Example 6a: Generalized Multilevel Models for Categorical Two-Level Nested Outcomes

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

These are the same real data featured in Example 4 from a midwestern rectangular state. These analyses include 13,802 students from 94 schools, with 31–515 students in each school (M = 139). Although admittedly this is not the most meaningful example, we will examine how student lunch status (0 = pay full price for lunch, 1= receive reduced lunch, 2= receive free lunch) can be predicted by student math test scores (i.e., the reverse of Example 4). This handout includes models treating lunch status as **binary** (0 = paid, 1 = reduced or free) or as **ordinal** (original coding). Adaptive quadrature with 7 points of integration (the default in STATA) was used for the random intercept models and the random slope models when possible (whereas the latter required the Laplace method via 1 point of integration instead in R). My attempts at **nominal** multilevel models in SAS, STATA, R, and M*plus* are only in the online files.

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_Example6a"

// Open trimmed example excel data file from sheet "grade10" and clear away existing data
   clear // clear memory in case of open data
   import excel "$filesave\Example6a_Data.xlsx", firstrow case(preserve) sheet("grade10")

// Add labels to original variables
   label variable districtID "districtID: District ID number"
   label variable studentID  "studentID: Student ID number"
   label variable schoolID   "schoolID: School ID number"
   label variable lunch      "lunch: 0=Paid, 1=Reduced, 2=Free"
   label variable math       "math: Math Test Score"

display "STATA Descriptive Statistics within Student-Level Data"
tabulate lunch
summarize math
```

lunch: 0=Paid, 1=Reduced, 2=Free				
lunch	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	9059	69.25	9059	69.25
1	1140	8.71	10199	77.96
2	2883	22.04	13082	100.00

Analysis Variable : math math				
N	Mean	Std Dev	Minimum	Maximum
13082	48.12	17.26	0.00	83.00

```
// Create new lunch variables for clarity
   gen lunch3 = lunch // 3-category version
   gen lunch2 = .     // 2-category version
   replace lunch2=1 if lunch>0
   replace lunch2=0 if lunch==0

// Filter to complete cases before computing cluster means
   egen nmiss=rowmiss(math lunch)
   drop if nmiss>0
// Create sample size per school cluster mean math
   sort schoolID
   egen schoolN = count(math), by(schoolID)
   egen CMmath  = mean(math),  by(schoolID)
   label variable schoolN "schoolN: # Students Sampled Per School"
   label variable CMmath  "CMmath: School Mean Math Score"

// Rescale and center cluster mean math to be per 10 points
   gen CMmath50 = (CMmath-50)/10
   label variable CMmath50 "CMmath: School Mean Math (0=50)"
```

```
// Rescale and cluster-MC student math to be per 10 points
   gen WCmath = (math-CMmath)/10
   label variable WCmath "WCmath: Within-School Math (0=CM)"

display "STATA Descriptive Statistics within School-Level Data"
preserve  // Save for later use, then compute school-level dataset
collapse  schoolN CMmath, by(schoolID)
summarize schoolN CMmath
restore   // Go back to student-level dataset
```

CMmath predictor SD is relevant for making plots to show low/high values

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
schoolN	schoolN: # Students Sampled Per School	94	139.17	138.20	31.00	515.00
CMmath	CMmath: School Mean Math Score	94	47.73	6.97	29.45	61.61

```
display "STATA Descriptive Statistics within Student-Level Data"
summarize WCmath, detail
```

Analysis Variable : WCmath					
N	Mean	Variance	Std Dev	Minimum	Maximum
13082	-0.000	**2.514**	1.586	-4.465	4.488

This predictor variance will be used in computing slope reliability later…

R Syntax for Importing and Preparing Data for Analysis (after loading packages *readxl*, *TeachingDemos*, *psych*, *lme4*, *lmerTest*, *performance*, *multcomp*, *prediction*, *ordinal*, and *mclogit*):

```
# Define variables for working directory and data name -- CHANGE THESE
filesave = "C:\\Dropbox/23_PSQF6272/PSQF6272_Example6a/"
filename = "Example6a_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 trimmed example excel data file from sheet "grade10"
Example6a = read_excel(paste0(filesave,filename), sheet="grade10")
# Convert to data frame to use in analysis
Example6a = as.data.frame(Example6a)

print("R Descriptive Statistics within Student-Level Data")
prop.table(table(x=Example6a$lunch, useNA="ifany"))

# Create new lunch variables for clarity
Example6a$lunch3 = Example6a$lunch    # 3-category version
Example6a$lunch2=NA                   # 2-category version
Example6a$lunch2[which(Example6a$lunch>0)]=1
Example6a$lunch2[which(Example6a$lunch==0)]=0

# Filter to only cases complete on all variables to be used below (before cluster means)
Example6a = Example6a[complete.cases(Example6a[ , c("math","lunch")]),]

# Create cluster mean math using Jonathan's function
Example6a = addUnitMeans(data=Example6a, unitVariable="schoolID",
                         meanVariables=c("math"), newNames=c("CMmath"))

print("R Descriptive Statistics within School-Level Data")
schoolMeans = unique(Example6a[,c("schoolID","NperschoolID","CMmath")])
describe(x=schoolMeans[ , c("NperschoolID","CMmath")])

# Rescale and center cluster mean math to be per 10 points
Example6a$CMmath50 = (Example6a$CMmath-50)/10
# CMmath50 = "CMmath50: School Mean Math (0=50)
# Rescale and cluster-MC student math to be per 10 points
Example6a$WCmath = (Example6a$math - Example6a$CMmath)/10
# WCmath= "WCmath: Within-School Math (0=CM)

print("R Descriptive Statistics within Student-Level Data")
describe(x=Example6a[ , c("WCmath")]); var(Example6$WCmath)
```

Model 1. Empty Means, Single-Level Logistic Model Predicting Lunch2: Binary Paid Lunch (=0) vs. Reduced or Free Lunch (=1)

lunch	Frequency	Percent	Cumulative Percent
0	9059	69.25	69.25
1	1140	8.71	77.96
2	2883	22.04	100.00

Lunch 2 collapsed 1 and 2 into 1 (= 30.75%)

Level 1: $Log\left[\frac{prob(lunch2_{pc}=1)}{prob(lunch2_{pc}=0)}\right] = Logit\left(lunch2_{pc} = 1\right) = \beta_{0c}$

Level 2: $\beta_{0c} = \gamma_{00}$ **(g = gamma in annotation below)**

```
display "STATA Model 1: Empty Means, Single-Level for Student Binary Lunch"
melogit lunch2 , nolog coeflegend

Log likelihood = -8072.9469                      Prob > chi2       =          .
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _cons |  -.8117308   .0189462   -42.84   0.000    -.8488647    -.774597  g00 in logits
------------------------------------------------------------------------------

display "-2LL = " e(ll)*-2 // Print -2LL for model
-2LL = 16145.894
```

$$Prob(y=1) = \frac{\exp(-0.8117)}{1+\exp(-0.8117)} = .3075$$

```
nlcom 1/(1+exp(-1*(_b[_cons]))) // Fixed intercept in probability
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .3075218   .0040346    76.22   0.000     .2996141    .3154295 g00 in prob
----------------------------------------------------------------------------

print("R Model1: Empty Means, Single-Level for Student Binary Lunch")
Model1 = glm(data=Example6a, family=binomial(link="logit"), formula=lunch2~1)
summary(Model1) # residual deviance = -2LL already

Coefficients:
             Estimate Std. Error z value  Pr(>|z|)
(Intercept) -0.811731   0.018946 -42.844 < 2.2e-16  g00 in logits

    Null deviance: 16145.9  on 13081  degrees of freedom
Residual deviance: 16145.9  on 13081  degrees of freedom  → -2LL for model

print("Convert logits to probability via inverse link")
Model1Prob=1/(1+exp(-1*coefficients(Model1))); Model1Prob
(Intercept)
 0.30752179 → g00 in probability
```

Model 2. Empty Means, Two-Level Logistic Model Predicting Paid (=0) vs. Reduced/Free Lunch (=1)

Level 1: $Log\left[\frac{prob(lunch2_{pc}=1)}{prob(lunch2_{pc}=0)}\right] = Logit\left(lunch2_{pc} = 1\right) = \beta_{0c}$

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

```
display "STATA Model 2: Empty Means, Two-Level Logistic Model Predicting Binary Lunch"
melogit lunch2 , || schoolID: , intpoints(7) nolog

------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _cons |   -1.17212   .1488225    -7.88   0.000    -1.463807   -.8804335  g00 in logits
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|   1.955621   .3312813                       1.40311    2.725699  Var(U0c)
------------------------------------------------------------------------------
LR test vs. logistic model: chibar2(01) = 2973.47     Prob >= chibar2 = 0.0000
```

