# Example 5a: Generalized Linear Mixed Models for Logistic Clustered Outcomes using SAS PROC GLIMMIX, STATA MELOGIT, and MPLUS (last model only)

These are the same real data featured in PSQF 7375 Clustered Example 3b from a $10^{th}$ grade math test in a Midwestern Rectangular State. These data include 13,802 students from 94 schools, with 31–515 students per school (M = 275). We will examine how student free and reduced lunch status (0 = pay for lunch, 1= receive free or reduced lunch) can be predicted by math test scores (i.e., the reverse of Example 3b).

## SAS Syntax for Data Import, Manipulation, and Description:

```
* Define global variable for file location to be replaced in code below;
* \\Client\ precedes actual path when using UIowa Virtual Desktop;
%LET filesave=C:\Dropbox\19_PSQF7375_Clustered\PSQF7375_Clustered_Example3b;
LIBNAME example "&filesave.";

* Import data into work library;
DATA work.grade10; SET example.grade10school;
     LABEL studentID= "studentID: Student ID number"
           schoolID=  "schoolID: School ID number"
           frlunch=   "frlunch: 0=No, 1=Free/Reduced Lunch"
           math=      "math: Math Test Score Outcome";
       * Selecting cases that are complete for analysis variables;
       IF NMISS(studentID, schoolID, frlunch, math)>0 THEN DELETE;
RUN;

* Get school means;
PROC SORT DATA=work.grade10; BY schoolID studentID; RUN;
PROC MEANS NOPRINT N DATA=work.grade10;
     BY schoolID; VAR frlunch math;
     OUTPUT OUT=work.SchoolMeans MEAN(frlunch math)= SMfrlunch SMmath;
RUN;

* Label new school mean variables;
DATA work.SchoolMeans; SET work.SchoolMeans;
     Nperschool = _FREQ_; * Saving N per school;
     DROP _TYPE_ _FREQ_;    * Dropping unneeded SAS-created variables;
     LABEL Nperschool= "Nperschool: # Students Contributing Data"
           SMfrlunch=  "SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch"
           SMmath=     "SMmath: School Mean Math Outcome";
RUN;

* Merge school means back with individual data;
DATA work.grade10; MERGE work.grade10 work.SchoolMeans; BY schoolID;
     * Arbitrarily select only schools with at least 30 students;
     IF Nperschool < 31 THEN DELETE;
     * Center math predictors;
     WSmath = (math – SMmath)/10; LABEL WSmath= "WSmath: Within-School Math (0=SM)";
     SMmath50 = (SMmath - 50)/10; LABEL SMmath50= "SMmath50: School Mean Math (0=5)";
RUN;

TITLE "School-Level Descriptives";
PROC MEANS NDEC=2 DATA=work.SchoolMeans;
     VAR Nperschool SMmath SMfrlunch;
RUN; TITLE;

TITLE "Student-Level Descriptives";
PROC MEANS NDEC=2 DATA=work.grade10;
     VAR math frlunch;
RUN; TITLE;
```

School-Level Descriptives

| Variable | Label | N | Mean | Std Dev | Minimum | Maximum |
|---|---|---|---|---|---|---|
| Nperschool | Nperschool: # Students Contributing Data | 94 | 139.17 | 138.20 | 31.00 | 515.00 |
| SMmath | SMmath: School Mean Math Outcome | 94 | 47.73 | 6.97 | 29.45 | 61.61 |
| SMfrlunch | SMfrlunch: School Mean 0=No, 1=F/R Lunch | 94 | 0.30 | 0.21 | 0.00 | 0.80 |

Student-Level Descriptives

| Variable | Label | N | Mean | Std Dev | Minimum | Maximum |
|---|---|---|---|---|---|---|
| math | math: Math Test Score Outcome | 13082 | 48.12 | 17.26 | 0.00 | 83.00 |
| frlunch | frlunch: 0=No, 1=Free/Reduced Lunch | 13082 | 0.31 | 0.46 | 0.00 | 1.00 |

## STATA Syntax for Data Import, Manipulation, and Description:

```
/ Define global variable for file location to be replaced in code below
// \\Client\ precedes actual path when using UIowa Virtual Desktop
global filesave "C:\Dropbox\19_PSQF7375_Clustered\PSQF7375_Clustered_Example3b"

// Import example stata data file
use "$filesave\grade10school.dta", clear

// Label existing variables
label variable studentID "studentID: Student ID number"
label variable schoolID  "schoolID: School ID number"
label variable frlunch   "frlunch: Student Free/Reduced Lunch 0=No 1=Yes"
label variable math      "math: Student Free/Reduced Lunch 0=No 1=Yes"

// Get school means of variables and label them
egen SMfrlunch   = mean(frlunch),   by(schoolID)
egen SMmath      = mean(math),      by(schoolID)
label variable SMfrlunch "SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch"
label variable SMmath    "SMmath: School Mean Math Outcome"

// Get number of students per school
egen Nperschool = count(studentID), by(schoolID)
label variable Nperschool "Nperschool: # Students Contributing Data"

// Center school mean math
gen SMmath50 = (SMmath-50)/10
label variable SMmath50 "SMmath: School Mean Math (0=5)"

// Center to get within-school math
gen WSmath = (math-SMmath)/10
label variable SMmath "WSmath: Within-School Math (0=SM)"

// Drop schools with <= 30 students
drop if Nperschool < 31

display as result "STATA School-Level Descriptives"
preserve  // Save for later use, then compute school-level dataset
collapse  Nperschool SMfrlunch SMmath, by(schoolID)
format    Nperschool SMfrlunch SMmath  %4.2f
summarize Nperschool SMfrlunch SMmath, format

restore   // Go back to student-level dataset
display as result "STATA Student-Level Descriptives"
format    math frlunch %4.2f
summarize math frlunch, format

// Add option "or" to model options in melogit get odds ratios for fixed effects
```

