# Example 3: Fixed Slopes of Level-1 Predictors in General Multilevel Models for Two-Level Nested Outcomes

***(complete syntax, data, and output available for STATA, R, and SAS electronically)***

The data for this example come from the High School and Beyond dataset (HSB4) used in McNeish (2023). Using 7,185 students from 160 schools, we will be examining the extent to which student math can be predicted from student-level variables of SES and white versus nonwhite. Note that this example computes total-$R^2$ and pseudo-$R^2$ in SAS using two custom macros (available in the SAS syntax file online) as well as in R using two custom functions and a general package (available in the R syntax file and function files online).

## STATA Syntax for Importing and Preparing Data for Analysis:

```
// Define global variable for file location to be replaced in code below
// \\Client\ precedes path in Virtual Desktop outside H drive
   global filesave "C:\Dropbox\23_PSQF6272\PSQF6272_Example3"

// Open trimmed example excel data file from sheet "HSB4" and clear away any existing data
   clear // clear memory in case of open data
   import excel "$filesave\Example3_Data.xlsx", firstrow case(preserve) sheet("HSB4") clear

// Filter to only cases complete on all variables to be used below (before cluster means)
   egen nmiss=rowmiss(math ses nonwhite)
   drop if nmiss>0

// Compute cluster means for level-1 variables
   sort schoolID
   egen CMmath = mean(math),     by(schoolID)
   egen CMses  = mean(ses),      by(schoolID)
   egen CMnon  = mean(nonwhite), by(schoolID)

display "STATA Descriptive Statistics for Cluster Mean Level-2 Variables"
summarize CMmath CMses CMnon
pwcorr    CMmath CMses CMnon, sig
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      CMmath |      7,185    12.74785    3.005817   4.239781   19.71914
       CMses |      7,185    .0001434    .4135432  -1.193946   .8249825
       CMnon |      7,185     .274739    .3042263          0          1

             |   CMmath    CMses    CMnon
-------------+---------------------------
      CMmath |   1.0000
             |
             |
       CMses |   0.7865   1.0000
             |   0.0000
             |
       CMnon |  -0.4852  -0.4647   1.0000
             |   0.0000   0.0000
             |
```

Using a multivariate MLM to get these correlations among the 3 variables' L2 random intercepts (i.e., a latent mean rather than an observed mean) would pry be better—stay tuned for how to do so!

```
display "STATA Descriptive Statistics for Original Level-1 Variables"
summarize math ses nonwhite
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        math |      7,185    12.74785    6.878246     -2.832     24.993
         ses |      7,185    .0001434    .7793552     -3.758      2.692
    nonwhite |      7,185     .274739    .4464137          0          1
```

```
// Center L2 predictor at constant so 0 is meaningful
   gen CMnon30 = CMnon - .30
// 0 is otherwise already meaningful in other predictors

// Center L1 predictors at cluster means
   gen WCses = ses - CMses
   gen WCnon = nonwhite - CMnon
// Center outcome at cluster mean for correlations
   gen WCmath = math - CMmath

display "STATA Descriptive Statistics for Within-Cluster Level-1 Variables"
summarize WCmath WCses WCnon
pwcorr    WCmath WCses WCnon, sig

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      WCmath |      7,185   -3.77e-08    6.186706  -19.67464   17.96889
       WCses |      7,185   -1.62e-10     .660588  -3.650741   2.856078
       WCnon |      7,185   -1.79e-10     .326698  -.9824561   .9824561

             |   WCmath    WCses    WCnon
-------------+---------------------------
      WCmath |   1.0000
             |
             |
       WCses |   0.2340   1.0000
             |   0.0000
             |
       WCnon |  -0.1877  -0.1667   1.0000
             |   0.0000   0.0000
```

Using a multivariate MLM to get these correlations among the 3 variables' L1 residuals (i.e., a latent deviation rather than an observed deviation) could be better—stay tuned for how to do so!

## <u>R</u> Syntax for Importing and Preparing Data for Analysis (after loading packages *readxl, TeachingDemos, Hmisc, psych, lme4, lmerTest, performance, nlme,* and *psychometric*):

```
# Define variables for working directory and data name -- CHANGE THESE
filesave = "C:\\Dropbox/23_PSQF6272/PSQF6272_Example3/"
filename = "Example3_Data.xlsx"
setwd(dir=filesave)

# Load Jonathan's custom R functions from folder within working directory
functions = paste0("R functions/",dir("R functions/"))
temp = lapply(X=functions, FUN=source)

#  Import Example 3 excel data file from sheet "HSB4"
Example3 = read_excel(paste0(filesave,filename), sheet="HSB4")
# Convert to data frame to use in analysis
Example3 = as.data.frame(Example3)

# Filter to only cases complete on all variables to be used below (before cluster means)
Example3 = Example3[complete.cases(Example3[ , c("math","ses","nonwhite")]),]

# Compute cluster means for level-1 variables using Jonathan's function
Example3 = addUnitMeans(data=Example3, unitVariable="schoolID",
                        meanVariables=c("math","ses","nonwhite"),
                        newNames=c("CMmath","CMses","CMnon"))

print("R Descriptive Statistics for Cluster Mean Level-2 Variables")
describe(x=Example3[ , c("CMmath","CMses","CMnon")])

print("R Pearson Correlation Matrix with P-values using rcorr from Hmisc package")
print("Must convert data frame to matrix to use rcorr")
rcorr(x=as.matrix(Example3[,c("CMmath","CMses","CMnon")]), type="pearson")

print("R Descriptive Statistics for Original Level-1 Variables")
describe(x=Example3[ , c("math","ses","nonwhite")])
```

```
# Center L2 predictor at constant so 0 is meaningful
Example3$CMnon30 = Example3$CMnon - .30
# 0 is otherwise already meaningful in other predictors
# Center L1 predictors at cluster means
Example3$WCses = Example3$ses - Example3$CMses
Example3$WCnon = Example3$nonwhite - Example3$CMnon
# Center outcome at cluster mean for correlations
Example3$WCmath = Example3$math - Example3$CMmath;

print("R Descriptive Statistics for Within-Cluster Level-1 Variables")
describe(x=Example3[ , c("WCmath","WCses","WCnon")])

print("R Pearson Correlation Matrix with P-values using rcorr from Hmisc package")
rcorr(x=as.matrix(Example3[,c("WCmath","WCses","WCnon")]), type="pearson")
```