```
display "-2LL = " e(ll)*-2      // Print -2LL for model
-2LL = 13172.424

estat icc                       // ICC using 3.29 as residual variance
Intraclass correlation
------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                    schoolID |   .3728196   .0396099      .2989807    .4531076
------------------------------------------------------------------------------

nlcom 1/(1+exp(-1*(_b[_cons]))) // Fixed intercept in probability
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |    .236472   .0268703     8.80   0.000     .1838071    .2891369  g00 in prob
------------------------------------------------------------------------------

print("R Model 2: Empty Means, Random Intercept for Student Binary Lunch")
Model2 = glmer(data=Example6a, family=binomial(link="logit"), nAGQ=7,
               lunch2~1+(1|schoolID))
print("Show -2LL with more precision, results, and ICC using 3.29=residual variance")
-2*logLik(Model2); summary(Model2); icc(Model2)

'log Lik.' 13172.43 (df=2) → -2LL for model

     AIC      BIC   logLik deviance df.resid
 13176.4  13191.4  -6586.2  13172.4    13080

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept) 1.9545   1.398   Var(U_0c)

Fixed effects:
            Estimate Std. Error z value  Pr(>|z|)
(Intercept) -1.17212    0.14937 -7.8468 4.267e-15 g00 in logits

# Intraclass Correlation Coefficient
    Adjusted ICC: 0.373
  Unadjusted ICC: 0.373
```

Model-scale ICC for the correlation of students in the same school for lunch2:

$$\text{ICC} = \frac{1.9545}{1.9545 + 3.29} = .373$$

```
print("Convert logits to probability via inverse link")
Model2Prob=1/(1+exp(-1*fixef(Model2))); Model2Prob
(Intercept)
 0.23647289 → g00 in probability
```

The fixed intercept in probability (.2364) no longer matches the outcome mean (.3075 = proportion of 1 values). Instead, it is "unit-specific": it is the predicted outcome for a school with $U_{0c} = 0$, which is closer to the median of the school means than the mean of the school means according to Stroup's 2016 book); see also Hedeker & Gibbons' 2006 book (which our library has).

```
print("Compute LRT manually -- would not work any other way across different packages")
DevTest=-2*(logLik(Model1)-logLik(Model2))
Pvalue=pchisq((DevTest), df=1, lower.tail=FALSE)
print("Test Statistic and P-values for DF=1")
DevTest; Pvalue
'log Lik.' 2973.4638 (df=1)
'log Lik.' 0 (df=1)
```

Calculate a 95% random effect confidence interval for the school random intercept:

*CI = fixed effect ± 1.96*SQRT(random intercept variance)*

*CI = −1.1721 ± 1.96*SQRT(1.9545) = −3.91 to 1.57 in logits, or 0.02 to 0.83 in probability!*

Model 3. Add a Level-2 Fixed Effect of School Mean Student Math (0=50 per 10 points)

Level 1: $Logit\left(lunch2_{pc}=1\right)=\beta_{0c}$

Level 2: $\beta_{0c}=\gamma_{00}+\gamma_{01}([CMmath_c-50]/10)+U_{0c}$

```
display "STATA Model 3: Add Level-2 Fixed Slope of School Mean Math"
melogit lunch2 c.CMmath50, || schoolID: , intpoints(7) nolog
display "-2LL = " e(ll)*-2                       // Print -2LL for model
margins , at(c.CMmath50=(-1(1)1)) predict(xb)   // Unit-specific predicted logits (at U0c=0)
margins , at(c.CMmath50=(-1(1)1)) predict(mu)   // Marginal predicted probabilities
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    CMmath50 |  -1.442861   .1401903   -10.29   0.000    -1.717629   -1.168093  g01
       _cons |  -1.469553   .1039214   -14.14   0.000    -1.673236   -1.265871  g00
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|    .765871   .1448655                      .5286275    1.109587  Var(U0c)
------------------------------------------------------------------------------

display "STATA Model 3: Odds Ratios Instead"
melogit lunch2 c.CMmath50, || schoolID: , intpoints(7) nolog or
------------------------------------------------------------------------------
      lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    CMmath50 |   .2362508   .0331201   -10.29   0.000     .1794911    .3109593  exp(g01)
       _cons |   .2300282   .0239049   -14.14   0.000      .187639    .2819936  exp(g00)
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|    .765871   .1448655                      .5286275    1.109587 → not exp
------------------------------------------------------------------------------

print("R Model 3: Add Level-2 Fixed Slope of School Mean Math")
Model3 = glmer(data=Example6a, family=binomial(link="logit"), nAGQ=7,
               lunch2~1+CMmath50+(1|schoolID))
print("Show -2LL with more precision, results, and odds ratios")
-2*logLik(Model3); summary(Model3); exp(fixef(Model3))

'log Lik.' 13103.224 (df=3) → -2LL for model
     AIC      BIC   logLik deviance df.resid
 13109.2  13131.7  -6551.6  13103.2    13079

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept) 0.76572  0.87506  Var(U_0c)

Fixed effects:
            Estimate Std. Error z value  Pr(>|z|)
(Intercept) -1.46957    0.10400 -14.130 < 2.2e-16  g00
CMmath50    -1.44290    0.14027 -10.287 < 2.2e-16  g01

(Intercept)      CMmath50
 0.23002323     0.23624240 → exp(g)
(odds of y=1)   (odds ratio for unit change in CMmath50)
```

Converting the logit intercept into probability:
$$Prob(y=1)=\frac{\exp(-1.470)}{1+\exp(-1.470)}=.230$$

What does the fixed intercept represent? *The logit = −1.4696 for the probability of getting reduced/free lunch for a student in a school with a random intercept U_{0c} = 0 and school mean math = 50, which corresponds to a probability = .230 (as found from the inverse link function above).*

What does the main effect of school mean math represent? *Without controlling for student math, for every 10 units higher school mean math, the logit for the probability of getting reduced/free lunch is significantly lower by 1.4429, which translates into an odds ratio of 0.236. This is the level-2 between-school math effect, which accounted for 60.82% of the level-2 school random intercept variance (as computed in SAS, see next page).*

```
print("Yhat in logits and probabilities for unit-specific values of predictor")
Model3Logits = prediction(model=Model3, type="link",     re.form=NA, at=list(CMmath50=-1:1))
Model3Probs  = prediction(model=Model3, type="response", re.form=NA, at=list(CMmath50=-1:1))
summary(Model3Logits); summary(Model3Probs)

at(CMmath50) Prediction SE  z  p lower upper     at(CMmath50) Prediction SE  z  p lower upper
          -1   -0.02668 NA NA NA    NA    NA               -1    0.49333 NA NA NA    NA    NA
           0   -1.46957 NA NA NA    NA    NA                0    0.18701 NA NA NA    NA    NA
           1   -2.91247 NA NA NA    NA    NA                1    0.05154 NA NA NA    NA    NA
```

The ***predicted logits on the left (which each reflect a one unit difference in CMmath50) all differ by exactly −1.4429****, the slope of CMmath50. In contrast, the predicted probabilities on the right do not have a constant one-unit difference—and that's why you can't talk about slopes directly in terms of what they do to predicted probabilities!*

Pseudo-R2 Relative to 2.CovEmpty (from SAS)

Name	CovParm	Subject	Estimate	PseudoR2
2.CovEmpty	UN(1,1)	schoolID	1.9545	.
3.CovCMmath	UN(1,1)	schoolID	0.7657	0.60823

Because the level-2 random intercept variance is a freely estimated quantity, it can be reduced as usual by adding level-2 cluster characteristics (like school mean math here).

This will not be true for level-1 predictors, as shown next!

Model 4. Add a Level-1 Fixed Slope of Cluster-Mean-Centered Student Math

Level 1: $Logit\left(lunch2_{pc}=1\right)=\beta_{0c}+\beta_{1c}\left(\left[math_{pc}-CMmath_{c}\right]/10\right)$

Level 2: $\beta_{0c}=\gamma_{00}+\gamma_{01}([CMmath_{c}-50]/10)+U_{0c}$

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

