Example 6a: Two-Level Clustered Data Example: Students within Schools

(only syntax and output available for SAS, SPSS, and STATA electronically; last model also in Mplus)

These are real data from a math test given at the end of 10th grade in a Midwestern Rectangular State. These data include 13,802 students from 94 schools, with 31–515 students in each school (M = 275). We will examine how student free and reduced lunch status (0 = pay for lunch, 1= receive free or reduced lunch) predicts math test scores.

SAS Code for Data Manipulation:

```
* Importing data into work library;
DATA work.grade10; SET example.grade10;
       * Selecting cases that are complete for analysis variables;
       WHERE NMISS(studentID, schoolID, frlunch, math)=0;
       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"; RUN;

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

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

* Merging school means back with individual data;
DATA work.grade10; MERGE work.grade10 work.SchoolMeans; BY schoolID;
       * Selecting only schools with data from at least 30 students;
       IF SchoolN < 31 THEN DELETE; RUN;

TITLE "Getting means to center predictors with";
PROC MEANS MEAN STDDEV MIN MAX DATA=work.grade10;
       VAR math frlunch SMmath SMfrlunch SchoolN; RUN; TITLE;

* Centering school mean predictors;
DATA work.grade10; SET work.grade10;
       SMfrlunch30 = SMfrlunch - .30; LABEL SMfrlunch30= "SMfrlunch30: 0=.30"; RUN;
```

SPSS Code for Data Manipulation:

```
* SPSS code to import data and create/center predictors.
DATASET NAME grade10 WINDOW=FRONT.
VARIABLE LABELS
       studentID     "studentID: Student ID number"
       schoolID      "schoolID: School ID number"
       frlunch       "frlunch: 0=No, 1=Free/Reduced Lunch"
       math          "math: Math Test Score".

* Selecting complete cases for analysis.
SELECT IF (NMISS(studentID, schoolID, frlunch, math)=0).
EXECUTE.
```

```
* Getting school means to use as level-2 predictors - SPSS 14+ can merge them back automatically.
SORT CASES BY schoolID studentID.
AGGREGATE
       /OUTFILE=* MODE=ADDVARIABLES
       /PRESORTED
       /BREAK = schoolID
       /SMfrlunch = MEAN(frlunch)
       /SMmath = MEAN(math)
       /SchoolN = N.

* Labeling new school mean variables.
VARIABLE LABELS
       SMfrlunch    "SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch"
       SMmath       "SMmath: School Mean Math Outcome"
       SchoolN      "SchoolN: # Students Contributing Data".

* Selecting schools with data from at least 30 students.
SELECT IF (SchoolN GT 30).

* Descriptive statistics.
DESCRIPTIVES VARIABLES=math frlunch SMmath SMfrlunch SchoolN
  /STATISTICS=MEAN STDDEV MIN MAX.

* Centering school mean predictor.
COMPUTE SMfrlunch30 = SMfrlunch - .30.
VARIABLE LABELS SMfrlunch30 "SMfrlunch30: 0=.30".
EXECUTE.
```

STATA Code for Data Manipulation:

```
* 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 SchoolN = count(studentID), by (schoolID)
label variable SchoolN= "SchoolN: # Students Contributing Data"

* then drop schools with <= 30 students
drop if SchoolN < 31

* get means to center with
summarize math frlunch SMmath SMfrlunch SchoolN

* centering school mean predictor
gen SMfrlunch30 = SMfrlunch - .30
label variable SMfrlunch30 "SMfrlunch30: Percentage Students with Free Lunch (0=30%)"
```

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        math |     13082    48.11856    17.25905          0         83
     frlunch |     13082    .3075218     .461485          0          1
      SMmath |     13082    48.11856     6.81813   29.45098   61.61364
   SMfrlunch |     13082    .3075218    .2220852          0   .8032787
     SchoolN |     13082    274.9502    155.3319         31        515
```

<u>Model 1: Two-Level Empty Means, Random Intercept for Math Outcome</u>

$$\begin{aligned}&\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + e_{ij}\\&\text{Level 2: } \beta_{0j} = \gamma_{00} + U_{0j}\end{aligned}$$

```
TITLE1 "SAS Model 1: 2-Level Empty Means, Random Intercept for Math";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
      CLASS schoolID studentID;
      MODEL math = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
      ODS OUTPUT CovParms=CovEmpty InfoCrit=FitEmpty; RUN;

TITLE "SPSS Model 1: 2-Level Empty Means, Random Intercept for Math".
MIXED math BY schoolID studentID
      /METHOD = ML
      /PRINT  = SOLUTION TESTCOV
      /FIXED  =
      /RANDOM = INTERCEPT | SUBJECT(schoolID) COVTYPE(UN).