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

$$\text{Level 1: } \mathrm{Logit}\left(\mathrm{FRlunch}_{ks}=1\right)=\beta_{0s}$$
$$\text{Level 2: } \text{Intercept: } \beta_{0s}=\gamma_{00}$$

```
TITLE "SAS Empty Means, Single-Level Logistic Model Predicting Student Free/Reduced Lunch";
PROC GLIMMIX DATA=work.grade10 NOCLPRINT NAMELEN=100 METHOD=QUAD(QPOINTS=15) GRADIENT;
      CLASS schoolID;
      * Descending makes us predict the 1 instead of the default-predicted 0;
      MODEL frlunch (DESCENDING) =  / SOLUTION LINK=LOGIT DIST=BINARY DDFM=Satterthwaite;
      ESTIMATE "Intercept" intercept 1 / ILINK; * ILINK is inverse link (to un-logit);
RUN; TITLE;

display as result "STATA Model 1: Empty Means, Single-Level Logistic Model Predicting FRlunch"
melogit frlunch ,
       estat ic, n(94), // getting AIC and BIC equivalent to SAS
       nlcom 1/(1+exp(-1*(_b[_cons]))) // fixed intercept in probability
```

```
        Convergence criterion (GCONV=1E-8) satisfied.

           Fit Statistics
-2 Log Likelihood           16145.89
AIC  (smaller is better)    16147.89
AICC (smaller is better)    16147.89
BIC  (smaller is better)    16155.37
CAIC (smaller is better)    16156.37
HQIC (smaller is better)    16150.39
Pearson Chi-Square          13082.00
Pearson Chi-Square / DF         1.00
```

***What table is missing that would normally be here?***

To go from logits to probability for predicted outcomes (i.e., to apply the inverse logit link):

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

```
                              Parameter Estimates
                          Standard
Effect       Estimate        Error       DF    t Value    Pr > |t|    Gradient
Intercept     -0.8117      0.01895    13081     -42.84      <.0001    2.155E-9

                                         Estimates                          Standard
                          Standard                                             Error
Label        Estimate        Error       DF    t Value    Pr > |t|       Mean       Mean
Intercept     -0.8117      0.01895    13081     -42.84      <.0001     0.3075   0.004035
```

**What does the fixed intercept represent?**

---

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

$$\text{Level 1: } \mathrm{Logit}\left(\mathrm{FRlunch}_{ks}=1\right)=\beta_{0s}$$
$$\text{Level 2: } \text{Intercept: } \beta_{0s}=\gamma_{00}+U_{0s}$$

```
TITLE "SAS Empty Means, Two-Level Logistic Model Predicting Student Free/Reduced Lunch";
PROC GLIMMIX DATA=work.grade10 NOCLPRINT NAMELEN=100 METHOD=QUAD(QPOINTS=15) GRADIENT;
      CLASS schoolID;
      MODEL frlunch (DESCENDING) = / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
      ESTIMATE "Intercept" intercept 1 / ILINK; * ILINK is inverse link (to un-logit);
      COVTEST "Random School Intercept?" 0;     * Test if G matrix UN(1,1)=0;
      ODS OUTPUT CovParms=CovEmpty;             * Save random int var for pseudo-R2;
RUN; TITLE;
```

DDFM=Satterthwaite or KR is not available in METHOD=QUAD, so we switch to DDFM=BW (Between-Within).

```
display as result "STATA Model 2: Empty Means, Two-Level Logistic Model Predicting FRlunch"
melogit frlunch,  ||  schoolID:  , covariance(unstructured) intpoints(15),
        estat ic, n(94),
        nlcom 1/(1+exp(-1*(_b[_cons]))) // fixed intercept in probability
```

```
         Convergence criterion (GCONV=1E-8) satisfied.

           Fit Statistics
-2 Log Likelihood           13172.43
AIC  (smaller is better)    13176.43
AICC (smaller is better)    13176.43
BIC  (smaller is better)    13181.52
CAIC (smaller is better)    13183.52
HQIC (smaller is better)    13178.48

                  Covariance Parameter Estimates
Cov                                  Standard
Parm       Subject     Estimate        Error    Gradient
UN(1,1)    schoolID      1.9545       0.3315    0.000164
```

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

$$\text{ICC} = \frac{1.9545}{1.9545 + 3.29} = .3737$$

```
                         Solution for Fixed Effects
                         Standard
Effect       Estimate       Error       DF    t Value    Pr > |t|    Gradient
Intercept     -1.1721      0.1494       93      -7.85      <.0001    0.000085

                                      Estimates
                                                                          Standard
                         Standard                                            Error
Label        Estimate       Error       DF    t Value    Pr > |t|       Mean      Mean
Intercept     -1.1721      0.1494       93      -7.85      <.0001     0.2365   0.02697

                         Tests of Covariance Parameters
                            Based on the Likelihood
Label                          DF    -2 Log Like      ChiSq    Pr > ChiSq    Note
Random School Intercept?        1          16146    2973.46        <.0001    MI
MI: P-value based on a mixture of chi-squares.
```

**The COVTEST tells us whether adding the random intercept variance across schools significantly improves model fit:**
−2LL single-level = 16,145.89    −2LL two-level = 13,172.43    −2ΔLL (df=~1) = 2,973.46

COVTEST can be used for any nested model comparisons involving variance components, but I have seen it get the answer wrong, so be careful when using it!