```
display "STATA Model 4: Add Level-1 Fixed Slope of Cluster-MC Student Math"
melogit lunch2 c.CMmath50 c.WCmath, || schoolID: , intpoints(7) nolog
estimates store FixMath               // Save LL for LRT
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    CMmath50 |  -1.517373   .1465674   -10.35   0.000     -1.80464   -1.230106   g01
      WCmath |  -.3719926   .0144957   -25.66   0.000    -.4004037   -.3435814   g10
       _cons |  -1.559778   .1087199   -14.35   0.000    -1.772865   -1.346691   g00
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|   .8416117   .1576131                      .5830459    1.214845   Var(U0C)
------------------------------------------------------------------------------
LR test vs. logistic model: chibar2(01) = 716.80      Prob >= chibar2 = 0.0000

display "-2LL = " e(ll)*-2           // Print -2LL for model
-2LL = 12390.671

lincom c.WCmath*-1 + c.CMmath50*1    // Math Contextual Slope
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |  -1.145381   .1467268    -7.81   0.000     -1.43296   -.8578014   g01-g10
------------------------------------------------------------------------------

display "STATA Model 4: Odds Ratios Instead"
melogit lunch2 c.CMmath50 c.WCmath, || schoolID: , intpoints(7) nolog or
------------------------------------------------------------------------------
      lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    CMmath50 |   .2192872   .0321404   -10.35   0.000     .1645337    .2922615   exp(g01)
      WCmath |   .6893594   .0099928   -25.66   0.000     .6700495    .7092257   exp(g10)
       _cons |   .2101828    .022851   -14.35   0.000     .1698457    .2600996   exp(g00)
-------------+----------------------------------------------------------------
```

```
lincom c.WCmath*-1 + c.CMmath50*1, or  // Math Contextual Slope
------------------------------------------------------------------------------
      lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .3181028   .0466742    -7.81   0.000     .2386016    .4240935   exp(g01-g10)
------------------------------------------------------------------------------
```

```
print("R Model 4: Add Level-1 Fixed Slope of Cluster-MC Student Math")
Model4 = glmer(data=Example6a, family=binomial(link="logit"), nAGQ=7,
               lunch2~1+CMmath50+WCmath+(1|schoolID))
print("Show -2LL with more precision, results, and odds ratios")
-2*logLik(Model4); summary(Model4); exp(fixef(Model4))
```

```
'log Lik.' 12390.672 (df=4) → -2LL for model

     AIC      BIC   logLik deviance df.resid
 12398.7  12428.6  -6195.3  12390.7    13078

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept) 0.84145  0.91731  Var(U_0c)

Fixed effects:
             Estimate Std. Error z value  Pr(>|z|)
(Intercept) -1.559803   0.108808 -14.335 < 2.2e-16  g00
CMmath50    -1.517410   0.146655 -10.347 < 2.2e-16  g01
WCmath      -0.371993   0.014496 -25.662 < 2.2e-16  g10

(Intercept)    CMmath50      WCmath
 0.21017743  0.21927903  0.68935917 → exp(g)
```

```
Model 3 Fixed effects:
            Estimate Std. Error
(Intercept) -1.46957    0.10400 gamma00
CMmath50    -1.44290    0.14027 gamma01
```

Note that the L2 math slope is **stronger** below than before (above). This is an **artifact** of the model scale changing—the **model total variance had to increase** in order for the L1 residual variance to still be 3.29!

```
Model4glht = summary(glht(model=Model4, linfct=rbind(
    "Math Contextual Slope and Odds Ratio"= c(0,1,-1))),test=adjusted("none"))
Model4glht; data.frame(OR=exp(Model4glht$test$coefficients))
```

```
Linear Hypotheses:
                                          Estimate Std. Error z value  Pr(>|z|)
Math Contextual Slope and Odds Ratio == 0 -1.14542    0.14681 -7.8018 5.995e-15  g01-g11

                                             OR
Math Contextual Slope and Odds Ratio 0.31809112 exp(g01-g10)
```

What does the fixed intercept NOW represent? *The logit = −1.5598 for the probability of getting reduced or free lunch for a student in a school with a random intercept U_{0c} = 0 and school mean math = 50 and within-school math = 0 (e.g., an average student), which translates into a probability = .210.*

What does the main effect of school mean math NOW represent? *The interpretation is the same: Without controlling for student math, for every 10 units higher school mean math, the logit for the probability of getting reduced/free lunch is significantly lower by 1.5174, which translates into an odds ratio of 0.219. This effect is still significant after controlling for student math (as indicated by a level-2 contextual effect = −1.1454).*

What does the main effect of student math represent? *For every 10 units higher student math relative to the rest of your school, the logit for the probability of getting reduced/free lunch is significantly lower by 0.372, which translates into an odds ratio of 0.689. We cannot compute a pseudo-R^2 for the residual variance, which remains 3.29 in logits.*

Model 5. Add a Random Slope of Cluster-Mean-Centered Student Math

Level 1: $Logit\left(lunch2_{pc}=1\right)=\beta_{0c}+\beta_{1c}\left(\left[math_{pc}-CMmath_c\right]/10\right)$

Level 2: $\beta_{0c}=\gamma_{00}+\gamma_{01}([CMmath_c-50]/10)+U_{0c}$

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

```
display "STATA Model 5: Add Random Slope of Cluster-MC Student Math"
melogit lunch2 c.CMmath50 c.WCmath, || schoolID: WCmath, cov(un) intpoints(7) nolog

------------------------------------------------------------------------------------
          lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------------
        CMmath50 |  -1.561634   .1476518   -10.58   0.000    -1.851026   -1.272241  g01
          WCmath |  -.3434491   .0242425   -14.17   0.000    -.3909636   -.2959346  g10
           _cons |  -1.566559   .1075517   -14.57   0.000    -1.777357   -1.355762  g00
-----------------+------------------------------------------------------------------
schoolID         |
     var(WCmath)|    .0160741    .005429                      .0082915    .0311616  Var(U1c)
      var(_cons)|    .8119266   .1540144                       .559825    1.177555  Var(U0c)
-----------------+------------------------------------------------------------------
schoolID         |
 cov(WCmath,_cons)|  -.0351931   .0290211    -1.21   0.225    -.0920734    .0216872  Cov(U0c,U1c)
------------------------------------------------------------------------------------
LR test vs. logistic model: chi2(3) = 755.46              Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

display "-2LL = " e(ll)*-2  // Print -2LL for model
-2LL = 12352.007

estimates store RandMath    // Save LL for LRT
lrtest RandMath FixMath     // LRT against fixed-only WCmath slope
Likelihood-ratio test
Assumption: FixMath nested within RandMath

 LR chi2(2) =  38.66
Prob > chi2 = 0.0000

display "STATA Model 5: Odds Ratios Instead"
melogit lunch2 c.CMmath50 c.WCmath, || schoolID: WCmath, cov(un) intpoints(7) nolog or

------------------------------------------------------------------------------------
          lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------------
        CMmath50 |   .2097931   .0309763   -10.58   0.000     .1570759    .2802029 exp(g01)
          WCmath |   .7093196   .0171957   -14.17   0.000     .6764048    .7438361 exp(g10)
           _cons |   .2087623   .0224527   -14.57   0.000     .1690845    .2577509 exp(g00)
------------------------------------------------------------------------------------

print("R Model 5: Add Random Slope of Cluster-MC Student Math")
print("Switched to Laplace estimation with 1 quadrature point")
Model5 = glmer(data=Example6a, family=binomial(link="logit"), nAGQ=1,
               lunch2~1+CMmath50+WCmath+(1+WCmath|schoolID))
print("Show -2LL with more precision, results, and odds ratios")
-2*logLik(Model5); summary(Model5); exp(fixef(Model5))

'log Lik.' 12353.038 (df=6) → -2LL for model

     AIC       BIC    logLik deviance df.resid
 12365.0   12409.9   -6176.5  12353.0    13076

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 schoolID (Intercept) 0.805419 0.89745
          WCmath      0.015833 0.12583  -0.309

Fixed effects:
             Estimate Std. Error z value  Pr(>|z|)
(Intercept) -1.565979   0.106814 -14.661 < 2.2e-16
CMmath50    -1.561167   0.146741 -10.639 < 2.2e-16
WCmath      -0.343533   0.023989 -14.321 < 2.2e-16

(Intercept)    CMmath50      WCmath
 0.20888332  0.20989101  0.70926033 → exp(g)
```

SAS output including each parameter's **gradient** = slope for the partial derivative with respect to each parameter, which should be ~0 at most likely estimate:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Gradient
UN(1,1)	schoolID	0.8118	0.1540	-0.00188
UN(2,1)	schoolID	-0.03524	0.02906	0.00736
UN(2,2)	schoolID	0.01608	0.005433	**0.324488**

Note that the level-2 random slope variance across schools for the effect of student math is not estimated very well—the gradient is far away from 0!

Note that in R glmer I had to reduce the number of quadrature points from 7 to 1, which is then the "laplace" method. So re-estimated Model 4 the same way (using 1 quadrature point) to ensure their comparability for an LRT (and thus its results differ a bit from those of STATA and SAS using 7 quadrature points instead).