---

## Model 1: Empty Means, Random Intercept Model for the Math Outcome

Level 1: $Math_{pc} = \beta_{0c} + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + U_{0c}$

```
print("R Model 1: Empty Means, Random Intercept for Math")
Model1 = lmer(data=Example3, REML=TRUE, formula=math~1+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model1, chkREML=FALSE); summary(Model1, ddf="Satterthwaite")
print("Show intraclass correlation and its LRT")
icc(Model1); ranova(Model1)

display "STATA Model 1: Empty Means, Random Intercept for Math"
mixed math , || schoolID: , reml dfmethod(satterthwaite) dftable(pvalue) nolog
------------------------------------------------------------------
        math |      Coef.   Std. Err.           DF        t    P>|t|
-------------+----------------------------------------------------
       _cons |   12.63697   .2443943         158.8    51.71    0.000 gamma00
------------------------------------------------------------------
```

The fixed intercept is the grand mean of the school means, which will differ from the overall grand mean (as given by the model without a random intercept variance) whenever level-2 units have different level-1 sizes (i.e., are "unbalanced").

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
schoolID: Identity           |
                  var(_cons) |   8.614081   1.078813      6.739162    11.01062 Var(U_0c)
-----------------------------+------------------------------------------------
               var(Residual) |   39.14832   .6606445      37.87466    40.46481 Var(e_pc)
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 986.12        Prob >= chibar2 = 0.0000
```

The **`chibar2`** test above is a likelihood ratio (LR) test comparing this model to a single-level regression (without a random intercept, as **`linear model`**) using a chi-square ($\chi^2$) distribution with a mixture of DF=0 (for which $\chi^2 = 0$ always) and DF=1. Consequently, in this case you can obtain the mixture *p*-value by weighting each contribution to the $\chi^2$ by 0.5, which means cutting the regular *p*-value in half. **Here, this LRT is a significance test of the intraclass correlation (ICC), which in turn provides an effect size for the amount of constant dependency attributed to school mean differences in math.**

```
display "-2LL = " e(ll)*-2      // Print -2LL for model
-2LL = 47116.793
```

Likelihood Ratio Test (LRT) Statistic: $= 48{,}102.917 - 47{,}116.793 = 986.12$

```
estat icc  // Intraclass correlation
Intraclass correlation
```

$$\text{ICC} = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{8.614}{8.614 + 39.148} = \mathbf{.180}$$

```
------------------------------------------------------------------------------
                       Level |        ICC   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
                    schoolID |   .1803528   .0187219      .1465168    .2199886
------------------------------------------------------------------------------
```

## Empty Models for the Level-1 Student Predictors: Quantitative SES and Binary Nonwhite

```
display "STATA: Empty Means, Random Intercept for SES"
mixed ses , || schoolID: , reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2      // Print -2LL for model
estat icc                       // Intraclass correlation

print("R: Empty Means, Random Intercept for SES")
EmptySES = lmer(data=Example3, REML=TRUE, formula=ses~1+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(EmptySES, chkREML=FALSE); summary(EmptySES, ddf="Satterthwaite")

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept) 0.16082  0.40102 Var(U_0)
 Residual             0.44624  0.66802 Var(e_pc)
Number of obs: 7185, groups:  schoolID, 160

Fixed effects:
              Estimate  Std. Error          df t value Pr(>|t|)
(Intercept)  -0.0056771   0.0327494 159.1872578 -0.1733   0.8626

print("Show intraclass correlation and its LRT")
icc(EmptySES); ranova(EmptySES)

# Intraclass Correlation Coefficient
    Adjusted ICC: 0.265
  Unadjusted ICC: 0.265

ANOVA-like table for random-effects: Single term deletions
               npar   logLik     AIC    LRT Df Pr(>Chisq)
<none>            3 -7523.01 15052.0
(1 | schoolID)    2 -8407.21 16818.4 1768.4  1 < 2.22e-16


print("R: Empty Means, Random Intercept for Binary Nonwhite")
EmptyNon = glmer(data=Example3, family=binomial(link="logit"), nonwhite~1+(1|schoolID))
summary(EmptyNon) # deviance = -2LL already

print("Compute ICC using pi^2/3 = 3.29 as L1 residual variance")
icc(EmptyNon)

print("R: Single-level empty model predicting observed free/reduced lunch ignoring school")
SingleNon = glm(data=Example3, family=binomial(link="logit"), formula=nonwhite~1)
print("Likelihood Ratio Test for Addition of Random Intercept Variance")
DevTest=-2*(logLik(SingleNon)-logLik(EmptyNon))
Pvalue=pchisq((DevTest), df=1, lower.tail=FALSE)
print("Test Statistic and P-values for DF=1")
DevTest; Pvalue

display "STATA: Empty Means, Random Intercept for Binary Nonwhite"
melogit nonwhite , || schoolID: , nolog

Integration method: mvaghermite                 Integration pts.  =          7
                                                Wald chi2(0)      =          .
Log likelihood = -2741.2796                     Prob > chi2       =          .
------------------------------------------------------------------------------
    nonwhite |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |  -1.627512   .1832885    -8.88   0.000    -1.986751   -1.268273
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|   5.536585   .7849087                      4.193427    7.309956
------------------------------------------------------------------------------
LR test vs. logistic model: chibar2(01) = 2965.79     Prob >= chibar2 = 0.0000
```

```
display "-2LL = " e(ll)*-2       // Print -2LL for model
-2LL = 5482.5592

estat icc                        // Intraclass correlation
------------------------------------------------------------------------------
                       Level |        ICC    Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
                    schoolID |   .6272718    .0331455      .5603717    .6896299
------------------------------------------------------------------------------
```

---

## Model 2a: Add Level-1 Cluster-Mean-Centered Student SES

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}\left(SES_{pc} - \overline{SES}_c\right) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + U_{0c}$

$\beta_{1c} = \gamma_{10}$

Cluster-mean-centering removes all L2 mean differences from the L1 predictor, such that **WCses** contains within-school person-to-person variation only.