* STATA Model 1: 2-Level Empty Means, Random Intercept for Math
mixed math  , || schoolID: , ///
      variance ml covariance(un) residuals(independent),
      estat ic, n(94),
      estat icc         // intraclass correlation
```

SAS output:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	schoolID	44.9335	7.0391	6.38	<.0001
Residual		253.18	3.1415	80.59	<.0001

Calculate the ICC for the correlation of students in the same school for math:

$$ICC = \frac{44.94}{44.94 + 253.18} = .15$$

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	1857.08	<.0001

This LR test tells us that the random intercept variance is significantly greater than 0, and thus so is the ICC.

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
109791	3	109797	109797	109800	109805	109808

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	47.7561	0.7192	94.9	66.40	<.0001

Design effect using mean #students per school: = 1 + ((n – 1) * ICC) → 1 + [(275−1)*.15] = **42.1**

Effective sample size: $N_{effective}$ = (#Total Obs) / Design Effect → 13,082 / 42.1 = **311!!!**

95% random effect confidence interval for the intercept across schools:
Fixed effect ± 1.96*SQRT(variance)
48 ± 1.96*SQRT(45) = 35 to 61 → 95% of schools are predicted to have school mean math from 35 to 61

Model 2: Adding a Fixed Effect of Student Free/Reduced Lunch (Level 1)

Level 1: $\text{Math}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{FRlunch}_{ij}\right) + e_{ij}$

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

Free/Reduced Lunch: $\beta_{1j} = \gamma_{10}$

```
TITLE1 "SAS Model 2: Adding Fixed Effect of Student Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
       CLASS schoolID studentID;
       MODEL math = frlunch / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovFR1 InfoCrit=FitFR1; RUN;

TITLE "SPSS 2: Adding Fixed Effect of Student Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch
       /METHOD = ML
       /PRINT  = SOLUTION TESTCOV
       /FIXED  = frlunch
       /RANDOM = INTERCEPT | SUBJECT(schoolID) COVTYPE(UN).

* STATA Model 2: Adding Fixed Effect of Student Free/Reduced Lunch
mixed math c.frlunch, || schoolID: , ///
       variance ml covariance(un) residuals(independent),
       estat ic, n(94)
```

SAS output:

```
                  Covariance Parameter Estimates
                                   Standard         Z
Cov Parm      Subject     Estimate      Error     Value      Pr > Z
UN(1,1)       schoolID     26.8873     4.4382      6.06      <.0001
Residual                    239.33     2.9700     80.58      <.0001

                            Information Criteria
Neg2LogLike     Parms        AIC       AICC       HQIC        BIC       CAIC
     109016         4     109024     109024     109028     109034     109038

                   Solution for Fixed Effects
                         Standard
Effect        Estimate      Error       DF    t Value    Pr > |t|
Intercept      50.6161     0.5766       98      87.78      <.0001
frlunch        -9.4316     0.3318     13E3     -28.43      <.0001

* Calculate PseudoR2 relative to empty model;
%PseudoR2(NCov=2, CovFewer=CovEmpty, CovMore=CovFR1);

PsuedoR2 (% Reduction) for CovEmpty vs. CovFR1
                                                                               Pseudo
  Name       CovParm      Subject      Estimate      StdErr     ZValue      ProbZ       R2
CovEmpty     UN(1,1)      schoolID      44.9335      7.0391       6.38     <.0001       .
CovEmpty     Residual                    253.18      3.1415      80.59     <.0001       .
CovFR1       UN(1,1)      schoolID      26.8873      4.4382       6.06     <.0001     0.40162
CovFR1       Residual                    239.33      2.9700      80.58     <.0001     0.05469
```

What does the effect of student free/reduced lunch represent in model 2?

Children who get free/reduced lunch are predicted to score 9.43 points lower in math than children who don't.

What are we assuming about the effect of student free/reduced lunch in model 2?

We are assuming no contextual effect (that the between-school and within-school effects of FRlunch are equal).

Model 3: Adding a Fixed Effect of School Proportion Free/Reduced Lunch (Level 2)

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{FRlunch}_{ij}\right) + e_{ij}$$

$$\text{Level 2: } \quad \text{Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}\left(\overline{\text{SchoolFRLunch}}_j - .30\right) + U_{0j}$$

$$\text{Free/Reduced Lunch: } \beta_{1j} = \gamma_{10}$$