```
print("Re-estimate Model 4 with same laplace method for LRT")
Model4R = glmer(data=Example6a, family=binomial(link="logit"), nAGQ=1,
                lunch2~1+CMmath50+WCmath+(1|schoolID))
print("LRT for random slope"); anova(Model5, Model4R)

        npar     AIC     BIC   logLik deviance   Chisq Df       Pr(>Chisq)
Model4R    4 12399.4 12429.3 -6195.69  12391.4
Model5     6 12365.0 12409.9 -6176.52  12353.0 38.3353  2 0.0000000047381
```

Does the level-2 random slope of within-school math improve model fit? *Yes, −2ΔLL(~2) = 38.34, p < .001*

Calculate a 95% random effect confidence interval for the student math slope:
*CI = fixed effect ± 1.96*SQRT(random slope variance)*
*CI = −0.3435 ± 1.96*SQRT(0.015833) = −0.59 to −0.10 in logits (there is no analog in probability terms)*

Random slope reliability: $\mathrm{SR} = \dfrac{\tau^2_{U_1}}{\tau^2_{U_1} + [\sigma^2_e/(L1n * var)]} = \dfrac{.0160741}{.0160741 + [3.29/(139 * 2.514)]} = \mathbf{.627}$

Btw, 2.514 is the variance of the cluster-mean-centered L1 WCmath predictor that has the random slope.

So what does this mean? *The extent to which within-school student differences in math predicts student reduced/free lunch status varies significantly across schools, but across 95% of schools, higher student math is still predicted to relate to a lower probability of receiving reduced or free lunch.*

Model 6. Adding Intra-Variable Interactions of School Mean Math and Cluster-MC Student Math

Level 1: $Logit(lunch2_{pc} = 1) = \beta_{0c} + \beta_{1c}([math_{pc} - CMmath_c]/10)$

Level 2: $\beta_{0c} = \gamma_{00} + \gamma_{01}([CMmath_c - 50]/10) + \gamma_{02}([CMmath_c - 50]/10)^2 + U_{0c}$

$\beta_{1c} = \gamma_{10} + \gamma_{11}([CMmath_c - 50]/10) + U_{1c}$

7 quadrature points, so results won't match R

```
display "STATA Model 6: Add Interactions of School Mean and Cluster-MC Student Math"
melogit lunch2 c.CMmath50 c.WCmath c.CMmath50#c.WCmath c.CMmath50#c.CMmath50, ///
        || schoolID: WCmath, cov(un) intpoints(7) nolog

------------------------------------------------------------------------------------
             lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
--------------------+---------------------------------------------------------------
           CMmath50 |  -1.583295   .1996128    -7.93   0.000    -1.974529   -1.192061  g01
             WCmath |  -.3687824   .0263208   -14.01   0.000    -.4203701   -.3171946  g10
  c.CMmath50#c.WCmath |  -.0696396   .0336319    -2.07   0.038    -.1355569   -.0037223  g11
c.CMmath50#c.CMmath50 |  -.0685714   .1759445    -0.39   0.697    -.4134163    .2762735  g02
              _cons |  -1.546039   .1231085   -12.56   0.000    -1.787328   -1.304751  g00
--------------------+---------------------------------------------------------------
schoolID            |
        var(WCmath)|    .0134755    .004905                        .0066025     .027503  Var(U1c)
         var(_cons)|    .8158619   .1553881                        .5616913    1.185047  Var(U0c)
--------------------+---------------------------------------------------------------
schoolID            |
   cov(WCmath,_cons)|  -.0276683   .0279353    -0.99   0.322    -.0824204    .0270838 Cov(U0,U1)
------------------------------------------------------------------------------------

display "-2LL = " e(ll)*-2                          // Print -2LL for model
-2LL = 12347.844
```

```
lincom c.WCmath*-1 + c.CMmath50*1                          // Math Contextual Simple Main Effect
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |  -1.214512   .1992125    -6.10   0.000    -1.604962    -.824063  g01-g11
------------------------------------------------------------------------------
lincom c.CMmath50#c.WCmath*-1 + c.CMmath50#c.CMmath50*1  // Math Contextual Interaction
------------------------------------------------------------------------------
      lunch2 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0010682    .177141     0.01   0.995    -.3461217    .3482581  g02-g11
------------------------------------------------------------------------------

display "STATA Model 6: Odds Ratios Instead"
melogit lunch2 c.CMmath50 c.WCmath c.CMmath50#c.WCmath c.CMmath50#c.CMmath50, ///
        || schoolID: WCmath, cov(un) intpoints(7) nolog or
---------------------------------------------------------------------------------------
               lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
----------------------+----------------------------------------------------------------
             CMmath50 |   .2052976     .04098    -7.93   0.000     .1388267     .303595  exp(g01)
               WCmath |   .6915759   .0182028   -14.01   0.000     .6568037     .728189  exp(g10)
    c.CMmath50#c.WCmath |   .9327299   .0313695    -2.07   0.038     .8732295    .9962846  exp(g11)
c.CMmath50#c.CMmath50 |   .9337268   .1642841    -0.39   0.697     .6613869    1.318208  exp(g02)
                _cons |   .2130903   .0262332   -12.56   0.000     .1674069      .27124  exp(g00)
----------------------+----------------------------------------------------------------
lincom c.WCmath*-1 + c.CMmath50*1, or                   // Math Contextual Simple Main Effect
------------------------------------------------------------------------------
      lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .2968547   .0591372    -6.10   0.000     .2008972    .4386458  exp(g01-g10)
------------------------------------------------------------------------------
lincom c.CMmath50#c.WCmath*-1 + c.CMmath50#c.CMmath50*1, or  // Math Contextual Interaction
------------------------------------------------------------------------------
      lunch2 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   1.001069   .1773303     0.01   0.995     .7074264    1.416598  exp(g02-g11)
------------------------------------------------------------------------------
```

```
print("R Model 6: Add Interactions of School Mean and Cluster-MC Student Math")
print("Switched to Laplace estimation with 1 quadrature point")
Model6 = glmer(data=Example6a, family=binomial(link="logit"), nAGQ=1,
               lunch2~1+CMmath50+WCmath+CMmath50:WCmath+I(CMmath50^2)+(1+WCmath|schoolID))
print("Show -2LL with more precision, results, and odds ratios")
-2*logLik(Model6); summary(Model6); exp(fixef(Model6))

'log Lik.' 12348.838 (df=8) → -2LL for model

     AIC      BIC   logLik deviance df.resid 
 12364.8  12424.7  -6174.4  12348.8    13074 

Random effects:
 Groups   Name        Variance Std.Dev. Corr  
 schoolID (Intercept) 0.80938  0.89965          Var(U0c)
          WCmath      0.01327  0.11520  -0.264  Var(U1c) Cor(U0,U1)

Fixed effects:
                 Estimate Std. Error  z value  Pr(>|z|)    
(Intercept)     -1.545376   0.122411 -12.6244 < 2.2e-16  g00
CMmath50        -1.582693   0.198230  -7.9841 1.415e-15  g01
WCmath          -0.368932   0.025974 -14.2039 < 2.2e-16  g10
I(CMmath50^2)   -0.068349   0.174916  -0.3908    0.6960  g02
CMmath50:WCmath -0.069660   0.033344  -2.0891    0.0367  g11

optimizer (Nelder_Mead) convergence code: 0 (OK)  (Pry random slope variance not estimated well)
Model failed to converge with max|grad| = 0.00321814 (tol = 0.002, component 1)

    (Intercept)        CMmath50          WCmath   I(CMmath50^2) CMmath50:WCmath 
     0.21323165      0.20542111      0.69147232      0.93393412      0.93271105 → exp(g)
```

```
Model6glht = summary(glht(model=Model6, linfct=rbind(
  "Math Contextual Simple Main Effect" = c(0,1,-1,0, 0),
  "Math Contextual Interaction"        = c(0,0, 0,1,-1))),test=adjusted("none"))
Model6glht; data.frame(OR=exp(Model6glht$test$coefficients))

Linear Hypotheses:
                                              Estimate Std. Error z value         Pr(>|z|)
Math Contextual Simple Main Effect == 0 -1.2137610  0.1978949 -6.1334 0.0000000008604  g01-g10
Math Contextual Interaction == 0         0.0013104  0.1761612  0.0074          0.9941  g02-g11
(Adjusted p values reported -- none method)