```
display "STATA Model 2a: Add L1 Cluster-MC Student SES"
mixed math c.WCses, || schoolID: , reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2       // Print -2LL for model

print("R Model 2a: Add L1 Cluster-MC Student SES")
Model2a = lmer(data=Example3, REML=TRUE, formula=math~1+WCses+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model2a, chkREML=FALSE); summary(Model2a, ddf="Satterthwaite")

'log Lik.' -23361.998 (df=4) → LL for model (with 5 parameters)

       AIC         BIC      logLik    deviance    df.resid
 46731.996   46759.515  -23361.998   46723.996    7181.000 → Deviance = −2LL for model

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept)  8.6723  2.9449   Var(U_0c)
 Residual             37.0104  6.0836   Var(e_pc)

Fixed effects:
              Estimate Std. Error         df t value  Pr(>|t|)
(Intercept)   12.63614    0.24449  156.74052  51.684 < 2.2e-16   gamma00
WCses          2.19117    0.10865 7022.02373  20.166 < 2.2e-16   gamma10
```

Intercept $\gamma_{00}$ =

WCses $\gamma_{10}$ =

```
print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model2a)

R2 Random.(Intercept)          R2 L1.sigma2
        -0.0067638737          0.0546104890
```

### Pseudo-R2 Relative to CovEmpty (from SAS)

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 |
|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . |
| CovSESWithin | UN(1,1) | schoolID | 8.6676 | 1.0784 | -0.006725 |
| CovSESWithin | Residual | | 37.0108 | 0.6246 | 0.054610 |

**Pseudo-$R^2$ Results:** The fixed slope of our L1 **WCses** predictor accounted for 5.5% of the L1 residual variance and NONE of the L2 random intercept variance—this is because **WCses** contains only L1 info! Consequently, the reduction in L1 $\sigma_e^2$ made L2 $\tau_{U_0}^2$ increase (= negative pseudo-$R_{U_0}^2$).

## Model 2b: Add Level-2 Cluster Mean SES to Level-1 Cluster-Mean-Centered Student SES

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}\left(SES_{pc} - \overline{SES}_c\right) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + \gamma_{01}\left(\overline{SES}_c - 0\right) + U_{0c}$

$\beta_{1c} = \gamma_{10}$

Because SES already had a mean ~0, the level-2 cluster mean SES predictor is left uncentered (i.e., centered at a constant of 0).

```
display "STATA Model 2b: L1 Cluster-MC Student SES + L2 Cluster Mean SES"
mixed math c.WCses c.CMses, || schoolID: , ///
           reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2      // Print -2LL for model
lincom c.WCses*-1 + c.CMses*1   // SES L2 Contextual Slope
predict predSES                 // Save fixed-effect predicted outcomes for total-R2
```

```
print("R Model 2b: Add L1 Cluster-MC Student SES")
Model2b = lmer(data=Example3, REML=TRUE, formula=math~1+WCses+CMses+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model2b, chkREML=FALSE); summary(Model2b, ddf="Satterthwaite")
```

```
       AIC        BIC     logLik   deviance   df.resid
 46578.584  46612.983 -23284.292  46568.584   7180.000

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept)  2.6925  1.6409   Var(U_0c)
 Residual             37.0191  6.0843   Var(e_pc)

Fixed effects:
              Estimate Std. Error         df t value  Pr(>|t|)
(Intercept)   12.68331    0.14938  153.65182  84.906 < 2.2e-16   gamma00
WCses          2.19117    0.10867 7021.50918  20.164 < 2.2e-16   gamma10
CMses          5.86617    0.36170  153.36659  16.218 < 2.2e-16   gamma01
```

Intercept $\gamma_{00}$ =

WCses $\gamma_{10}$ =

CMses $\gamma_{01}$ =

```
print("SES L2 Contextual Slope"); contest1D(Model2b, L=c(0,-1,1))
   Estimate Std. Error        df   t value      Pr(>|t|)
1 3.6750018  0.3776705 182.28632 9.7307094 2.7568765e-18  gamma01 - gamma10
```

```
print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model2b)
R2 Random.(Intercept)           R2 L1.sigma2
          0.687424952            0.054389572
```

**Pseudo-R2 Relative to CovEmpty (from SAS)**
**Change in Pseudo-R2 for CovSESwithin vs. CovSESbetween**

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 | PseudoR2Change |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . | . |
| CovSESwithin | UN(1,1) | schoolID | 8.6676 | 1.0784 | -0.00672 | . |
| CovSESwithin | Residual | | 37.0108 | 0.6246 | 0.05461 | . |
| CovSESbetween | UN(1,1) | schoolID | 2.6887 | 0.4043 | **0.68772** | 0.69444 |
| CovSESbetween | Residual | | 37.0200 | 0.6248 | 0.05438 | -0.00023 |

## Model 3a: What If We Had Just Used Grand-Mean-Centered Level-1 Student SES?

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}\left(SES_{pc} - 0\right) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + U_{0c}$

$\beta_{1c} = \gamma_{10}$

> The L1 predictor (centered at 0) still includes both L1 within-school and L2 between-school variance…

```
display "STATA Model 3a: Use Smushed Main Effect of L1 Student SES Instead"
mixed math c.ses, || schoolID: , reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2       // Print -2LL for model

print("R Model 3a: Use Smushed Main Effect of L1 Student SES Instead")
Model3a = lmer(data=Example3, REML=TRUE, formula=math~1+ses+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model3a, chkREML=FALSE); summary(Model3a, ddf="Satterthwaite")
```

```
       AIC        BIC      logLik   deviance   df.resid
 46653.169  46680.688 -23322.585   46645.169   7181.000