```
TITLE1 "SAS Model 3: Adding Fixed Effect of School Proportion Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
       CLASS schoolID studentID;
       MODEL math = frlunch SMfrlunch30 / SOLUTION DDFM=Satterthwaite OUTPM=work.LunchSave;
       RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovFR2 InfoCrit=FitFR2;
       ESTIMATE "FR Lunch Between-School Effect"  frlunch 1 SMfrlunch30 1; RUN;
PROC CORR NOSIMPLE DATA=work.LunchSave; VAR math pred; RUN;

TITLE "SPSS Model 3: Adding Fixed Effect of School Proportion Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
       /METHOD = ML
       /PRINT  = SOLUTION TESTCOV
       /FIXED  = frlunch SMfrlunch30
       /RANDOM = INTERCEPT | SUBJECT(schoolID) COVTYPE(UN)
        /SAVE = FIXPRED(lunchpred)
        /TEST = "FR Lunch Between-School Effect" frlunch 1 SMfrlunch30 1.
CORRELATIONS /VARIABLES = math lunchpred.

* STATA Model 3: Adding Fixed Effect of School Proportion Free/Reduced Lunch
mixed math c.frlunch c. SMfrlunch30, || schoolID: , ///
       variance ml covariance(un) residuals(independent),
       estat ic, n(94),
       predict lunchpred, xb,            // save fixed-effect predicted outcomes
       estimates store FixFRLunch,       // save LL for LRT
       lincom 1*frlunch + 1*SMfrlunch30 // FR lunch between-school effect
corr math lunchpred                      // calculate total R2
```

SAS output:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	schoolID	13.4819	2.5421	5.30	<.0001
Residual		239.40	2.9716	80.56	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108965	5	108975	108975	108980	108988	108993

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.6054	0.4341	93.7	116.57	<.0001
frlunch	-9.1729	0.3344	13E3	-27.43	<.0001
SMfrlunch30	-16.8502	2.0007	84.9	-8.42	<.0001

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
FR Lunch Between-School Effect	-26.0231	1.9725	80.2	-13.19	<.0001

What does the effect of school proportion free/reduced lunch represent in model 3?
This is the level-2 contextual effect for FRlunch: holding child lunch status constant, for every 10% more children in your school who get free/reduced lunch, school mean math is predicted to be lower by 1.69 points. Before controlling for individual kid lunch status, the reduction is 2.60 points per 10% (the level-2 between-school effect, given in separate estimate).

What does the effect of student free/reduced lunch NOW represent in model 3?
This is the pure within-school effect: holding school lunch status constant, children who receive free/reduced lunch are predicted to score 9.17 points lower in math than children who don't.

```
      Pearson Correlation Coefficients, N = 13082
               Prob > |r| under H0: Rho=0
                                        math          Pred
math                                 1.00000       0.40382
math: Math Test Score Outcome                      <.0001
```

R = .4038, so total R^2 ~ .163

```
* Calculate PseudoR2 relative to previous model;
%PseudoR2(NCov=2, CovFewer=CovFR1, CovMore=CovFR2);
```

PsuedoR2 (% Reduction) for CovFR1 vs. CovFR2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
CovFR1	UN(1,1)	schoolID	26.8873	4.4382	6.06	<.0001	.
CovFR1	Residual		239.33	2.9700	80.58	<.0001	.
CovFR2	UN(1,1)	schoolID	13.4819	2.5421	5.30	<.0001	0.49858
CovFR2	Residual		239.40	2.9716	80.56	<.0001	-0.00029

```
* Calculate PseudoR2 relative to empty model (total for FRlunch);
%PseudoR2(NCov=2, CovFewer=CovEmpty, CovMore=CovFR2);
```

Total reduction from **both** lunch effects:
Intercept variance → 69.99% (of 15%)
Residual variance → 5.44% (of 85%)

PsuedoR2 (% Reduction) for CovEmpty vs. CovFR2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
CovEmpty	UN(1,1)	schoolID	44.9335	7.0391	6.38	<.0001	.
CovEmpty	Residual		253.18	3.1415	80.59	<.0001	.
CovFR2	UN(1,1)	schoolID	13.4819	2.5421	5.30	<.0001	0.69996
CovFR2	Residual		239.40	2.9716	80.56	<.0001	0.05442

Model 4: Adding a Random Effect of Student Free/Reduced Lunch (over Schools)

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{FRlunch}_{ij}\right) + e_{ij}$$

$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}\left(\overline{\text{SchoolFRLunch}}_j - .30\right) + U_{0j}$$

$$\text{Free/Reduced Lunch: } \beta_{1j} = \gamma_{10} + U_{1j}$$