                                           OR
Math Contextual Simple Main Effect 0.29707786  exp(g01-g10)
Math Contextual Interaction        1.00131130  exp(g02-g11)
```

Pseudo-R2 Relative to 5.CovRandMath (from SAS)

Name	CovParm	Subject	Estimate	PseudoR2
5.CovRandMath	UN(1,1)	schoolID	0.8118	.
5.CovRandMath	UN(2,2)	schoolID	0.01608	.
6.CovInteract	UN(1,1)	schoolID	0.8157	-0.00479
6.CovInteract	UN(2,2)	schoolID	0.01348	0.16163

Because the level-2 random intercept and WCmath slope variances still freely estimated quantities, they could be reduced as usual by adding level-2 interactions or cross-level interactions with WCmath, respectively. However, intercept variance increased!

What does the Within-School*Between-School math interaction represent? *For every 10 units higher school mean math, the effect of within-school student differences in math on student reduced/free lunch (which is −0.369 as evaluated at school mean math = 50) becomes significantly more negative by 0.070. So the effect of being "smarter than the others" is even stronger in a "smart" school, which accounted for 16.16% of the level-2 school random slope variance in the level-1 effect of within-school student math.*

What does the Between-School*Between-School math interaction represent? *Without controlling for student math, for every 10 units higher school mean math, the effect of school mean math on school mean reduced/free lunch (which is −1.583 as evaluated at school mean math = 50) becomes nonsignificantly more negative by 2*0.068. So the effect of being in a "smart" school is predominantly linear. The quadratic effect of school mean math did not account for any level-2 school random intercept variance (which increased by 0.479% instead).*

What do the contextual math effects represent? *After controlling for student math, there is a contextual effect of school mean math = −1.214 per 10 units as evaluated at school mean math = 50 for an average student (math = 50). However, there is not a contextual effect of how school mean math moderates the effect of within-school student math (incremental interaction = 0.0011). —OR — The between-school math effect is significantly more negative by 1.214 as evaluated at school mean math = 50 for an average student. However, school mean math does not moderate the level-2 between-school math effect (–0.068) differently than the level-1 within-school math effect (–0.070).*

Sample Results Section for Binary Multilevel Models using STATA Output

[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]

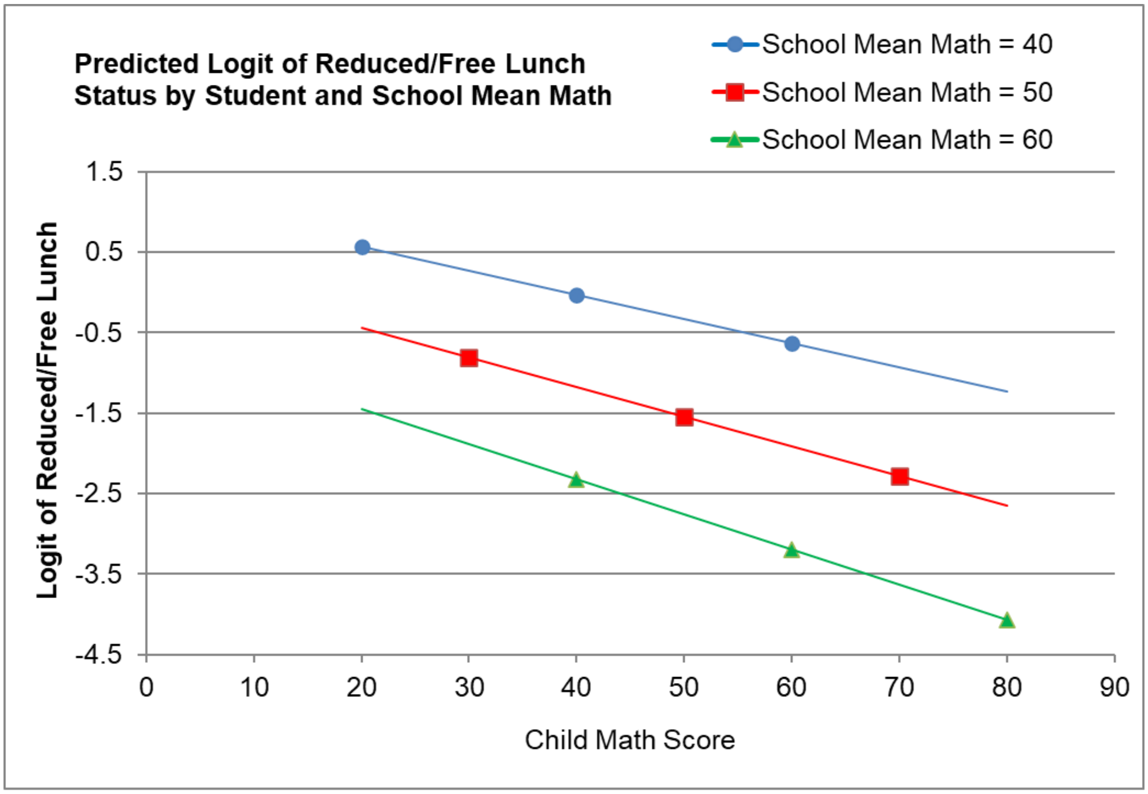
Overall, 30.75% of the sample students received reduced or free lunch; the proportion of students receiving reduced or free lunch in each school ranged from 0 to 80.33%. The extent to which student math outcomes could predict student reduced or free lunch status was examined in a series of multilevel models in which the 13,802 students were modeled as nested at level 1 within their 94 schools at level 2, and school differences were captured via school-level random effects. The binary lunch status outcome was predicted using a logit link function and Bernoulli conditional outcome distribution. All model parameters were estimated via full-information marginal maximum likelihood (MML) using adaptive Gaussian quadrature with 7 points of integration per random effect dimension in STATA MELOGIT v. 17. Accordingly, all fixed effects should be interpreted as unit-specific (i.e., as the fixed effect specifically for schools in which the corresponding random effect = 0). The significance of fixed effects was evaluated with Wald tests (i.e., the *z*-test of the ratio of each estimate to its standard error without denominator degrees of freedom), whereas the significance of random effects was evaluated via likelihood ratio tests (i.e., −2ΔLL with degrees of freedom equal to the number of new random effects variances and covariances). Effect size was evaluated via pseduo-R^2 values for the proportion reduction in each variance component for level-2 school variances when appropriate. Odds ratios were also computed for individual slopes (exponentiated coefficients), which for negative slopes range from 0–1 and for positive slopes range from 1 to positive infinity.

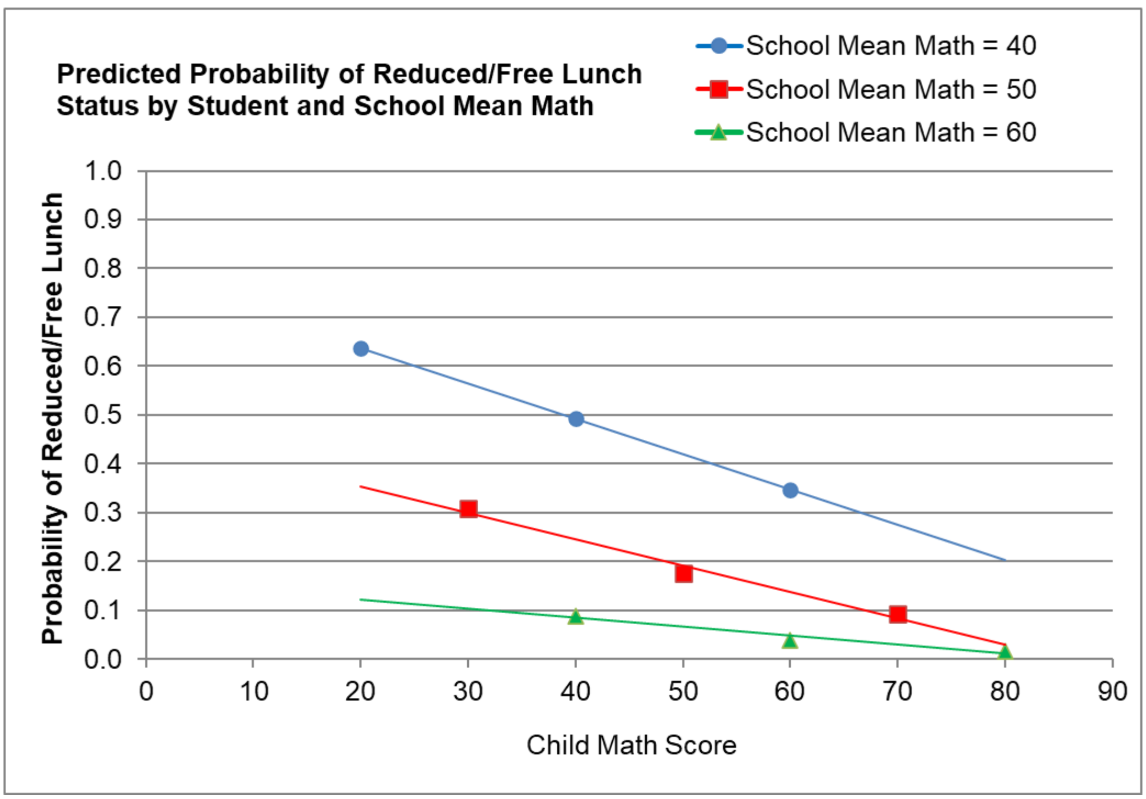
As derived from an empty means, random intercept model, student lunch status had an intraclass correlation of ICC = .373, indicating that 37.3% of the variance in lunch status was between schools, which was significant, −2ΔLL(1) = 2,973.47 , $p <$ .0001. A 95% random effects confidence interval, calculated as fixed intercept ± 1.96*SQRT(random intercept variance), revealed that 95% of the sample schools were predicted to have intercepts for school proportion reduced or free lunch between .02 and .83. The fixed intercept estimate for the logit (log-odds) of receiving reduced or free lunch in an average school (random intercept = 0) was –1.172, or probability = .236. We then examined the impact of student math scores in predicting student lunch status. Given that previous analyses had revealed that approximately 15% of the variance in math was between schools, the level-1 variance in student math was represented by cluster-mean-centering, in which the level-1 predictor was calculated by substracting the school's mean math score from each student's math score. The level-2 school variance in student math was then represented by centering the school mean math score at 50 (near the mean of the distribution). To aid the numeric stability of the solution, both predictors were rescaled by diving by 10, such that a one-unit increase indicated a 10-point increase in each level of math score.

The effect of school mean math was first added to the model. The fixed intercept indicated that the logit for getting reduced or free lunch for a child in a school with a random intercept = 0 and school mean math = 50 was −1.470, or a probability = .230. The level-2 between-school effect of math indicated that for every 10 units higher school mean math, the logit of getting reduced or free lunch was significantly lower by 1.443, which translates into an odds ratio of 0.236. This effect accounted for 60.82% of the level-2 school random intercept variance.

Next, the effect of cluster-mean-centered student math was added to the model. The fixed intercept indicated that the logit of getting reduced or free lunch for a child in a school with a random intercept = 0 and school mean math = 50 and within-school math = 0 (i.e., an average student) was −1.560, or a probability = .210. The level-1 within-school effect of math indicated that for every 10 units higher student math relative to the rest of a student's school, the logit for the probability of getting reduced or free lunch was significantly lower by 0.372, which translates into an odds ratio of 0.689. After controlling for student math, the model-implied contextual math effect (i.e., the between effect minus the within effect) of −1.517 + 0.372 = −1.145 per additional 10 points of math was still significant (odds ratio = 0.318). We then examined to what extent the within-school effect of student math varied across schools. A level-2 random slope variance for the effect of level-1 within-school math resulted in a significant improvement in model fit, −2ΔLL(2) = 38.66, $p < .001$, indicating that the size of the disadvantage related to student math differed significantly across schools. A 95% random effects confidence interval for the student math effect, calculated as fixed slope ± 1.96 *SQRT(random slope variance), revealed that 95% of the schools were predicted to have math-related slopes on the logit scale ranging from −0.59 to −0.10. Slope reliability was 0.627, as computed based on Willett (1989).

Finally, the extent to which school differences in the math-related disadvantage in predicting student lunch status could be predicted from school math scores was then examined by adding a cross-level intra-variable interaction between the student and school math predictors, as well as the quadratic effect of school math to allow comparable between-school moderation as well. The within-school student math effect was significantly moderated by school mean math (which reduced its random slope variance by 16.2%), although the moderation of the between-school and contextual effects was not significant and did not reduce the random intercept variance. The significant intra-variable cross-level interaction is shown by the nonparallel slopes of the lines in Figure 1, in which the top panel depicts predicted logit (log-odds), and the bottom panel translates those predictions in probability. The decrease in the logit for the probability of receiving reduced or free lunch per 10-point increase in within-school student math (of −0.369, as found for students with school mean math = 50), became significantly more negative by 0.070 for per 10 points of school mean math. Alternatively, the between-school school effect (of −1.583 per 10 points of school mean math in students at their school's mean) became significantly more negative by 0.070 per 10 points higher student math relative to their school's mean. Thus, the effect of relatively better math on student lunch status was more pronounced in better performing schools. The level-2 quadratic effect indicated that the between-school math effect became nonsignificantly more negative by 2*0.069 for every additional 10 points of school mean math. (see excel spreadsheet for original figures)


Predicted Logit of Reduced/Free Lunch Status by Student and School Mean Math
School Mean Math = 40
School Mean Math = 50
School Mean Math = 60
Logit of Reduced/Free Lunch
1.5
0.5
-0.5
-1.5
-2.5
-3.5
-4.5
0
10
20
30
40
50
60
70
80
90
Child Math Score


Predicted Probability of Reduced/Free Lunch Status by Student and School Mean Math
School Mean Math = 40
School Mean Math = 50
School Mean Math = 60
Probability of Reduced/Free Lunch
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
60
70
80
90
Child Math Score

Model 7. Empty Means, Single-Level Ordinal Model Predicting Lunch3: Paid Lunch (=0) vs. Reduced (=1) or Free Lunch (=2)

lunch	Frequency	Percent	Cumulative Percent
0	9059	69.25	69.25
1	1140	8.71	77.96
2	2883	22.04	100.00

Level 1: $Log\left[\frac{prob(lunch3_{pc}=1or2)}{prob(lunch3_{pc}=0)}\right] = Logit\left(lunch3_{pc} > 0\right) = \beta_{0c1}$

$Log\left[\frac{prob(lunch3_{pc}=2)}{prob(lunch3_{pc}=0or\ 1)}\right] = Logit\left(lunch3_{pc} > 1\right) = \beta_{0c2}$

Level 2: $\beta_{0c1} = -\gamma_{001}$ $\quad \beta_{0c2} = -\gamma_{002}$

Note the negative sign in front—the model returns **thresholds** (= logit of lower category), which become **intercepts** (= logit of higher category) if *** −1**. In SAS GLIMMIX, the DESCENDING option returns intercepts instead, but to the best of my knowledge this is not possible in STATA MEOLOGIT or R CLM/CLMM (from ORDINAL package).