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept)  4.7682  2.1836   Var(U_0c)
 Residual             37.0344  6.0856   Var(e_pc)

Fixed effects:
             Estimate Std. Error         df t value  Pr(>|t|)
(Intercept)  12.65748    0.18799  148.30225  67.332 < 2.2e-16   gamma00
ses           2.39020    0.10572 6838.07757  22.609 < 2.2e-16   gamma10
```

> **Fixed effects from Model 2b:**
> ```
> (Intercept)  12.68331
> WCses         2.19117
> CMses         5.86617
> ```

Intercept $\gamma_{00}$ =

ses $\gamma_{10}$ =

```
print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model3a)
```

```
R2 Random.(Intercept)          R2 L1.sigma2
          0.446463794           0.053997803
```

### Pseudo-R2 Relative to CovEmpty (from SAS)

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 |
|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . |
| CovSESsmush | UN(1,1) | schoolID | 4.7665 | 0.6549 | **0.44637** |
| CovSESsmush | Residual | | 37.0346 | 0.6254 | **0.05400** |

> **Pseudo-$R^2$ Results:** The fixed slope of our L1 **ses** predictor accounted for **5.4%** of the L1 residual variance (which was **5.5%** for **WCses** before). However, L1 **ses** also accounted for **44.6%** of the L2 random intercept variance—this is because L1 SES still contains L2 info! But this model assumes no contextual effect—that the WC and BC slopes are equal—and so it does not account for the correct amount of L2 variance: **69.4%**. **The L2 model is wrong!**

---

## Model 3b: Let's Fix It—Add Level-2 Cluster Mean SES to Grand-Mean-Centered Level-1 Student SES

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}\left(SES_{pc} - 0\right) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + \gamma_{01}(\overline{SES}_c - 0) + U_{0c}$

$\beta_{1c} = \gamma_{10}$

> The L1 predictor (centered at 0) still includes both L1 within-school and L2 between-school variance… but now we've controlled for the L2 between-school variance explicitly (i.e., **the model is unsmushed**)!

```
display "STATA Model 3b: L1 Grand-MC Student SES + L2 Cluster Mean SES"
mixed math c.ses c.CMses, || schoolID: , reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2       // Print -2LL for model
lincom c.ses*1 + c.CMses*1       // SES L2 Between Slope
```

```
print("R Model 3b: L1 Grand-MC Student SES + L2 Cluster Mean SES")
Model3b = lmer(data=Example3, REML=TRUE, formula=math~1+ses+CMses+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model3b, chkREML=FALSE); summary(Model3b, ddf="Satterthwaite")

       AIC        BIC     logLik   deviance   df.resid
 46578.584  46612.983 -23284.292  46568.584   7180.000

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept)  2.6925  1.6409   Var(U_0c)
 Residual             37.0191  6.0843   Var(e_pc)

Fixed effects:
             Estimate Std. Error         df t value  Pr(>|t|)
(Intercept)  12.68331    0.14938  153.65182 84.9062 < 2.2e-16  gamma00
ses           2.19117    0.10867 7021.50918 20.1640 < 2.2e-16  gamma10
CMses         3.67500    0.37767  182.28632  9.7307 < 2.2e-16  gamma10
```

**Fixed effects from equivalent Model 2b:**

```
(Intercept)  12.68331
WCses         2.19117
CMses         5.86617
```

Intercept $\gamma_{00}$ =

ses $\gamma_{10}$ =

CMses $\gamma_{01}$ =

```
print("SES L2 Between Slope"); contest1D(Model3b, L=c(0,1,1))
   Estimate Std. Error        df   t value      Pr(>|t|)
1 5.8661738 0.36169936 153.36659 16.218369 4.3919409e-35

print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model3b)
R2 Random.(Intercept)          R2 L1.sigma2
          0.687424952           0.054389572
```

### Pseudo-R2 Relative to CovEmpty (from SAS)
### Change in Pseudo-R2 for CovSESsmush vs. CovSEScontext

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 | PseudoR2Change |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . | . |
| CovSESsmush | UN(1,1) | schoolID | 4.7665 | 0.6549 | **0.44637** | . |
| CovSESsmush | Residual | | 37.0346 | 0.6254 | **0.05400** | . |
| CovSEScontext | UN(1,1) | schoolID | 2.6887 | 0.4043 | **0.68772** | **0.24134** |
| CovSEScontext | Residual | | 37.0200 | 0.6248 | **0.05438** | **0.00037** |

The L2 contextual slope accounted for another 24.1% of the L2 random intercept variance—model 3b is fully equivalent to model 2b that used cluster-mean-centered L1 SES instead (see below for variance progression from model 2a to 2b). Note that the L1 model was slightly misspecified when using the smushed effect, as shown by an increase in $R_e^2$ in model 3b).

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 | PseudoR2Change |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . | . |
| CovSESwithin | UN(1,1) | schoolID | 8.6676 | 1.0784 | **-0.00672** | . |
| CovSESwithin | Residual | | 37.0108 | 0.6246 | **0.05461** | **.** |
| CovSESbetween | UN(1,1) | schoolID | 2.6887 | 0.4043 | **0.68772** | **0.69444** |
| CovSESbetween | Residual | | 37.0200 | 0.6248 | **0.05438** | **-0.00023** |

## Model 4a: Add Binary Student Nonwhite Predictor (to Model 3b)

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}\left(SES_{pc} - 0\right) + \beta_{2c}\left(Nonwhite_{pc}\right) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + \gamma_{01}(\overline{SES}_c - 0) + U_{0c}$

$\beta_{1c} = \gamma_{10}$

$\beta_{2c} = \gamma_{20}$

```
display "STATA Model 4a: Add L1 Binary Student Nonwhite"
mixed math c.ses c.CMses c.nonwhite, || schoolID: , ///
           reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2      // Print -2LL for model
```

```
print("R Model 4a: Add L1 Binary Student Nonwhite")
Model4a = lmer(data=Example3, REML=TRUE, formula=math~1+ses+CMses+nonwhite+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model4a, chkREML=FALSE); summary(Model4a, ddf="Satterthwaite")
```