```
TITLE1 "SAS Model 4: Adding Random Effect of Student Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
       CLASS schoolID studentID;
       MODEL math = frlunch SMfrlunch30 / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT frlunch / G TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovFR2RandFR1 InfoCrit=FitFR2RandFR1; RUN;

TITLE "SPSS Model 4: Adding Random Effect of Student Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
       /METHOD = ML
       /PRINT  = SOLUTION TESTCOV G
       /FIXED  = frlunch SMfrlunch30
       /RANDOM = INTERCEPT frlunch | SUBJECT(schoolID) COVTYPE(UN).
```

```
* STATA Model 4: Adding Random Effect of Student Free/Reduced Lunch
mixed math c.frlunch c. SMfrlunch30, || schoolID: frlunch,  ///
       variance ml covariance(un) residuals(independent),
       estat recovariance, relevel(schoolID),
       estat ic, n(94),
       estimates store RandFRLunch       // save LL for LRT
       lrtest RandFRLunch FixFRLunch     // LRT against fixed effect model
```

SAS output:

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	schoolID	19.9147	3.7405	5.32	<.0001	random intercept variance
UN(2,1)	schoolID	-11.9055	3.1625	-3.76	0.0002	intercept-lunch covariance
UN(2,2)	schoolID	12.6853	3.3090	3.83	<.0001	random slope variance for frlunch
Residual		236.84	2.9468	80.37	<.0001	residual variance

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108877	7	108891	108891	108899	108909	108916

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.2593	0.5144	84.9	97.70	<.0001
frlunch	-8.4501	0.5611	98.7	-15.06	<.0001
SMfrlunch30	-17.0867	1.9157	77.3	-8.92	<.0001

```
* Calculate difference in model fit relative to fixed-FRlunch-only model;
%FitTest(FitFewer=FitFR2, FitMore=FitFR2RandFR1);
```

Likelihood Ratio Test for FitFR2 vs. FitFR2RandFR1

Is model 4 better than model 3?
Yes, $-2\Delta LL(2) = 87, p < .0001$

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitFR2	108965	5	108975	108988	.	.	.
FitFR2RandFR1	108877	7	108891	108909	87.4448	2	0

So what does this mean about the effect of student free/reduced lunch?
The difference in math between kids who get free/reduced lunch and kids who don't varies significantly over schools.

95% random effects CI for the random FRlunch slope: → $-8.45 \pm 1.96*SQRT(12.69) = -15.43$ *to* -1.47
On average, the gap in math related to lunch status is 8.45 points, but across 95% of the schools, that gap is predicted to be anywhere from 1.47 to 15.43 points.

Model 5: Adding a Cross-Level Interaction of Student by School Free/Reduced Lunch

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{FRlunch}_{ij}\right) + e_{ij}$$
$$\text{Level 2: } \quad \text{Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}\left(\overline{\text{SchoolFRLunch}}_j - .30\right) + U_{0j}$$
$$\text{Free/Reduced Lunch: } \beta_{1j} = \gamma_{10} + \gamma_{11}\left(\overline{\text{SchoolFRLunch}}_j - .30\right) + U_{1j}$$

```
TITLE1 "SAS Model 5: Adding Cross-Level Interaction of Student by School Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
       CLASS schoolID studentID;
       MODEL math = frlunch SMfrlunch30 frlunch*SMfrlunch30 / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT frlunch / TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovInt1 InfoCrit=FitInt1; RUN;

TITLE "SPSS Model 5: Adding Cross-Level Interaction of Student by School Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
       /METHOD = ML
       /PRINT  = SOLUTION TESTCOV
       /FIXED  = frlunch SMfrlunch30 frlunch*SMfrlunch30
       /RANDOM = INTERCEPT frlunch | SUBJECT(schoolID) COVTYPE(UN).