**What does the fixed intercept NOW represent?**

To go from logits to predicted probability:

$$\text{Prob}(y = 1) = \frac{\exp(-1.1721)}{1 + \exp(-1.1721)} = .2365$$

**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 .02 to .83 in probability*

## Model 3. Adding a Level-2 Fixed Effect of School Mean Student Math

$$\text{Level 1: } \mathrm{Logit}\left(\mathrm{FRlunch}_{ks}=1\right)=\beta_{0s}$$
$$\text{Level 2: } \quad \text{Intercept: } \beta_{0s}=\gamma_{00}+\gamma_{01}\left(\mathrm{SMmath}_{s}-50\right)+U_{0s}$$

```
TITLE "SAS Add Level-2 Fixed Effect of School Mean Math";
PROC GLIMMIX DATA=work.grade10 NOCLPRINT NAMELEN=100 METHOD=QUAD(QPOINTS=15) GRADIENT;
      CLASS schoolID;
      MODEL frlunch (DESCENDING) = SMmath50 / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW ODDSRATIO;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
      ESTIMATE "Intercept if SMmath=49"  intercept 1 SMmath50 -1 / ILINK;
      ESTIMATE "Intercept if SMmath=50"  intercept 1 SMmath50  0 / ILINK;
      ESTIMATE "Intercept if SMmath=51"  intercept 1 SMmath50  1 / ILINK;
      ESTIMATE "L2 Math Slope"           SMmath50 1 / ILINK; * Example of non-sense ILINK;
      ODS OUTPUT CovParms=CovSMmath;                     * Save random int var for pseudo-R2;
RUN; TITLE; %PseudoR2G(NCov=1, CovFewer=CovEmpty, CovMore=CovSMmath);

display as result "STATA Model 3: Add Level-2 Fixed Effect of School Mean Math"
melogit frlunch c.SMmath50, || schoolID: , covariance(unstructured) intpoints(15),
        estat ic, n(94),
        margins , at(c.SMmath50=(-1(1)1)) predict(xb) // unit-specific logits
        margins , at(c.SMmath50=(-1(1)1))             // marginal probabilities
```

```
        Convergence criterion (GCONV=1E-8) satisfied.
          Fit Statistics
-2 Log Likelihood           13103.22
AIC  (smaller is better)    13109.22
AICC (smaller is better)    13109.23
BIC  (smaller is better)    13116.85
CAIC (smaller is better)    13119.85
HQIC (smaller is better)    13112.31
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate | Standard Error | Gradient |
|---|---|---|---|---|
| UN(1,1) | schoolID | 0.7657 | 0.1448 | -0.00005 |

Solutions for Fixed Effects

| Effect | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Gradient |
|---|---|---|---|---|---|---|
| Intercept | -1.4696 | 0.1040 | 92 | -14.13 | <.0001 | 0.000025 |
| SMmath50 | -1.4429 | 0.1403 | 92 | -10.29 | <.0001 | -0.00002 |

Odds Ratio Estimates

| SMmath50 | _SMmath50 | Estimate | DF | 95% Confidence Limits | |
|---|---|---|---|---|---|
| 0.8119 | -0.188 | **0.236** | 92 | 0.179 | 0.312 |

Effects of continuous variables are assessed as one unit offsets from the mean.
The AT suboption modifies the reference value and the UNIT suboption modifies the offsets.

Estimates

| Label | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Mean | Standard Error Mean |
|---|---|---|---|---|---|---|---|
| Intercept if SMmath=49 | -0.02668 | 0.1421 | 92 | -0.19 | 0.8515 | 0.4933 | 0.03552 |
| Intercept if SMmath=50 | -1.4696 | 0.1040 | 92 | -14.13 | <.0001 | 0.1870 | 0.01581 |
| Intercept if SMmath=51 | -2.9125 | 0.2020 | 92 | -14.42 | <.0001 | 0.05154 | 0.009873 |
| L2 Math Slope | -1.4429 | 0.1403 | 92 | -10.29 | <.0001 | 0.1911 | 0.02168 |

PsuedoR2 (% Reduction) for CovEmpty vs. CovSMmath

| Name | CovParm | Subject | Estimate | StdErr | Gradient | Pseudo R2 |
|---|---|---|---|---|---|---|
| CovEmpty | UN(1,1) | schoolID | 1.9545 | 0.3315 | 0.000164 | . |
| CovSMmath | UN(1,1) | schoolID | 0.7657 | 0.1448 | -0.00005 | 0.60824 |

**What does the fixed intercept NOW represent?** *The logit for the probability of getting free/reduced lunch for a kid in a school with a random intercept $U_{0s} = 0$ and school mean math = 50 is −1.4696, which is a probability = .187.*

**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 free/reduced lunch is significantly lower by 1.4429, which translates into an odds ratio of 0.236. This is the "total" between-school effect. This effect accounted for 60.824% of the level-2 school random intercept variance.*

******Note that the probability estimate of 0.1911 is meaningless, because a one-unit difference in the predictor does not imply the same difference in probability at all points along the predictor.*******

## Model 4. Adding a Level-1 Fixed Effect of Group-Mean-Centered Student Math

$$\text{Level 1: } \text{Logit}\left(\text{FRlunch}_{ks}=1\right) = \beta_{0s} + \beta_{1s}\left(\text{math}_{ks} - \text{SMmath}_s\right)$$
$$\text{Level 2: } \text{Intercept: } \beta_{0s} = \gamma_{00} + \gamma_{01}\left(\text{SMmath}_s - 50\right) + U_{0s}$$
$$\text{Within-School Math: } \beta_{1s} = \gamma_{10}$$