```
       AIC         BIC      logLik    deviance    df.resid
 46408.480   46449.759 -23198.240   46396.480    7179.000

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept)  2.6376  1.6241   Var(U_0c)
 Residual             36.1388  6.0116   Var(e_pc)

Fixed effects:
             Estimate Std. Error         df  t value             Pr(>|t|)
(Intercept)  13.42427    0.15802  191.56698  84.9550            < 2.2e-16  gamma00
ses           1.96746    0.10869 7063.56468  18.1021            < 2.2e-16  gamma10
CMses         2.95308    0.37755  186.91931   7.8216 0.0000000000003719  gamma01
nonwhite     -2.71378    0.20484 4400.81812 -13.2486            < 2.2e-16  gamma20
```

Intercept $\gamma_{00}$ =

ses $\gamma_{10}$ =

CMses $\gamma_{01}$ =

nonwhite $\gamma_{20}$ =

```
print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model4a)
```

```
R2 Random.(Intercept)           R2 L1.sigma2
          0.693806993             0.076875598
```

**Pseudo-R2 Relative to CovEmpty (from SAS)**
**Change in Pseudo-R2 for CovSESbetween vs. CovNWsmush**

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 | PseudoR2Change |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . | . |
| CovSESbetween | UN(1,1) | schoolID | 2.6887 | 0.4043 | 0.68772 | . |
| CovSESbetween | Residual | | 37.0200 | 0.6248 | 0.05438 | . |
| CovNWsmush | UN(1,1) | schoolID | 2.6340 | 0.3976 | **0.69406** | **0.006345** |
| CovNWsmush | Residual | | 36.1396 | 0.6100 | **0.07686** | **0.022488** |

## Model 4b: Let's Fix It—Add Level-2 Cluster Mean of Nonwhite to Grand-Mean-Centered Level-1 Student Nonwhite

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}(SES_{pc} - 0) + \beta_{2c}(Nonwhite_{pc}) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + \gamma_{01}(\overline{SES}_c - 0) + \gamma_{02}(\overline{Nonwhite}_c - .30) + U_{0c}$

$\beta_{1c} = \gamma_{10}$

$\beta_{2c} = \gamma_{20}$

The L2 Nonwhite predictor was centered at .30 to prevent weirdness in interpreting its L1 slope (which would be impossible at **CMnon** = 0).

```
display "STATA Model 4b: L1 Binary Student Nonwhite + L2 Cluster Mean Nonwhite"
mixed math c.ses c.CMses c.nonwhite c.CMnon30, || schoolID: , ///
           reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2         // Print -2LL for model
lincom c.ses*1 + c.CMses*1         // SES L2 Between Slope
lincom c.nonwhite*1 + c.CMnon30*1  // Nonwhite L2 Between Slope

print("R Model 4b: L1 Binary Student Nonwhite + L2 Cluster Mean Nonwhite")
Model4b = lmer(data=Example3, REML=TRUE,
               formula=math~1+ses+CMses+nonwhite+CMnon30+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model4b, chkREML=FALSE); summary(Model4b, ddf="Satterthwaite")

       AIC         BIC      logLik   deviance    df.resid
 46404.554   46452.712 -23195.277  46390.554    7178.000

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept)  2.5588   1.5996   Var(U_0c)
 Residual             36.1361   6.0113   Var(e_pc)

Fixed effects:
              Estimate Std. Error         df  t value           Pr(>|t|)
(Intercept)   13.51085    0.16102  219.69072  83.9060          < 2.2e-16  gamma00
ses            1.95248    0.10889 7020.17024  17.9313          < 2.2e-16  gamma10
CMses          3.37468    0.41713  172.76081   8.0903 0.0000000000001011  gamma01
nonwhite      -2.89558    0.22017 7020.17024 -13.1515          < 2.2e-16  gamma20
CMnon30        1.35124    0.59432  199.13638   2.2736            0.02406  gamma02
```

Intercept $\gamma_{00}$ =

ses $\gamma_{10}$ =

CMses $\gamma_{01}$ =

nonwhite $\gamma_{20}$ =

CMnon30 $\gamma_{02}$ =

```
print("SES L2 Between Slope");      contest1D(Model4b, L=c(0,1,1,0,0)) gamma10 + gamma01
print("Nonwhite L2 Between Slope"); contest1D(Model4b, L=c(0,0,0,1,1)) gamma20 + gamma02

    Estimate Std. Error        df   t value       Pr(>|t|)
1  5.3271536 0.40266351 150.0205  13.22979  5.7141854e-27  gamma10 + gamma01
1 -1.5443398 0.55203705 148.27006 -2.7975292 0.0058334273  gamma20 + gamma02

print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model4b)
R2 Random.(Intercept)          R2 L1.sigma2
          0.702954095           0.076944633
```

**Pseudo-R2 Relative to CovEmpty (from SAS)**
**Change in Pseudo-R2 for CovNWsmush vs. CovNWcontext**

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 | PseudoR2Change |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . | . |
| CovNWsmush | UN(1,1) | schoolID | 2.6340 | 0.3976 | 0.69406 | . |
| CovNWsmush | Residual | | 36.1396 | 0.6100 | 0.07686 | . |
| CovNWcontext | UN(1,1) | schoolID | 2.5571 | 0.3889 | **0.70300** | **.008938830** |
| CovNWcontext | Residual | | 36.1365 | 0.6099 | **0.07694** | **.000079702** |

---

## Model 4c: Switch to Cluster-Mean-Centered Versions of Level-1 Predictors Instead

Level 1: $Math_{pc} = \beta_{0c} + \beta_{1c}(SES_{pc} - \overline{SES}_c) + \beta_{2c}(Nonwhite_{pc} - \overline{Nonwhite}_c) + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + \gamma_{01}(\overline{SES}_c - 0) + \gamma_{02}(\overline{Nonwhite}_c - .30) + U_{0c}$

$\beta_{1c} = \gamma_{10}$

$\beta_{2c} = \gamma_{20}$