* STATA Model 5: Adding Cross-Level Interaction of Student by School Free/Reduced Lunch
mixed math c.frlunch c.smfrlunch30 c.frlunch#c.smfrlunch30, ///
       || schoolID: frlunch, variance ml covariance(un) residuals(independent),
       estat ic, n(94)
```

SAS output:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	schoolID	19.8071	3.6954	5.36	<.0001
UN(2,1)	schoolID	-11.3587	3.0847	-3.68	0.0002
UN(2,2)	schoolID	11.7963	3.1631	3.73	<.0001
Residual		236.83	2.9465	80.38	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108875	8	108891	108891	108899	108911	108919

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.2228	0.5139	86.6	97.73	<.0001
frlunch	-8.6880	0.5673	113	-15.32	<.0001
SMfrlunch30	-19.4595	2.4725	89.9	-7.87	<.0001
frlunch*SMfrlunch30	4.1377	2.6329	104	1.57	0.1191

```
* Calculate PseudoR2 for interaction relative to random FRlunch;
%PseudoR2(NCov=4, CovFewer=CovFR2RandFR1, CovMore=CovInt1);
```

PsuedoR2 (% Reduction) for CovFR2RandFR1 vs. CovInt1

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
CovFR2RandFR1	UN(1,1)	schoolID	19.9147	3.7405	5.32	<.0001	.
CovFR2RandFR1	UN(2,2)	schoolID	12.6853	3.3090	3.83	<.0001	.
CovFR2RandFR1	Residual		236.84	2.9468	80.37	<.0001	.
CovInt1	UN(1,1)	schoolID	19.8071	3.6954	5.36	<.0001	0.005401
CovInt1	UN(2,2)	schoolID	11.7963	3.1631	3.73	<.0001	0.070080
CovInt1	Residual		236.83	2.9465	80.38	<.0001	0.000056

What does the effect of student free/reduced lunch NOW represent in model 5?
This is the difference between kids who get free/reduced lunch and those who don't in schools where 30% of the kids get free/reduced lunch: those kids who get free/reduced lunch are predicted to be lower in math by 8.69.

What does the effect of school proportion free/reduced lunch NOW represent in model 5?
This is the level-2 contextual (incremental between-school) effect for a kid who does not receive free/reduced lunch: for those kids, for every 10% more kids in their school that receive free/reduced lunch, their school mean math is predicted to be lower by 1.94.

What does the cross-level interaction of student by school free/reduced lunch represent in model 5?
The effect of being a kid who receives free/reduced lunch is reduced nonsignificantly by 0.41 for every 10% more children in their school who get free/reduced lunch. But this effect is currently smushed—it assumes without testing that school FRlunch moderates the within-school and between-school effects of FRlunch to the same extent.

Model 6: Adding a Level-2 Interaction of Quadratic School Free/Reduced Lunch

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{FRlunch}_{ij}\right) + e_{ij}$$

$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}\left(\overline{\text{SchoolFRLunch}}_j - .30\right) + \gamma_{02}\left(\overline{\text{SchoolFRLunch}}_j - .30\right)^2 + U_{0j}$$

$$\text{Free/Reduced Lunch: } \beta_{1j} = \gamma_{10} + \gamma_{11}\left(\overline{\text{SchoolFRLunch}}_j - .30\right) + U_{1j}$$

```
TITLE1 "SAS Model 6: Adding Level-2 Interaction of Quadratic School Free/Reduced Lunch";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
       CLASS schoolID studentID;
       MODEL math = frlunch SMfrlunch30 frlunch*SMfrlunch30 SMfrlunch30*SMfrlunch30
                    / SOLUTION DDFM=Satterthwaite OUTPM=work.TotalSave;
       RANDOM INTERCEPT frlunch / TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovInt2 InfoCrit=FitInt2;
ESTIMATE "FR Lunch Between-School Main Effect" frlunch 1 SMfrlunch30 1;
ESTIMATE "FR Lunch Between-School Interaction" frlunch*SMfrlunch30 1 SMfrlunch30*SMfrlunch30 1;
RUN; PROC CORR NOSIMPLE DATA=work.TotalSave; VAR math pred; RUN;
```

```
TITLE "SPSS Model 6: Adding Level-2 Interaction of Quadratic School Free/Reduced Lunch".
MIXED math BY schoolID studentID WITH frlunch SMfrlunch30
       /METHOD = ML
       /PRINT  = SOLUTION TESTCOV
       /FIXED  = frlunch SMfrlunch30 frlunch*SMfrlunch30 SMfrlunch30*SMfrlunch30
       /RANDOM = INTERCEPT frlunch | SUBJECT(schoolID) COVTYPE(UN)
       /SAVE = FIXPRED(totalpred)
 /TEST = "FR Lunch Between-School Main Effect" frlunch 1 SMfrlunch30 1
 /TEST = "FR Lunch Between-School Interaction" frlunch*SMfrlunch30 1 SMfrlunch30*SMfrlunch30 1.
CORRELATIONS /VARIABLES = math totalpred.
```