```
TITLE "SAS Add Level-1 Fixed Effect of Group-MC Student Math";
PROC GLIMMIX DATA=grade10 NOCLPRINT NOITPRINT NAMELEN=100 METHOD=QUAD (QPOINTS=15) GRADIENT;
      CLASS schoolID studentID;
      MODEL frlunch (DESCENDING) = SMmath50 WSmath
          / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW ODDSRATIO;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
      ESTIMATE "Between-School Contextual Effect of Math" WSmath -1 SMmath50 1;
      CONTRAST "Multivariate Wald test for Math Effects"  SMmath50 1, WSmath 1 / CHISQ;
RUN; TITLE;

display as result "STATA Model 4: Add Level-1 Fixed Effect of Group-MC Student Math"
melogit frlunch c.SMmath50 c.WSmath,  ||  schoolID:  , covariance(unstructured) intpoints(15),
        estat ic, n(94),
        estimates store FixMath,            // save LL for LRT
        lincom c.WSmath*-1 + c.SMmath50*1   // Between-School Contextual Effect of Math
```

```
        Convergence criterion (GCONV=1E-8) satisfied.
          Fit Statistics
-2 Log Likelihood           12390.67
AIC  (smaller is better)    12398.67
AICC (smaller is better)    12398.67
BIC  (smaller is better)    12408.85
CAIC (smaller is better)    12412.85
HQIC (smaller is better)    12402.78
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate | Standard Error | Gradient |
|---|---|---|---|---|
| UN(1,1) | schoolID | 0.8414 | 0.1576 | 0.000012 |

Note the increase in the level-2 random intercept variance and in the math fixed effect—it is rescaled due to the reduction of the level-1 residual variance (which stays at 3.29 no matter what).

Solutions for Fixed Effects

| Effect | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Gradient |
|---|---|---|---|---|---|---|
| Intercept | -1.5598 | 0.1088 | 92 | -14.34 | <.0001 | -0.00046 |
| SMmath50 | -1.5174 | 0.1467 | 92 | -10.35 | <.0001 | 0.00009 |
| WSmath | -0.3720 | 0.01450 | 12987 | -25.66 | <.0001 | 0.000823 |

Odds Ratio Estimates

| SMmath50 | WSmath | _SMmath50 | _WSmath | Estimate | DF | 95% Confidence Limits | |
|---|---|---|---|---|---|---|---|
| 0.8119 | -1E-17 | -0.188 | -1E-17 | 0.219 | 92 | 0.164 | 0.293 |
| -0.188 | 1 | -0.188 | -1E-17 | 0.689 | 12987 | 0.670 | 0.709 |

Estimates

| Label | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Contextual Between-School Effect of Math | -1.1454 | 0.1468 | 92 | -7.80 | <.0001 |

**What does the fixed intercept NOW represent?** *The logit for the probability of getting free/reduced lunch for a kid in a school with a random intercept $U_{0s} = 0$ and school mean math = 50 and within-school math = 0 (e.g., an average student) is −1.5598, 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 one-unit higher school mean math, the logit for the probability of getting free/reduced lunch is significantly lower by 0.1517, which translates into an odds ratio of 0.219. This effect is still significant after controlling for kid math (as indicated by a contextual between-school 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 free/reduced 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 un-estimated.*

Contrasts

| Label | Num DF | Den DF | Chi-Square | F Value | Pr > ChiSq | Pr > F |
|---|---|---|---|---|---|---|
| Multivariate Wald test for Math Effects | 2 | 12987 | **746.30** | 373.15 | <.0001 | <.0001 |

There are two ways to test multiple fixed effects at once. The above output is an example of a multivariate Wald test (from CONTRAST) that you can use for any model and with either REML or ML. Given that we are using ML here, we can also do an LRT: −2ΔLL(2) = **781.76**, $p$ < .0001. These tests should agree (asymptotically).