```
display "STATA Model 7: Empty Means, Single-Level for Student Ordinal Lunch"
meologit lunch3 , nolog  // coeflegend
------------------------------------------------------------------------------
       lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        /cut1 |   .8117308   .0189462                       .774597    .8488647   logit of 0
        /cut2 |   1.263458   .0210929                      1.222117      1.3048   logit of 0 or 1
------------------------------------------------------------------------------

display "-2LL = " e(ll)*-2       // Print -2LL for model
-2LL = 20942.184
```

$$Prob(y=1) = \frac{\exp(0.8117)}{1+\exp(0.8117)} = .6925$$

```
nlcom 1/(1+exp(-1*(_b[/cut1])))  // 0 vs 12 threshold in probability
------------------------------------------------------------------------------
       lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        _nl_1 |   .6924782   .0040346   171.63   0.000     .6845705    .7003859   prob of 0
------------------------------------------------------------------------------

nlcom 1/(1+exp(-1*(_b[/cut2])))  // 01 vs 2 threshold in probability
------------------------------------------------------------------------------
       lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        _nl_1 |   .7796209    .003624   215.13   0.000     .7725179    .7867238   prob of 0 or 1
------------------------------------------------------------------------------

print("R Model 7: Empty Means, Single-Level for Student Ordinal Lunch")
print("Using clm from the ordinal package")
Model7 = clm(data=Example6a, link="logit", formula=as.factor(lunch3)~1)
print("Show -2LL and results")
-2*logLik(Model7); summary(Model7)

'log Lik.' 20942.184 (df=2) → -2LL for model

link  threshold nobs  logLik    AIC      niter max.grad cond.H
 logit flexible  13082 -10471.09 20946.18 5(1)  6.39e-09 9.0e+00

Threshold coefficients:
    Estimate Std. Error z value
0|1 0.811731   0.018946  42.844  logit of 0
1|2 1.263458   0.021093  59.900  logit of 0 or 1

print("Convert logits to probability via inverse link")
Model7Prob=1/(1+exp(-1*coefficients(Model7))); Model7Prob
       0|1        1|2
0.69247821 0.77962085
```

Model 8. Empty Means, Two-Level Ordinal Model Predicting Lunch3

Level 1: $Log\left[\frac{prob(lunch3_{pc}=1or2)}{prob(lunch3_{pc}=0)}\right] = Logit\left(lunch3_{pc} > 0\right) = \beta_{0c1}$

$Log\left[\frac{prob(lunch3_{pc}=2)}{prob(lunch3_{pc}=0or\ 1)}\right] = Logit\left(lunch3_{pc} > 1\right) = \beta_{0c2}$

Level 2: $\beta_{0c1} = -\gamma_{001} + U_{0c}$ $\quad\beta_{0c2} = -\gamma_{002} + U_{0c}$

Note that the *single* random intercept predicts the higher category—it's only the fixed threshold that refers to the logit of the lower category instead (which become intercepts when −*1).