```
* STATA Model 6: Adding Level-2 Interaction of Quadratic School Free/Reduced Lunch
mixed math c.frlunch c.SMfrlunch30 c.frlunch#c.SMfrlunch30 c.SMfrlunch30#c.SMfrlunch30, ///
       || schoolID: frlunch, variance ml covariance(un) residuals(independent),
       estat ic, n(94),
       predict totalpred,xb,                             // save fixed-effect predicted outcomes
lincom 1*c.frlunch + 1*c.SMfrlunch30                     // FR lunch between-school main effect
lincom 1*c.frlunch#c.SMfrlunch30 + 1*c.SMfrlunch30#c.SMfrlunch30    // FR lunch BS interaction
margins, at(c.frlunch=(0 1) c.SMfrlunch30=(-.2 0 .2 .4)) vsquish      // create predicted values
marginsplot, noci name(predicted_lunch, replace) xdimension(frlunch)  // plot predicted, no CI
corr math totalpred                                                   // calculate total R2
```

SAS output:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	schoolID	18.9359	3.5682	5.31	<.0001
UN(2,1)	schoolID	-10.9387	3.0306	-3.61	0.0003
UN(2,2)	schoolID	11.8139	3.1752	3.72	<.0001
Residual		236.82	2.9465	80.38	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108872	9	108890	108890	108900	108913	108922

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.8588	0.6394	104	79.54	<.0001
frlunch	-8.8358	0.5768	115	-15.32	<.0001
SMfrlunch30	-17.9865	2.5944	88.3	-6.93	<.0001
frlunch*SMfrlunch30	5.4271	2.7642	108	1.96	0.0522
SMfrlunch30*SMfrlunch30	-14.1873	8.8055	88.3	-1.61	0.1107

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
FR Lunch Between-School Main Effect	-26.8224	2.6022	90.7	-10.31	<.0001
FR Lunch Between-School Interaction	-8.7603	8.4064	76.1	-1.04	0.3007

What does the cross-level interaction of student by school free/reduced lunch NOW represent?
The effect of being a kid who receives free/reduced lunch (now after allowing for differential moderation across levels of the effects of free/reduced lunch at both levels by school mean free/reduced lunch) is reduced significantly-ish by 0.54 for every 10% more children in their school who get free/reduced lunch.

What does the level-2 interaction of quadratic school free/reduced lunch represent?
*After controlling for kid free/reduced lunch status, the contextual (incremental between-school) effect of school mean free/reduced lunch (as evaluated at 30% FRlunch here) becomes nonsignificantly more negative by 2*1.42 for every 10% more kids in their school with free/reduced lunch.*

*If we don't control for kid free/reduced lunch, the between-school effect of −2.68 per 10% of school mean free/reduced lunch (as evaluated at 30% FRlunch here) becomes nonsignificantly more negative by 2*0.88 for every 10% more kids in their school with free/reduced lunch.*

So school mean free/reduced lunch moderates the within-school FRlunch effect, but not the contextual (incremental between-school) or between-school effects.

Pearson Correlation Coefficients, N = 13082
Prob > |r| under H0: Rho=0

	math	Pred
math	1.00000	0.40513
math: Math Test Score Outcome		<.0001

R = .4051, so total R^2 = .164

```
* Calculate PseudoR2 relative to level-1 lunch interaction only model;
%PseudoR2(NCov=4, CovFewer=CovInt1, CovMore=CovInt2);
```

PsuedoR2 (% Reduction) for CovInt1 vs. CovInt2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
CovInt1	UN(1,1)	schoolID	19.8071	3.6954	5.36	<.0001	.
CovInt1	UN(2,2)	schoolID	11.7963	3.1631	3.73	<.0001	.
CovInt1	Residual		236.83	2.9465	80.38	<.0001	.
CovInt2	UN(1,1)	schoolID	18.9359	3.5682	5.31	<.0001	0.043982
CovInt2	UN(2,2)	schoolID	11.8139	3.1752	3.72	<.0001	-0.001498
CovInt2	Residual		236.82	2.9465	80.38	<.0001	0.000021