---

## Model 5. Adding a Random Effect of Group-MC Student Math

$$\begin{aligned}
&\text{Level 1: } \text{Logit}\left(\text{FRlunch}_{ks}=1\right)=\beta_{0s}+\beta_{1s}\left(\text{math}_{ks}-\text{SMmath}_{s}\right)\\
&\text{Level 2: } \quad \text{Intercept: } \beta_{0s}=\gamma_{00}+\gamma_{01}\left(\text{SMmath}_{s}-50\right)+U_{0s}\\
&\quad \text{Within-School Math: } \beta_{1s}=\gamma_{10}+U_{1s}
\end{aligned}$$

```
TITLE "SAS Add Random Effect of Group-MC Student Math";
PROC GLIMMIX DATA=work.grade10 NOCLPRINT NAMELEN=100 METHOD=QUAD (QPOINTS=15) GRADIENT;
       CLASS schoolID;
       MODEL frlunch (DESCENDING) = SMmath50 WSmath
          / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW ODDSRATIO;
       RANDOM INTERCEPT WSmath / TYPE=UN SUBJECT=schoolID;
       COVTEST "Random Student Math Slope?" . 0 0;    * Leave (1,1), test if (2,1) and (2,2) =0;
       ODS OUTPUT CovParms=CovRandMath; * Save random variances for pseudo-R2;
RUN; TITLE;

display as result "* STATA Model 5: Add Random Effect of Group-MC Student Math"
melogit frlunch c.SMmath50 c.WSmath,   ||  schoolID: WSmath, ///
          covariance(unstructured) intpoints(15),
         estat ic, n(94),
         estimates store RandMath   // save LL for LRT
         lrtest RandMath FixMath    // LRT against fixed effect model
```

Convergence criterion (GCONV=1E-8) satisfied.

**FROM THE LOG:** At least one element of the gradient is greater than 1e-3.

Fit Statistics

| | |
|---|---|
| -2 Log Likelihood | 12352.01 |
| AIC (smaller is better) | 12364.01 |
| AICC (smaller is better) | 12364.01 |
| BIC (smaller is better) | 12379.27 |
| CAIC (smaller is better) | 12385.27 |

```
HQIC (smaller is better)    12370.17

            Covariance Parameter Estimates
Cov                                 Standard
Parm        Subject     Estimate       Error    Gradient
UN(1,1)     schoolID      0.8118      0.1540    -0.00188
UN(2,1)     schoolID    -0.03524     0.02906    0.007376
UN(2,2)     schoolID     0.01608    0.005433    0.324555
```

Note that the level-2 random slope variance across schools for the effect of student math is not estimated very well: the gradient is the partial derivative with respect to each parameter, which should be ~0.

```
                          Solutions for Fixed Effects
                          Standard
Effect        Estimate       Error        DF    t Value     Pr > |t|     Gradient
Intercept      -1.5665      0.1076        92     -14.56      <.0001     0.003945
SMmath50       -1.5617      0.1477        92     -10.57      <.0001      -0.0015
WSmath         -0.3434     0.02425     12987     -14.16      <.0001     -0.04844

                                   Odds Ratio Estimates
                                                                   95% Confidence
SMmath50    WSmath    _SMmath50    _WSmath    Estimate       DF        Limits
  0.8119    -1E-17       -0.188     -1E-17       0.210       92      0.156       0.281
  -0.188         1       -0.188     -1E-17       0.709    12987      0.676       0.744

                          Tests of Covariance Parameters
                             Based on the Likelihood
Label                           DF    -2 Log Like      ChiSq    Pr > ChiSq    Note
Random Student Math Slope?       2          12391      38.66        <.0001    MI
MI: P-value based on a mixture of chi-squares.
```

**Does the level-2 random effect of level-1 student math improve model fit?** *Yes,* $-2\Delta LL(\sim 2) = 38.66$*,* $p < .001$

**Calculate a 95% random effect confidence interval for the student math slope:**
*CI = fixed effect ± 1.96*SQRT(random slope variance)*
*CI = −0.3434 ± 1.96*SQRT(0.01608) = −0.59 to –0.09 in logits (there is no analog in probability terms)*

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

---

## Model 6. Adding Intra-Variable Interactions of School Mean Math and GMC Student Math

$$\text{Level 1: } \text{Logit}\left(\text{FRlunch}_{ks}=1\right)=\beta_{0s}+\beta_{1s}\left(\text{math}_{ks}-\text{SMmath}_s\right)$$

$$\text{Level 2: Intercept: } \beta_{0s}=\gamma_{00}+\gamma_{01}\left(\text{SMmath}_s-50\right)+\gamma_{02}\left(\text{SMmath}_s-50\right)^2+U_{0s}$$