```
display "STATA Model 4c: Use Cluster-MC Versions of L1 Predictors Instead"
mixed math c.WCses c.CMses c.WCnon c.CMnon30, || schoolID: , ///
            reml dfmethod(satterthwaite) dftable(pvalue) nolog
display "-2LL = " e(ll)*-2          // Print -2LL for model
lincom c.WCses*-1 + c.CMses*1       // SES L2 Contextual Slope
lincom c.WCnon*-1 + c.CMnon30*1     // Nonwhite L2 Contextual Slope
lincom c.CMnon30*1/10                     // Nonwhite L2 Between Slope per 10%
lincom c.WCnon*-1/10 + c.CMnon30*1/10     // Nonwhite L2 Contextual Slope per 10%

predict predSESnon                // Save fixed-effect predicted outcomes for total-R2
corr math predSES                 // Get total r to make R2
display "Total-R2 = " r(rho)^2    // Print total R2 for SES relative to empty model
corr math predSESnon              // Get total r to make R2
display "Total-R2 = " r(rho)^2    // Print total R2 for SES+Nonwhite relative to empty model
```

```
print("R Model 4c: Use Cluster-MC Versions of L1 Predictors Instead")
Model4c = lmer(data=Example3, REML=TRUE,
                 formula=math~1+WCses+CMses+WCnon+CMnon30+(1|schoolID))
print("Show results using Satterthwaite DDF including -2LL as deviance")
llikAIC(Model4c, chkREML=FALSE); summary(Model4c, ddf="Satterthwaite")
```

```
      AIC        BIC     logLik   deviance    df.resid
 46404.554  46452.712 -23195.277  46390.554    7178.000

Random effects:
 Groups    Name         Variance Std.Dev.
 schoolID (Intercept)   2.5588   1.5996   Var(U_0c)
 Residual              36.1361   6.0113   Var(e_pc)

Fixed effects:
             Estimate Std. Error         df  t value  Pr(>|t|)
(Intercept)  12.64217    0.14685  152.08121  86.0872 < 2.2e-16   gamma00
WCses         1.95248    0.10889 7020.17024  17.9313 < 2.2e-16   gamma10
CMses         5.32715    0.40266  150.02050  13.2298 < 2.2e-16   gamma01
WCnon        -2.89558    0.22017 7020.17024 -13.1515 < 2.2e-16   gamma20
CMnon30      -1.54434    0.55204  148.27007  -2.7975  0.005833   gamma02
```

**Fixed effects from equivalent Model 4b:**

```
(Intercept)  13.51085
ses           1.95248
CMses         3.37468
nonwhite     -2.89558
CMnon30       1.35124
```

Which slopes changed relative to Model 4b, and why?

```
print("SES L2 Contextual Slope");      contest1D(Model4c, L=c(0,-1,1,0,0))
print("Nonwhite L2 Contextual Slope"); contest1D(Model4c, L=c(0,0,0,-1,1))

    Estimate Std. Error        df   t value              Pr(>|t|)
1  3.3746775 0.41712616 172.76081 8.0903042 0.00000000000010111879
1  1.3512422 0.59432313 199.13638 2.2735817 0.024057918

print("Nonwhite L2 Between Slope per 10%");    contest1D(Model4c, L=c(0,0,0,0,1/10))
print("Nonwhite L2 Contextual Slope per 10%"); contest1D(Model4c, L=c(0,0,0,-1/10,1/10))

     Estimate  Std. Error        df    t value     Pr(>|t|)
1 -0.15443398 0.055203705 148.27007 -2.7975292 0.0058334273
1  0.13512422 0.059432313 199.13638  2.2735817 0.024057918

print("Psuedo-R2 relative to empty model using Jonathan's function")
pseudoRSquaredinator(smallerModel=Model1, largerModel=Model4c)
R2 Random.(Intercept)          R2 L1.sigma2
          0.702954095           0.076944633
```

**Pseudo-R2 Relative to CovEmpty (from SAS)**
**Change in Pseudo-R2 for CovSESbetween vs. CovNWbetween**

| Name | CovParm | Subject | Estimate | StdErr | PseudoR2 | PseudoR2Change |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 8.6097 | 1.0778 | . | . |
| CovEmpty | Residual | | 39.1487 | 0.6607 | . | . |
| CovSESbetween | UN(1,1) | schoolID | 2.6887 | 0.4043 | 0.68772 | . |
| CovSESbetween | Residual | | 37.0200 | 0.6248 | 0.05438 | . |
| CovNWbetween | UN(1,1) | schoolID | 2.5571 | 0.3889 | **0.70300** | **0.015284** |
| CovNWbetween | Residual | | 36.1365 | 0.6099 | **0.07694** | **0.022568** |

```
print("Total-R2 for SES relative to empty model using Jonathan's function")
totalRSquaredinator(model=Model2b, dvName="math", data=Example3)
Total R2: 0.16241