```
* Calculate PseudoR2 for both interactions relative to main effects only model;
%PseudoR2(NCov=4, CovFewer=CovFR2RandFR1, CovMore=CovInt2);
```

Total reduction from **both** interactions:
Intercept variance → 4.92%
Lunch slope variance → 6.87%
Residual variance → 0.01%

PsuedoR2 (% Reduction) for CovFR2RandFR1 vs. CovInt2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
CovFR2RandFR1	UN(1,1)	schoolID	19.9147	3.7405	5.32	<.0001	.
CovFR2RandFR1	UN(2,2)	schoolID	12.6853	3.3090	3.83	<.0001	.
CovFR2RandFR1	Residual		236.84	2.9468	80.37	<.0001	.
CovInt2	UN(1,1)	schoolID	18.9359	3.5682	5.31	<.0001	0.049146
CovInt2	UN(2,2)	schoolID	11.8139	3.1752	3.72	<.0001	0.068687
CovInt2	Residual		236.82	2.9465	80.38	<.0001	0.000077

Sample Results Section (note that "smushed" models are not reported)…

The extent to which student free/reduced lunch status could predict student math outcomes was examined in a series of multilevel models in which the 13,802 students were modeled as nested within their 94 schools. Maximum likelihood (ML) was used in estimating and reporting all model parameters. The significance of fixed effects was evaluated with individual Wald tests (i.e., the t-test of the ratio of each estimate to its standard error using Satterthwaite denominator degrees of freedom), whereas random effects were 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, as well as with total R^2, the squared correlation between the actual math outcomes and the math outcomes predicted by the fixed effects.

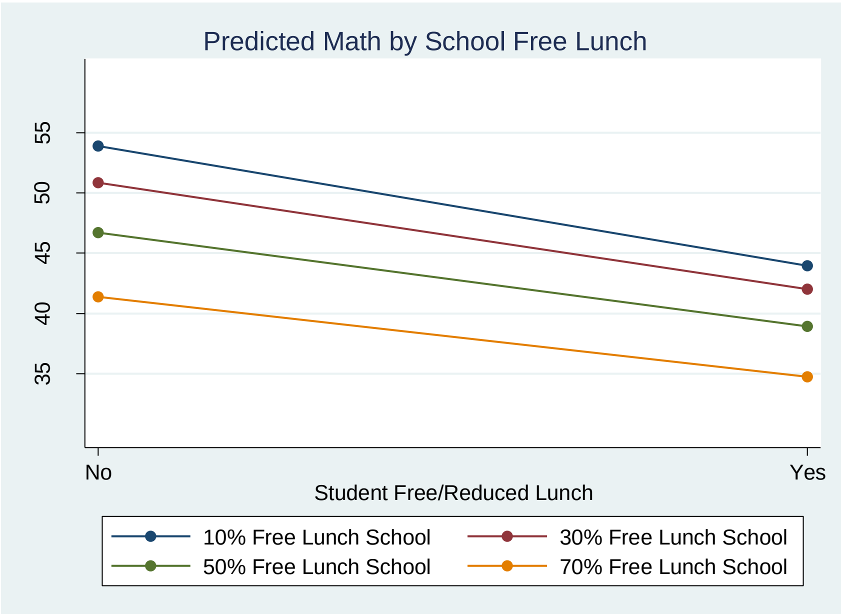
As derived from an empty means, random intercept model, student math scores had an intraclass correlation of .15, indicating that 15% of the variance in math scores was between schools, a significant amount, $-2\Delta LL(1) = 1857.08$, $p < .0001$. A 95% random effects confidence interval, calculated as fixed intercept ± 1.96* SQRT(random intercept variance), revealed that 95% of the sample schools were predicted to have intercepts for school mean math scores between 35 to 61. Children who did not receive free/reduced lunch were treated as the reference group. Given the large variability across schools in the proportion of students who received free/reduced lunch (0–80% of students), a contextual effect at level 2 was represented by the school proportion of students who receive free/reduced lunch centered near the sample mean of 30%.

The effects of free/reduced lunch status at each level were then added to the model. The within-school effect was significant and accounted for 5.44% of the residual variance, and indicated that students who receive free/reduced lunch are expected to have lower math scores than other students in their school by 9.17. The between-school effect was also significant and accounted for 70% of the random intercept variance, and indicated that for every additional 10% of students who receive free/reduced lunch, that school's mean math score is expected to be lower by 2.60. After controlling for student free/reduced lunch, the contextual free/reduced lunch effect of −1.69 per additional 10% of students was still significant. A random slope for the effect of free/reduced lunch also resulted in a significant improvement in model fit, −2ΔLL(2) = 87.4, $p < .001$, indicating that the size of the disadvantage related to free/reduced lunch differed significantly across schools. A 95% random effects confidence interval for the student free/reduced lunch effect, calculated as fixed slope ± 1.96*SQRT(random slope variance), revealed that 95% of the schools were predicted to have lunch-related gaps between students ranging from −15.43 to −1.47.

The extent to which school differences in the lunch-related disadvantage in math could be predicted from school lunch composition was then examined by adding a cross-level intra-variable interaction between the student and school lunch predictors, as well as the quadratic effect of school lunch composition to control for a contextual interaction effect. The within-school lunch effect was significantly moderated by school lunch composition (which reduced its random slope variance by 6.87%), although the moderation of the between-school and contextual effects was not significant, reducing the random intercept variance by another 4.92%, for a total R^2 = .164.

The significant intra-variable cross-level interaction, as shown by the nonparallel slopes of the lines in Figure 1, indicated that the lunch-related disadvantage in math scores of 8.84, as found for students receiving free/reduced lunch in schools in which 30% of students received free/reduced lunch, became significantly less negative by 0.54 for every additional 10% of students who received free/reduced lunch. Alternatively, the contextual school effect of −1.80 per 10% free/reduced lunch students (in baseline students in schools with 30% free/reduced lunch students) was reduced by 0.54 in free/reduced lunch students. The level-2 quadratic effect, seen by the widening distance between the lines in Figure 1, indicated that the same contextual school effect became nonsignificantly more negative by 1.42 for every additional 10% free/reduced lunch students (i.e., controlling for student lunch status), or that the between-school effect of −2.68 per 10% students became nonsignificantly more negative by 0.88 per 10% students (i.e., not controlling for student lunch status).

Figure 1: Plot of model-predicted math by free/reduced lunch status



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