$$\text{Within-School Math: } \beta_{1s}=\gamma_{10}+\gamma_{11}\left(\text{SMmath}_s-50\right)+U_{1s}$$

```
TITLE "SAS Add Intra-Variable Interactions of School Mean and Group-MC Student Math";
PROC GLIMMIX DATA=work.grade10 NOCLPRINT NAMELEN=100 METHOD=QUAD (QPOINTS=15) GRADIENT;
      CLASS schoolID;
      MODEL frlunch (DESCENDING) = SMmath50 WSmath SMmath50*WSmath SMmath50*SMmath50
                  / SOLUTION LINK=LOGIT DIST=BINARY DDFM=BW ODDSRATIO;
      RANDOM INTERCEPT WSmath / TYPE=UN SUBJECT=schoolID;
      ESTIMATE "Contextual Math Main Effect" WSmath -1 SMmath50 1;
      ESTIMATE "Contextual Math Interaction" SMmath50*WSmath -1 SMmath50*SMmath50 1;
CONTRAST "Multiv Wald test for Interactions"  SMmath50*WSmath 1, SMmath50*SMmath50 1 / CHISQ;
RUN; TITLE; %PseudoR2G(NCov=3, CovFewer=CovRandMath, CovMore=CovInteract);
```

```
display as result "STATA Model 6: Add Intra-Variable Interactions of School Mean Math and GMC
Student Math"
melogit frlunch c.SMmath50 c.WSmath c.SMmath50#c.WSmath c.SMmath50#c.SMmath50, ///
            || schoolID: WSmath, covariance(unstructured) intpoints(15),
        estat ic, n(94),
        lincom c.WSmath*-1 + c.SMmath50*1                       // Contextual Math Main Effect
        lincom c.SMmath50#c.WSmath*-1 + c.SMmath50#c.SMmath50*1  // Contextual Math Interaction
```

Convergence criterion (GCONV=1E-8) satisfied.

**FROM THE LOG:** At least one element of the gradient is greater than 1e-3.

Fit Statistics

| | |
|---|---|
| -2 Log Likelihood | 12347.84 |
| AIC (smaller is better) | 12363.84 |
| AICC (smaller is better) | 12363.86 |
| BIC (smaller is better) | 12384.19 |
| CAIC (smaller is better) | 12392.19 |
| HQIC (smaller is better) | 12372.06 |

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate | Standard Error | Gradient |
|---|---|---|---|---|
| UN(1,1) | schoolID | 0.8157 | 0.1553 | -0.00526 |
| UN(2,1) | schoolID | -0.02773 | 0.02798 | -0.05393 |
| UN(2,2) | schoolID | 0.01348 | 0.004909 | **0.332867** |

LRT agrees closely with tests of two new interactions: $-2\Delta LL(2) = \mathbf{4.17}$, $p = .124$

Contrasts

| Label | Num DF | Den DF | Chi-Square | F Value | Pr > ChiSq | Pr > F |
|---|---|---|---|---|---|---|
| Multiv Wald test for Interactions | 2 | 91 | **4.35** | 2.18 | 0.1133 | 0.1192 |

Solutions for Fixed Effects

| Effect | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Gradient |
|---|---|---|---|---|---|---|
| Intercept | -1.5460 | 0.1231 | 91 | -12.55 | <.0001 | 0.003075 |
| SMmath50 | -1.5833 | 0.1998 | 91 | -7.93 | <.0001 | -0.00215 |
| WSmath | -0.3688 | 0.02633 | 12986 | -14.01 | <.0001 | -0.10677 |
| SMmath50*WSmath | -0.06962 | 0.03364 | 12986 | -2.07 | **0.0385** | **0.055708** |
| SMmath50*SMmath50 | -0.06850 | 0.1760 | 91 | -0.39 | **0.6980** | **0.0059** |

Odds Ratio Estimates

| SMmath50 | WSmath | _SMmath50 | _WSmath | Estimate | DF | 95% Confidence Limits | |
|---|---|---|---|---|---|---|---|
| 0.8119 | -1E-17 | -0.188 | -1E-17 | 0.197 | 91 | 0.111 | 0.348 |
| -0.188 | 1 | -0.188 | -1E-17 | 0.701 | 12986 | 0.668 | 0.735 |

Estimates

| Label | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Contextual Math Main Effect | -1.2145 | 0.1994 | 91 | -6.09 | <.0001 |
| Contextual Math Interaction | 0.001114 | 0.1772 | 91 | 0.01 | 0.9950 |

PsuedoR2 (% Reduction) for CovRandMath vs. CovInteract

| Name | CovParm | Subject | Estimate | StdErr | Gradient | PseudoR2 |
|---|---|---|---|---|---|---|
| CovRandMath | UN(1,1) | schoolID | 0.8118 | 0.1540 | -0.00188 | . |
| CovRandMath | UN(2,2) | schoolID | 0.01608 | 0.005433 | 0.324555 | . |
| CovInteract | UN(1,1) | schoolID | 0.8157 | 0.1553 | -0.00526 | -0.00479 |
| CovInteract | UN(2,2) | schoolID | 0.01348 | 0.004909 | 0.332867 | 0.16163 |

**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 free/reduced lunch (which is −0.3688 as evaluated at school mean math = 50) becomes significantly more negative by 0.06962. So the effect of being "smarter than the others" is even stronger in a "smart" school, which accounted for 16.162% 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?** *<u>Without controlling for student math</u>, for every 10 units higher school mean math, the effect of school mean math on school mean free/reduced lunch (which is −1.5833 as evaluated at school mean math = 50) becomes nonsignificantly more negative by 2*0.06850. 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?** *<u>After controlling for student math</u>, there is a contextual effect of school mean math 1.2145 per 10 units as evaluated at school mean math = 50 for an average student. 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 1.2145 as evaluated at school mean math = 50 for an average student. However, school mean math does not moderate the between-school math effect (–0.06850) differently than the within-school math effect (–0.06962).*

## Sample Results Section using SAS Output

Overall, 30.75% of the sample students received free or reduced lunch; the proportion of students receiving free or reduced lunch in each school ranged from 0 to 80.33%. The extent to which student math outcomes could predict student free/reduced 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 15 points of integration per random effect dimension in SAS GLIMMIX. 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., $-2\Delta LL$ 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.