print("Total-R2 for SES+Nonwhite relative to empty model using Jonathan's function")
totalRSquaredinator(model=Model4c, dvName="math", data=Example3)
Total R2: 0.18428
```

**Total-R2 Per Model (from SAS)**
**Change in Total-R2 for PredSES vs. PredSESnon**

| Name | PredCorr | TotalR2 | TotalR2Diff |
|---|---|---|---|
| PredSES | 0.40300 | 0.16241 | . |
| PredSESnon | 0.42928 | 0.18428 | 0.021873 |

**Total-$R^2$ Results:** The two fixed slopes for SES accounted for 16.24% of the total math variance. The two fixed slopes for nonwhite accounted for an additional 2.2% of the total math variance (up to total-$R^2$ = .184).

---

**Sample Results Section starts here—see last two new pages for figures that could be included!**

**[indicates notes about what to customize or also include; note that SE and p-values are not needed if you provide tables for the model solutions]**

**Note that the smushed results are not reported, and results are combined across models to give all fixed slopes of interest (so not all models are reported)…**

The extent to which student math outcomes could be predicted from student-level (and corresponding school-level) variables of socio-economic status (SES) and white versus nonwhite identity was examined in a series of multilevel models in which the 7,185 students were modeled as nested within their 160 schools. Restricted Maximum likelihood (REML) within SAS MIXED [or STATA MIXED or R `lmer`] was used in estimating and reporting all model parameters. The significance of fixed effects was evaluated with univariate Wald tests using Satterthwaite denominator degrees of freedom. Alpha was chosen as .01. Model-implied fixed effects were requested via ESTIMATE [or LINCOM or `contest1D`] statements. Effect size for the fixed effects was evaluated via pseduo-$R^2$ values for the proportion reduction in each variance component relative to a nested model without the predictors in question, as well as with total-$R^2$, the squared correlation between the actual math outcomes and those predicted by the model fixed effects.

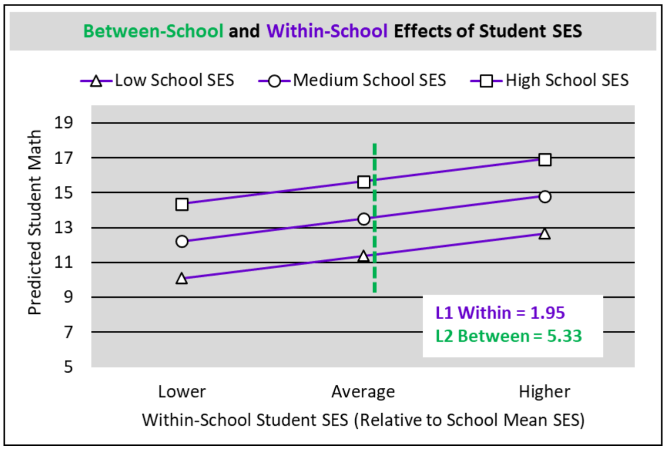
As derived from an empty means, random intercept model, student math had an intraclass correlation of ICC = .180, indicating that 18.0% of the variance in student math was between schools, a significant amount, $-2\Delta LL(1) = 986.12$, $p < .0001$. Likewise, student SES had an intraclass correlation of ICC = .265, which was significantly greater than 0, $-2\Delta LL(1) = 1{,}768.40$, $p < .0001$. For the binary student nonwhite identity variable, a logistic version of the two-level model (i.e., with a logit link function and a Bernoulli level-1 conditional distribution) was estimated instead. Using $\pi^2 / 3$ for the model-scale residual variance, the ICC = .627, which was also significantly greater than 0, $-2\Delta LL(1) = 2{,}965.79$, $p < .0001$. Consequently, cluster-mean-centering was used to partition isolate the level-1 student variability in each predictor, whereas the cluster mean was used to represent the level-2 school variability in each predictor. The cluster mean of SES was left uncentered given its mean near 0, whereas the cluster mean of nonwhite identity (i.e., the proportion of students who identified as nonwhite at each school) was centered at .30 to facilitate interpretation of the other fixed effects.

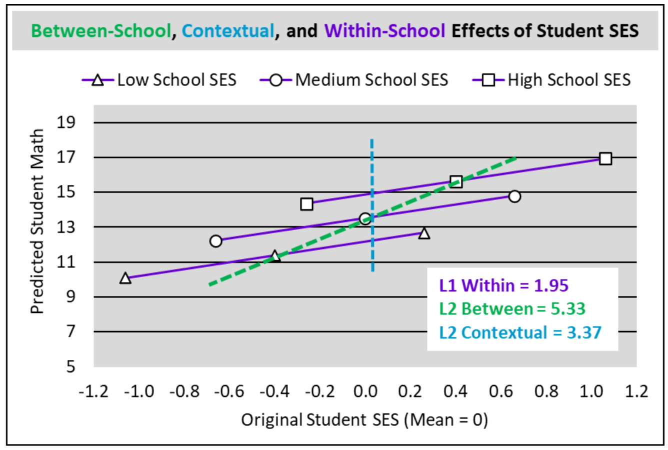
We first examined the effects of level-1 student SES and level-2 school SES, which accounted for 5.4% of the level-1 residual variance in math and 68.8% of the level-2 random intercept variance in math, respectively (total-$R^2$ = .162). The fixed intercept was 12.683 (SE = 0.149), which represented the expected math outcome for a student with average SES for their school from a school whose average SES = 0 (near the mean). At level 1, the within-school slope for student SES was significantly positive, indicating that student math was expected to be higher by 2.191 (SE = 0.109, $p < .0001$) for each additional unit of SES. At level 2, the between-school slope for school mean SES was also significantly positive, indicating that school mean math was expected to be higher by 5.866 (SE = 0.362, $p < .0001$) for each additional unit of school mean SES. The level-2 between-school slope was significantly larger (more positive) than the level-1 within-school slope, as indicated by the level-2 contextual slope for their difference (Est = 3.675). Said differently, the level-2 contextual slope indicated that, after controlling for the effect of student SES, school mean math was expected to be higher by 3.675 (SE = 0.378, $p < .0001$) for each additional unit of school mean SES.