```
display "STATA Model 8: Empty Means, Random Intercept for Student Ordinal Lunch"
meologit lunch3 , ||  schoolID:  , intpoints(7) nolog  // coeflegend
------------------------------------------------------------------------------
      lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       /cut1 |   1.203654   .1428788                      .9236164    1.483691  g001
       /cut2 |   1.767587   .1433791                      1.486569    2.048605  g002
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|   1.793843   .3054923                      1.284768    2.504635  Var(U0c)
------------------------------------------------------------------------------
LR test vs. ologit model: chibar2(01) = 2981.76        Prob >= chibar2 = 0.0000

display "-2LL = " e(ll)*-2       // Print -2LL for model
-2LL = 17960.419
```

Probabilities do not match raw data proportions anymore because they are "unit-specific" (conditional on $U_{0c} = 0$)

```
nlcom 1/(1+exp(-1*(_b[/cut1])))  // 0 vs 12 threshold in probability
------------------------------------------------------------------------------
      lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .7691741   .0253675    30.32   0.000     .7194548    .8188935  prob of 0
------------------------------------------------------------------------------
nlcom 1/(1+exp(-1*(_b[/cut2])))  // 01 vs 2 threshold in probability
------------------------------------------------------------------------------
      lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .8541573   .0178611    47.82   0.000     .8191502    .8891645  prob of 0 or 1
------------------------------------------------------------------------------
```

```
print("R Model 8: Empty Means, Random Intercept for Student Ordinal Lunch")
print("Using clmm from the ordinal package")
Model8 = clmm(data=Example6a, link="logit", nAGQ=7,
              formula=as.factor(lunch3)~1+(1|schoolID))
print("Show -2LL, results, and ICC using 3.29=residual variance")
-2*logLik(Model8); summary(Model8); icc(Model8)

'log Lik.' 17960.425 (df=3) → -2LL for model

link  threshold nobs  logLik   AIC      niter    max.grad cond.H
 logit flexible  13082 -8980.21 17966.42 103(713)  6.05e-05 3.2e+02

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept) 1.7929   1.339    Var(U0c)

No Coefficients

Threshold coefficients:
    Estimate Std. Error z value
0|1  1.20365    0.14334  8.3973  g001
1|2  1.76759    0.14384 12.2889  g002

# Intraclass Correlation Coefficient
    Adjusted ICC: 0.353
  Unadjusted ICC: 0.353
```

Model-scale ICC for the correlation of students in the same school for lunch3:

$$\text{ICC} = \frac{1.794}{1.794 + 3.29} = .353$$

```
print("LRT for Random Intercept Variance"); anova(Model8, Model7)

Likelihood ratio tests of cumulative link models:
       no.par     AIC    logLik LR.stat df Pr(>Chisq)
Model7      2 20946.2 -10471.09
Model8      3 17966.4  -8980.21 2981.76  1 < 2.22e-16

print("Convert logits to probability via inverse link")
Model8Prob=1/(1+exp(-1*coefficients(Model8))); Model8Prob
       0|1        1|2
0.76917435 0.85415731
```

Model 9. Add CMmath and Fixed Slope of Cluster-MC Student Math

Level 1: $Logit\left(lunch3_{pc} > 0\right) = \beta_{0c1} + \beta_{1c}\left(\left[math_{pc} - CMmath_{c}\right]/10\right)$

$Logit\left(lunch3_{pc} > 1\right) = \beta_{0c2} + \beta_{1c}\left(\left[math_{pc} - CMmath_{c}\right]/10\right)$

Level 2: $\beta_{0c1} = -\gamma_{001} + \gamma_{01}([CMmath_{c} - 50]/10) + U_{0c}$ $\beta_{1c} = \gamma_{10}$

$\beta_{0c2} = -\gamma_{002} + \gamma_{01}([CMmath_{c} - 50]/10) + U_{0c}$

```
display "STATA Model 9: Add CMmath and Fixed Slope of Cluster-MC Student Math"
meologit lunch3 c.CMmath50 c.WCmath, || schoolID: , intpoints(7) nolog
estimates store FixMath                 // Save LL for LRT
------------------------------------------------------------------------------
      lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    CMmath50 |  -1.393751   .1478313    -9.43   0.000    -1.683495   -1.104007   g01
      WCmath |  -.3776075   .0140071   -26.96   0.000    -.4050609   -.3501541   g10
-------------+----------------------------------------------------------------
       /cut1 |   1.569321   .1102684                      1.353199    1.785443   g001
       /cut2 |    2.17208   .1111773                      1.954176    2.389983   g002
-------------+----------------------------------------------------------------
schoolID     |
   var(_cons)|   .8713822   .1606757                      .6070914    1.250729   Var(U0c)
------------------------------------------------------------------------------
LR test vs. ologit model: chibar2(01) = 901.20        Prob >= chibar2 = 0.0000

display "-2LL = " e(ll)*-2          // Print -2LL for model
-2LL = 17116.161

lincom c.WCmath*-1 + c.CMmath50*1    // Math Contextual Slope
------------------------------------------------------------------------------
      lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |  -1.016143   .1480279    -6.86   0.000    -1.306273   -.7260139   g01-g11
------------------------------------------------------------------------------

display "STATA Model 9: Odds Ratios Instead"
meologit lunch3 c.CMmath50 c.WCmath, || schoolID: , intpoints(7) nolog or
------------------------------------------------------------------------------
      lunch3 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    CMmath50 |   .2481428   .0366833    -9.43   0.000     .1857238      .33154   exp(g01)
      WCmath |   .6854995   .0096019   -26.96   0.000     .6669362    .7045795   exp(g10)
-------------+----------------------------------------------------------------
lincom c.WCmath*-1 + c.CMmath50*1, or    // Math Contextual Slope
------------------------------------------------------------------------------
      lunch3 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .3619884   .0535844    -6.86   0.000     .2708277    .4838338   exp(g01-g10)
------------------------------------------------------------------------------
```

```
print("R Model 9: Add CMmath and Fixed Slope of Cluster-MC Student Math")
Model9 = clmm(data=Example6a, link="logit", nAGQ=7,
              formula=as.factor(lunch3)~1+CMmath50+WCmath+(1|schoolID))
print("Show -2LL, results, and odds ratios")
-2*logLik(Model9); summary(Model9); exp(coefficients(Model9))
# Below does not work and I do not know why -- I have tried every combo I can think of
#Model9glht = summary(glht(model=Model9, linfct=rbind(
#   "Math Contextual Slope and Odds Ratio"= c(0,0,1,-1))),test=adjusted("none"))
#Model9glht; data.frame(OR=exp(Model9glht$test$coefficients))
```

```
'log Lik.' 17116.162 (df=5) → -2LL for model

link  threshold nobs  logLik   AIC      niter     max.grad cond.H
 logit flexible  13082 -8558.08 17126.16 269(1426) 2.35e-03 2.3e+02

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolID (Intercept) 0.87122  0.93339  Var(U0c)
Number of groups:  schoolID 94

Coefficients:
          Estimate Std. Error  z value  Pr(>|z|)
CMmath50 -1.393787   0.147927  -9.4221 < 2.2e-16  g01
WCmath   -0.377607   0.014007 -26.9583 < 2.2e-16  g10

Threshold coefficients:
    Estimate Std. Error z value
0|1  1.56935    0.11037  14.219  g001
1|2  2.17210    0.11128  19.520  g002

       0|1        1|2   CMmath50     WCmath
4.80350093 8.77671657 0.24813384 0.68549975  → exp(g)
```

Model 10. Add Random Slope of Cluster-MC Student Math

Level 1: $Logit\left(lunch3_{pc} > 0\right) = \beta_{0c1} + \beta_{1c}\left(\left[math_{pc} - CMmath_c\right]/10\right)$

$Logit\left(lunch3_{pc} > 1\right) = \beta_{0c2} + \beta_{1c}\left(\left[math_{pc} - CMmath_c\right]/10\right)$

Level 2: $\beta_{0c1} = -\gamma_{001} + \gamma_{01}([CMmath_c - 50]/10) + U_{0c}$ $\quad\quad \beta_{1c} = \gamma_{10} + U_{1c}$

$\beta_{0c2} = -\gamma_{002} + \gamma_{01}([CMmath_c - 50]/10) + U_{0c}$