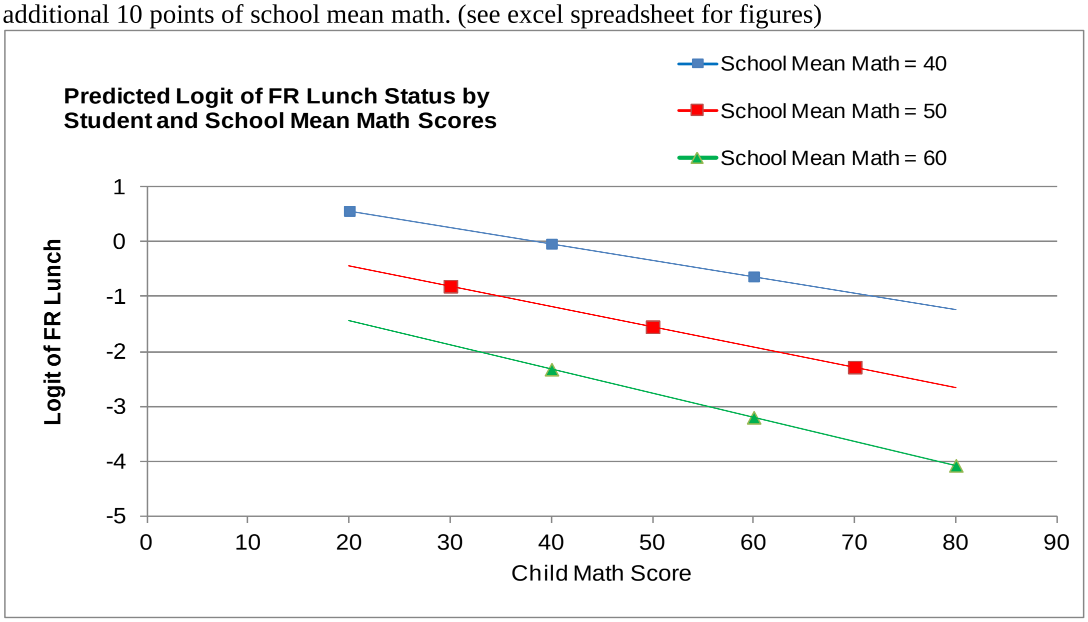
As derived from an empty means, random intercept model, student lunch status had an intraclass correlation of ICC = .373, indicating that 37.3% of the variance in lunch status was between schools, which was significant, $-2\Delta LL(1) = 2{,}973.46$ , $p <$ .0001. A 95% random effects confidence interval, 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 free or reduced lunch between .02 and .83. The fixed intercept estimate for the logit (log-odds) of receiving free or reduced lunch in an average school (random intercept = 0) was –1.172, or probability = .237. 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 group-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 in numeric stability, 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.

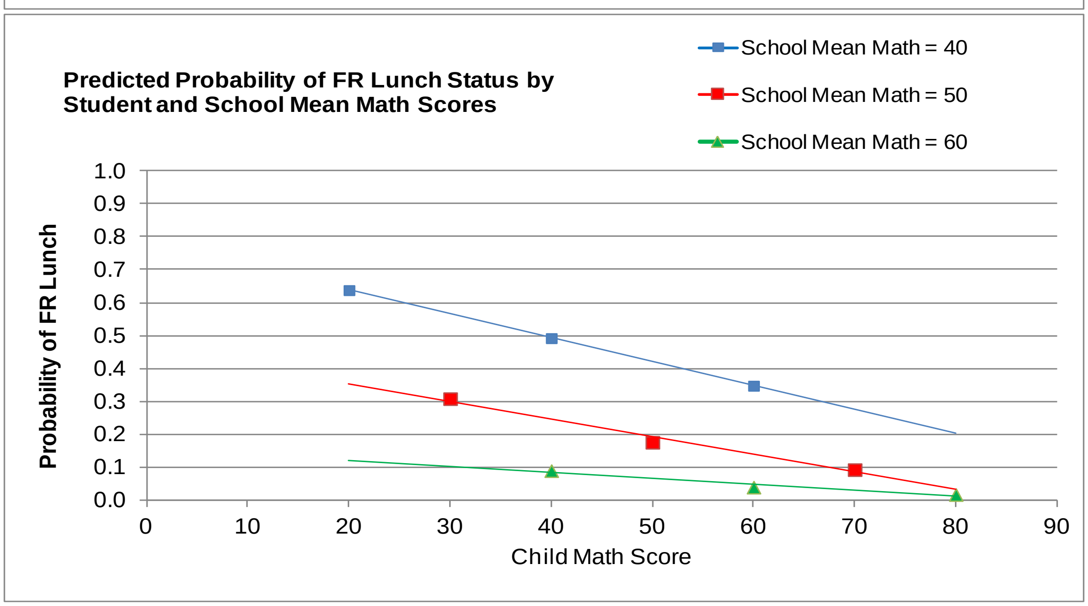
The effect of school mean math was first added to the model. The fixed intercept indicated that the logit for getting free or reduced lunch for a child in a school with a random intercept = 0 and school mean math = 50 was −1.470, or a probability = .187. The total between-school effect of math indicated that for every 10 units higher school mean math, the logit of getting free/reduced lunch was significantly lower by 1.4429, which translates into an odds ratio of 0.236. This effect accounted for 60.824% of the level-2 school random intercept variance.