We then added the effects of level-1 student binary nonwhite identity and level-2 school proportion nonwhite identity (i.e., the proportion of of students who identified as nonwhite at each school, centered at .30), which accounted for ~0% of the level-1 residual variance in math and another 0.89% of the level-2 random intercept variance in math, respectively (change in total-$R^2$ = .022). The fixed intercept was 13.511 (SE = 0.161), which represented the expected math outcome for a white-identifying student with average SES for their school who attended a school whose average SES = 0 (near the mean) and with 30% non-white-identifying students (near the mean). The effects of SES remained significant as previously described, and so we focus on the new effects of nonwhite identity. At level 1, the within-school slope for student nonwhite identity was significantly negative, indicating that student math was expected to be lower by 2.899 (SE = 0.220, $p < .0001$) for students identifying as nonwhite relative to white. At level 2, the between-school slope for school mean nonwhite identity was also significantly negative (i.e., Est = −1.544 for the difference between proportion = 0 and 1), indicating that school mean math was expected to be lower by 0.154 (SE = 0.055, $p = .006$) for each additional 10% of students identifying as nonwhite. The level-2 between-school slope was nonsignificantly weaker (less negative) than the level-1 within-school slope, as indicated by the level-2 contextual slope for their difference (Est = 1.351). Said differently, the level-2 contextual slope indicated that, after controlling for the effect of student nonwhite identity, school mean math was expected to be nonsignificantly higher by 0.135 (SE = 0.059, $p = .024$) for each additional 10% of students identifying as nonwhite.

## Figures to Illustrate SES Effects

The figures below demonstrating the SES effects holding the nonwhite predictors constant: at student nonwhite = 0 and school proportion nonwhite students = .30. The **slope of the purple lines shows the positive within-school SES effect** in both figures, but their differing x-axes translate into different level-2 effects being shown!



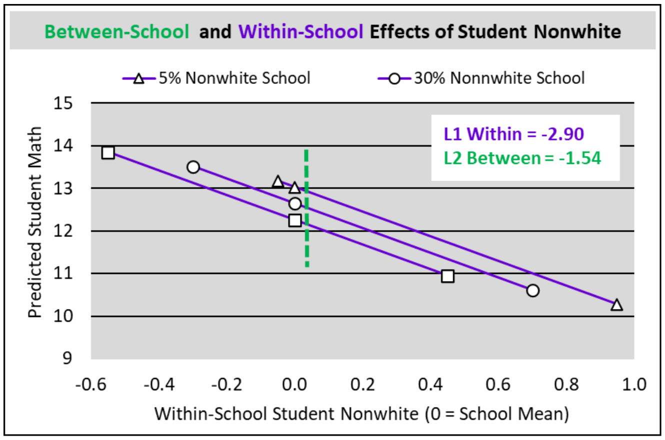



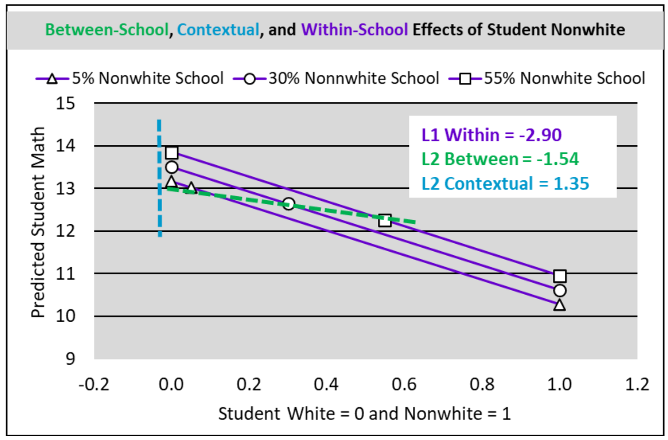

On the left, the x-axis shows level-1 within-school student SES relative to the school's mean (in which lower = −0.66 as −1 SD of school mean, average = 0 when at the school mean, and higher = 0.66 as +1 SD of the school mean for within-school SES). Consequently, the **positive level-2 between-school slope** is shown by the vertical distance between the purple lines (in which low = −0.40 as −1 SD, medium = 0 as near the mean, and high = 0.40 as + 1 SD for school mean SES)—it must be the "between" slope because it holds *within-school* student SES constant at "average" when at the school's mean, whose actual value varies by line! It is not possible to show the positive level-2 contextual slope in the left figure because there is no shown common point at which the *original* student SES predictor is held constant across schools—the x-axis only shows student SES *relative to school mean SES.*

On the right, the x-axis shows original level-1 student SES, in which the values plotted are the same as on the left (to show relatively lower, average, and higher student SES within each school). Consequently, the **positive level-2 contextual slope** is now shown by the vertical distance between the lines (in which low = −0.40 as −1 SD, medium = 0 as near the mean, and high = 0.40 as +1 SD for school mean SES) holding *original* student SES constant (at 0 here, but it could be any constant given the lack of an interaction term between the SES predictors). The **positive level-2 between-school slope** is then shown by diagonal distance between the lines at within-school SES = 0 (when at school mean SES).

### Figures to Illustrate Nonwhite Effects

The figures below demonstrating the nonwhite effects holding the SES predictors constant: at student SES = 0 and school mean SES = 0 (an average student in an average school). The **slope of the purple lines shows the negative within-school nonwhite effect** in both figures, but their differing x-axes translate into different level-2 effects being shown!




On the left, the x-axis shows level-1 within-school student nonwhite relative to the school's mean (in which the lowest = white, middle = school mean, and highest = nonwhite, respectively). Consequently, the **positive level-2 between-school slope** is shown by the vertical distance between the lines (in which low = 0.05, medium = 0.30, and high = 0.55 for school proportion nonwhite, chosen to be a smaller range than ± 1 SD = .30 to avoid a school without any nonwhite students)—it must be the "between" slope because it holds *within-school* student nonwhite constant at the intermediate value for the school's mean, which varies by line. It is not possible to show the positive level-2 contextual slope in the left figure because there is no shown common point at which the *original* student nonwhite predictor is held constant across schools—the x-axis shows student nonwhite relative to school mean nonwhite only.

On the right, the x-axis shows original level-1 student nonwhite, in which the values plotted are the same as on the left (to show white, school mean, and nonwhite students within each school). Consequently, the **positive level-2 contextual slope** is now shown by the vertical distance between the lines (in which low = 0.05, medium = 0.30, and high = 0.55 for school proportion nonwhite, chosen to be a smaller range than ± 1 SD = .30 to avoid a school without any nonwhite students) holding *original* student nonwhite constant (at 0 here, but it could be any constant given the lack of an interaction term between the nonwhite predictors). The **negative level-2 between-school slope** is then shown by diagonal distance between the lines when within-school nonwhite = 0 (at school mean). But this figure is likely to break your readers' brains because the intermediate values—of student nonwhite = school mean—are not possible! Therefore, **the figure below omitting that intermediate value may be more understandable**, but then it cannot be used to show the negative between-school slope.