```
display "STATA Model 10: Add Random Slope of Cluster-MC Student Math"
meologit lunch3 c.CMmath50 c.WCmath, || schoolID: WCmath, cov(un) intpoints(7) nolog
-----------------------------------------------------------------------------------
           lunch3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
------------------+----------------------------------------------------------------
         CMmath50 |  -1.475822   .1498346    -9.85   0.000    -1.769493   -1.182152  g01
           WCmath |  -.3403429   .0241359   -14.10   0.000    -.3876484   -.2930373  g10
------------------+----------------------------------------------------------------
            /cut1 |   1.580968   .1089267                      1.367476    1.794461  g001
            /cut2 |   2.188412   .1098383                      1.973133    2.403692  g002
------------------+----------------------------------------------------------------
schoolID          |
       var(WCmath)|   .0168303   .0054834                      .0088872    .0318728  Var(U1c)
        var(_cons)|   .8345144   .1558512                      .5787169    1.203376  Var(U0c)
------------------+----------------------------------------------------------------
schoolID          |
 cov(WCmath,_cons)|  -.0494273   .0290298    -1.70   0.089    -.1063247    .0074702  Cov(U0,U1)
-----------------------------------------------------------------------------------
LR test vs. ologit model: chi2(3) = 945.96                Prob > chi2 = 0.0000

display "-2LL = " e(ll)*-2  // Print -2LL for model
-2LL = 17071.401
```

```
estimates store RandMath    // Save LL for LRT
lrtest RandMath FixMath     // LRT against fixed-only WCmath slope

Likelihood-ratio test  Assumption: FixMath nested within RandMath
 LR chi2(2) =  44.76
Prob > chi2 = 0.0000

print("R Model 10: Add Random Slope of Cluster-MC Student Math")
Model10 = clmm(data=Example6a, link="logit", nAGQ=1,
               formula=as.factor(lunch3)~1+CMmath50+WCmath+(1+WCmath|schoolID))
print("Show -2LL, results, and odds ratios")
-2*logLik(Model10); summary(Model10); exp(coefficients(Model10))

'log Lik.' 17072.391 (df=7) → -2LL for model

link  threshold nobs  logLik   AIC      niter     max.grad cond.H
 logit flexible  13082 -8536.20 17086.39 673(3662) 6.64e-04 2.2e+02

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 schoolID (Intercept) 0.828745 0.91035          Var(U0c)
          WCmath      0.016615 0.12890  -0.418  Var(U1c) Cor(U0,U1)

Coefficients:
         Estimate Std. Error  z value  Pr(>|z|)
CMmath50 -1.475627   0.149510  -9.8698 < 2.2e-16  g01
WCmath   -0.340459   0.024075 -14.1415 < 2.2e-16  g10

Threshold coefficients:
    Estimate Std. Error z value
0|1  1.58070    0.10869  14.544  g001
1|2  2.18811    0.10960  19.965  g002

       0|1        1|2   CMmath50     WCmath
4.85833596 8.91834113 0.22863532 0.71144363  → -2LL for model

print("Re-estimate Model 9 with same laplace method for LRT")
Model9R = clmm(data=Example6a, link="logit", nAGQ=1,
               formula=as.factor(lunch3)~1+CMmath50+WCmath+(1|schoolID))
print("LRT for Random Slope Variance"); anova(Model10, Model9R)

Likelihood ratio tests of cumulative link models:

        no.par     AIC    logLik LR.stat df       Pr(>Chisq)
Model9R      5 17126.9 -8558.43
Model10      7 17086.4 -8536.20 44.4632  2 0.00000000022128
```

Does the level-2 random slope of within-school math improve model fit? *Yes, −2ΔLL(~2) = 44.76, p < .001*

Calculate a 95% random effect confidence interval for the student math slope (STATA output):
*CI = fixed effect ± 1.96*SQRT(random slope variance)*
*CI = −0.3403 ± 1.96*SQRT(0.0168303) = −0.59 to −0.09 in logits (there is no analog in probability terms)*

Random slope reliability: $$\mathrm{SR} = \frac{\tau^2_{U_1}}{\tau^2_{U_1} + [\sigma^2_e/(L1n * var)]} = \frac{.0168303}{.0168303 + [3.29/(139 * 2.514)]} = \mathbf{.641}$$

Btw, 2.514 is the variance of the cluster-mean-centered L1 WCmath predictor that has the random slope.

So what does this mean? *The extent to which within-school student differences in math predicts student reduced/free lunch status varies significantly across schools, but across 95% of schools, higher student math is still predicted to relate to a lower probability of receiving reduced or free lunch.*

Note that these ordinal models assume proportional odds—I could not find a direct way to test it in any package without having to write a custom model (e.g., in SAS NLMIXED).

Sample Results Section for Ordinal Multilevel Models using STATA Output and INTERCEPTS instead of thresholds (i.e., after reversing the sign of the thresholds)

[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]

Overall, 69.25% of the sample students paid full-price for lunch, 8.71% paid reduced price, and 22.04% received free lunch. The extent to which student math outcomes could predict student ordinal lunch status was examined in a series of multilevel models in which the 13,802 students were modeled as nested at level 1 within their 94 schools at level 2, and school differences were captured via school-level random effects. The ordinal lunch status outcome was predicted using a cumulative logit link function and multinomial conditional outcome distribution. All model parameters were estimated via full-information marginal maximum likelihood (MML) using adaptive Gaussian quadrature with 7 points of integration per random effect dimension in STATA MELOGIT v. 17. Accordingly, all fixed effects should be interpreted as unit-specific (i.e., as the fixed effect specifically for schools in which the corresponding random effect = 0). The significance of fixed effects was evaluated with Wald tests (i.e., the *t*-test of the ratio of each estimate to its standard error using between–within denominator degrees of freedom), whereas the significance of random effects was evaluated via likelihood ratio tests (i.e., −2ΔLL with degrees of freedom equal to the number of new random effects variances and covariances). Effect size was evaluated via pseduo-R^2 values for the proportion reduction in each variance component for level-2 school variances when appropriate, as well as odds ratios for individual slopes.

As derived from an empty means, random intercept model, student lunch status had an intraclass correlation of ICC = .353, indicating that 35.3% of the variance in lunch status was between schools, which was significant, −2ΔLL(1) = 2,981.76 , $p <$ .0001. The fixed intercept estimate for the logit (log-odds) of receiving reduced or free lunch (instead of paid lunch) in an average school (random intercept = 0) was –1.204, or probability = .231. The fixed intercept estimate for the logit (log-odds) of receiving free lunch (instead of paid or reduced-price lunch) in an average school (random intercept = 0) was –1.768, or probability = .146.

We then examined the impact of student math scores in predicting student lunch status assuming proportional odds (i.e., equal slopes across submodels). Given that previous analyses had revealed that approximately 15% of the variance in math was between schools, the level-1 variance in student math was represented by cluster-mean-centering, in which the level-1 predictor was calculated by substracting the school's mean math score from each student's math score. The level-2 school variance in student math was then represented by centering the school mean math score at 50 (near the mean of the distribution). To aid the numeric stability of the solution, both predictors were rescaled by diving by 10, such that a one-unit increase indicated a 10-point increase in each level of math score.

Both cluster-mean-centered student math and school mean math (as just described) were added to the model in a single step. The fixed intercept indicated that the logit of getting reduced or free lunch (instead of paid lunch) for a child in a school with a random intercept = 0 and school mean math = 50 and within-school math = 0 (i.e., an average student) was −1.569. The fixed intercept estimate for the logit (log-odds) of receiving free lunch (instead of paid or reduced-price lunch) in an average school (random intercept = 0) was –2.172. The level-2 between-school effect of math indicated that for every 10 units higher school mean math, the logit of getting reduced or free lunch was significantly lower by 1.394, which translates into an odds ratio of 0.248. The level-1 within-school effect of math indicated that for every 10 units higher student math relative to the rest of a student's school, the logit for the probability of getting reduced or free lunch was significantly lower by 0.378, which translates into an odds ratio of 0.685. After controlling for student math, the model-implied contextual math effect (i.e., the between effect minus the within effect) of −1.393 + 0.377 = −1.016 per additional 10 points of math was still significant (odds ratio = 0.362).

We then examined to what extent the within-school effect of student math varied across schools. A level-2 random slope variance for the effect of level-1 within-school math resulted in a significant improvement in model fit, −2ΔLL(2) = 44.46, $p < .001$, indicating that the size of the disadvantage related to student math differed significantly across schools. A 95% random effects confidence interval for the student math effect, calculated as fixed slope ± 1.96 *SQRT(random slope variance), revealed that 95% of the schools were predicted to have math-related slopes on the logit scale ranging from −0.59 to −0.09. Slope reliability was 0.641, as computed based on Willett (1989). [Models with interactions would be described as for the binary outcome Model 6).