Next, the effect of group-mean-centered student math was added to the model. The fixed intercept indicated that the logit of getting free or reduced 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 total within-school effect of math indicated that for every 10 units higher student math relative to the rest of your school, the logit for the probability of getting free/reduced lunch was significantly lower by 0.372, which translates into an odds ratio of 0.689.

After controlling for student math, the contextual between-school math effect of −1.145 per additional 10 points of math was still significant. 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 student math resulted in a significant improvement in model fit, $-2\Delta LL(2) = 38.66$, $p < .001$, 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 −0.09.

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 control for a contextual interaction effect. 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 free or reduced lunch per unit increase in within-school student math of of 3.69, as found for students with school mean math = 50, became significantly more negative by 0.070 for every additional 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 0.069 for every additional 10 points of school mean math. (see excel spreadsheet for figures)



## Mplus Syntax and Output for final model (using observed variables as predictors rather than latent)—results are very similar to SAS:

```
TITLE:  2-Level Model for Students within Schools Predicting FR Lunch;
DATA:    FILE = grade10M.csv;       ! Can just list file if in same directory;
         FORMAT = free;             ! FREE or FIXED format;
         TYPE = individual;         ! Individual or matrix data as input;

VARIABLE:
! List of ALL variables in stacked data file, in order;
! Mplus does NOT know what they used to be called, though;
    NAMES ARE Student School BvG FRlunch Math smvG smFR smMath SchoolN
              smBvG50 smFR30 WSmath smMath50;
! List of ALL variables used in model (DEFINED variables at end);
    USEVARIABLES ARE FRlunch WSmath smMath50 smMath2;
! Outcomes that are binary/ordinal;
    CATEGORICAL ARE FRlunch;
! Missing data codes (here, -999);
    MISSING ARE ALL (-999);
! Identify upper-level nesting;
    CLUSTER = School;
! Predictor variables with variation ONLY within at level 1;
    WITHIN = WSmath;
! Predictor variables with variation ONLY between at level 2;
    BETWEEN = smMath50 smMath2;

DEFINE:     smMath2 = smMath50*smMath50; ! Creating level-2 math quadratic;
ANALYSIS:   TYPE IS TWOLEVEL RANDOM;     ! 2-level model with random slopes;
            ESTIMATOR IS ML;             ! Can also use MLR for non-normality;

MODEL:
!!! MODEL 6
! Level-1, student-level model;
%WITHIN%
! NO residual variance is estimated for FRlunch at level 1;
    L1math | FRlunch ON WSmath;      ! B1s effect of 0/1 level-1 math;
! Level-2, school-level model;
%BETWEEN%
    FRlunch;                         ! Random intercept variance (is default);
    [FRlunch$1];                     ! Fixed "threshold" (is intercept*-1);
    [L1math]              (L1math);  ! Fixed WS effect of level-1 math;
    L1math;                          ! Yes random effect of level-1 math;
    FRlunch WITH L1math;             ! Covariance of intercept and math slope;
    FRlunch ON smMath50   (L2math);  ! Linear BS math on intercept;
    FRlunch ON smMath2    (L2math2); ! Quad BS math on intercept;
    L1math  ON smMath50   (L12math); ! Cross-level L1 by L2 math interaction;


!!!!! Adding NEW statements to show how to get ESTIMATE-type statements;
MODEL CONSTRAINT:
! Define new parameters not directly given by model;
NEW (conM conMint);
conM    = L2math - L1math;           ! Contextual main effect of math;
conMint = L2math2 - L12math;         ! Contextual L2 interaction of math;
```

```
UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

    FRLUNCH
      Category 1    0.692      9059.000
      Category 2    0.308      4023.000

THE MODEL ESTIMATION TERMINATED NORMALLY
MODEL FIT INFORMATION

Number of Free Parameters                        8
Loglikelihood
          H0 Value                       -6173.936

Information Criteria
          Akaike (AIC)                   12363.871
          Bayesian (BIC)                 12423.703
          Sample-Size Adjusted BIC       12398.280
            (n* = (n + 2) / 24)


MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
Within Level

Between Level

 L1MATH     ON
    SMMATH50          -0.069      0.034     -2.065      0.039

 FRLUNCH    ON
    SMMATH50          -1.587      0.200     -7.952      0.000
    SMMATH2           -0.083      0.176     -0.472      0.637

 FRLUNCH  WITH
    L1MATH            -0.027      0.028     -0.972      0.331

 Intercepts
    L1MATH            -0.369      0.026    -14.099      0.000

 Thresholds
    FRLUNCH$1          1.526      0.123     12.443      0.000

 Residual Variances
    FRLUNCH            0.813      0.155      5.251      0.000
    L1MATH             0.013      0.005      2.729      0.006

 New/Additional Parameters
    CONM              -1.218      0.199     -6.115      0.000
    CONMINT           -0.014      0.177     -0.077      0.939
```