```
TITLE:  2-Level Model for Students within Schools predicting student math;

DATA:    FILE = grade10.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;
! List of ALL variables used in model (DEFINED variables at end);
    USEVARIABLES ARE FRlunch Math smFR30 smFR302;
! 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 = FRlunch;
! Predictor variables with variation ONLY between at level 2;
    BETWEEN = smFR30 smFR302;

DEFINE:     smFR302 = smFR30*smFR30;      ! Creating level-2 FRlunch 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%
    math;                          ! Residual variance (is default);
    L1lunch | math ON FRlunch;     ! B1i effect of 0/1 level-1 FRlunch;
! Level-2, school-level model;
%BETWEEN%
    math;                          ! Random intercept variance (is default);
    [math];                        ! Fixed intercept (is default);
    [L1lunch]          (L1lunch);  ! Fixed WS effect of level-1 FRlunch;
    L1lunch;                       ! Yes random effect of level-1 FRlunch;
    math WITH L1lunch;             ! Covariance of intercept & FRlunch slope;
    math ON smFR30     (L2lunch);  ! Linear contextual FRlunch on intercept;
    math ON smFR302    (L2lunch2); ! Quad contextual FRlunch on intercept;
    L1lunch ON smFR30  (L12lunch); ! Cross-level L1 by L2 lunch interaction;

!!!!! Adding NEW statements to show how to get ESTIMATE-type statements;
MODEL CONSTRAINT:
! Define new parameters not directly given by model;
NEW (BSmainFR BSintFR);
BSmainFR = L1lunch + L2lunch;          ! BS main effect of FRlunch;
BSintFR = L12lunch + L2lunch2;         ! BS L2 interaction of FRlunch;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                        9

Loglikelihood

          H0 Value                      -54436.244

Information Criteria

          Akaike (AIC)                  108890.488
          Bayesian (BIC)                108957.799
          Sample-Size Adjusted BIC      108929.198
            (n* = (n + 2) / 24)



MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 Residual Variances
    MATH             236.814      2.946     80.376      0.000

Between Level

 L1LUNCH    ON
    SMFR30             5.326      2.774      1.920      0.055

 MATH       ON
    SMFR30           -17.995      2.599     -6.924      0.000
    SMFR302          -14.049      8.912     -1.576      0.115

 MATH     WITH
    L1LUNCH          -10.935      3.045     -3.591      0.000

 Intercepts
    MATH              50.851      0.644     78.970      0.000
    L1LUNCH           -8.811      0.586    -15.043      0.000

 Residual Variances
    MATH              18.953      3.574      5.303      0.000
    L1LUNCH           11.904      3.209      3.709      0.000

New/Additional Parameters
    BSMAINFR         -26.806      2.608    -10.280      0.000
    BSINTFR           -8.723      8.514     -1.024      0.306
```