_i - 3)(Ed_i - 12)$

- This model equation provides predictions for:
  - Expected outcome given any combination of predictor values
  - Any conditional (simple) main effect slopes implied by interaction term
  - **Any slope can be found as: what it is + what *modifies* it**

- Three steps to get any model-implied simple main effect slope:

1. **Identify** all terms in model involving the predictor of interest
2. **Factor out** common predictor variable to find slope linear combination
3. **Calculate** estimate and SE for slope linear combination
   - *By "calculate" I of course mean "ask a program to do this for you"*

# Model-Implied Predictor Simple Slopes

- Example equation for predicted GPA using centered predictors:
  $\widehat{GPA}_i = \beta_0 + \beta_1(Att_i - 3) + \beta_2(Ed_i - 12) + \beta_3(Att_i - 3)(Ed_i - 12)$

1. **Identify** all slopes in model involving the predictor of interest

   To get attitudes slope: $Est = \beta_1(Att_i - 3) + \beta_3(Att_i - 3)(Ed_i - 12)$

   To get education slope: $Est = \beta_2(Ed_i - 12) + \beta_3(Att_i - 3)(Ed_i - 12)$

2. **Factor out** predictor of interest to find slope linear combination

   To get attitudes slope: $Est = [\beta_1 + \beta_3 (Ed_i - 12)]$ **that will multiply** $(Att_i - 3)$

   To get education slope: $Est = [\beta_2 + \beta_3(Att_i - 3)]$ **that will multiply** $(Ed_i - 12)$

- Btw, the SEs for the new slopes provided by the program come from:
  - $SE^2$ = sampling variance of slope estimate → e.g., $Var(\beta_1) = SE^2_{\beta_1}$

  attitudes slope: $SE^2 = Var(\beta_1) + Var(\beta_3)(Ed_i - 12) + 2Cov(\beta_1, \beta_3)(Ed_i - 12)$

  education slope: $SE^2 = Var(\beta_2) + Var(\beta_3)(Att_i - 3) + 2Cov(\beta_2, \beta_3)(Att_i - 3)$

# Model-Implied Predictor Simple Slopes

- **To request <u>predicted simple slopes</u> (= simple main effects):**
  - **DO NOT include the intercept** ($\beta_0$ does **not** contribute to slopes)
  - **Include ONLY** the fixed effects that contain the predictor of interest

$$\widehat{GPA}_i = \beta_0 + \beta_1(Att_i - 3) + \beta_2(Ed_i - 12) + \beta_3(Att_i - 3)(Ed_i - 12)$$

→ attitudes slope: $Est = [\beta_1 + \beta_3(Ed_i - 12)]$ that multiplies $(Att_i - 3)$

→ education slope: $Est = [\beta_2 + \beta_3(Att_i - 3)]$ that multiplies $(Ed_i - 12)$

**STATA: Each line starts with lincom, label at end of line after // comment**

```
_cons*0 + att*__  + ed*__  + att#ed*__  // Att Slope if Ed=10
_cons*0 + att*__  + ed*__  + att#ed*__  // Att Slope if Ed=18
_cons*0 + att*__  + ed*__  + att#ed*__  // Ed Slope  if Att=2
_cons*0 + att*__  + ed*__  + att#ed*__  // Ed Slope  if Att=5
```

**R: Coefficients are multipliers in GLHT entered in order of fixed effects**

```
"Att Slope if Ed=10" = c(0,__,__,__),
"Att Slope if Ed=18" = c(0,__,__,__),
"Ed Slope  if Att=2" = c(0,__,__,__),
"Ed Slope  if Att=5" = c(0,__,__,__)
```

# Regions of Significance for Simple Slopes

- For quantitative predictors, there may not be specific values of the moderator at which you want to know the slope's significance…

- For example, with age*woman (in which 0=man, 1=woman here):

$$\hat{y}_i = \beta_0 + \beta_1(Age_i - 85) + \beta_2(Woman_i) + \beta_3(Age_i - 85)(Woman_i)$$

→ **age slope:** $Est =$ ________________ that multiplies $(Age_i - 85)$

→ **gender slope:** $Est =$ ________________ that multiplies $(Woman_i)$

- Age slopes are only relevant for two specific values of binary *woman*:

```
_cons*0 + age85*_ + woman*_ + age85*woman*_ // Age Slope for Men
_cons*0 + age85*_ + woman*_ + age85*woman*_ // Age Slope for Women
```

- But there are many ages to request gender differences for…

```
_cons*0 + age85*_ + woman*_ + age85*woman*_ //Gender Diff at Age=80
_cons*0 + age85*_ + woman*_ + age85*woman*_ //Gender Diff at Age=90
```

# Regions of Significance for Simple Slopes

- An alternative approach for continuous moderators is known as **regions of significance** (see Hoffman 2015 chapter 2 for refs)
- Rather than asking if the simple main effect of gender is still significant at a particular age, we can find the **boundary ages** at which the gender effect becomes non-significant
- We know that: EST / SE = *t*-value → if |*t*| > |1.96|, then $p < .05$
- So we work backwards to find the EST and SE such that:

$$\pm t = \pm 1.96 = \frac{\text{Slope Estimate}}{\sqrt{\text{Variance of Slope Estimate}}}, \text{ where:}$$

$$\text{Gender Slope (Gender Difference) Estimate} = \beta_2 + \beta_3(\text{Age} - 85)$$

$$\text{Variance of Slope Estimate} = \text{Var}(\beta_2) + \boxed{2\text{Cov}(\beta_2\beta_3)(\text{Age} - 85)} + \text{Var}(\beta_3)(\text{Age} - 85)^2$$

- Need to request "asymptotic covariance matrix" (COVB)
  - Covariance matrix of fixed effect estimates ($SE^2$ on diagonal)

# Regions of Significance for Simple Slopes

$$\pm t = \pm 1.96 = \frac{\text{Slope Estimate}}{\sqrt{\text{Variance of Slope Estimate}}}, \text{ where:}$$

$$\text{Gender Slope (Gender Difference) Estimate} = \beta_2 + \beta_3(\text{Age} - 85)$$

$$\text{Variance of Slope Estimate} = \text{Var}(\beta_2) + 2\text{Cov}(\beta_2\beta_3)(\text{Age} - 85) + \text{Var}(\beta_3)(\text{Age} - 85)^2$$

- For example, age*woman (0=man, 1=woman), age = moderator:

$$\hat{y}_i = \beta_0 + \beta_1(Age_i - 85) + \beta_2(Woman_i) + \beta_3(Age_i - 85)(Woman_i)$$

- **$\beta_2 = -0.5306$* at age=85, $Var(\beta_2)$ → $SE^2$ for $\beta_2$ was $0.06008$**
- **$\beta_3 = -0.1104$* unconditionally, $Var(\beta_3)$ → $SE^2$ for $\beta_3$ was $0.00178$**
- **Covariance of $\beta_2$ $SE$ and $\beta_3$ $SE$ was $0.00111$**

- **Regions of Significance for Moderator of Age = 60.16 to 79.52**
  - The gender effect $\beta_2$ is predicted to be significantly negative above age 79.52, non-significant from ages 79.52 to 60.16, and significantly positive below age 60.16 (because non-parallel lines will cross eventually).