Two-Level Clustered Data Example: Students within Schools
COMPLETED VERSION

These are real data taken from the results of 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 gender (0=boy, 1=girl) and student free and reduced lunch status (0=pay for lunch, 1= receive free or reduced lunch) predict math test scores.

SAS Code for Data Manipulation:

```
* Importing data into work library;
%LET example = F:\Example Data\School Data;
LIBNAME example "&example.";
DATA grade10; SET example.grade10;
LABEL  studentID=    "studentID: Student ID number"
       schoolID=     "schoolID: School ID number"
       boyvsgirl=    "boyvsgirl: Boy=0, Girl=1"
       frlunch=      "frlunch: 0=No, 1=Free/Reduced Lunch"
       math=         "math: Math Test Score Outcome";
       KEEP studentID--math;
       * Selecting cases that are complete for analysis variables;
       IF NMISS(studentID, schoolID, boyvsgirl, frlunch, math)>0 THEN DELETE; RUN;
* Getting school means to use as predictors;
PROC SORT DATA=grade10; BY schoolID studentID; RUN;
PROC MEANS NOPRINT N DATA=grade10;
       BY schoolID;
       VAR boyvsgirl frlunch math;
       OUTPUT OUT=SchoolMeans
              MEAN(boyvsgirl frlunch math)= SMboyvsgirl SMfrlunch SMmath;  RUN;
* Labeling new school mean variables;
DATA SchoolMeans; SET SchoolMeans;
       SchoolN = _FREQ_;    * Saving N per school;
       DROP _TYPE_ _FREQ_; * Dropping unneeded SAS-created variables;
LABEL  SMboyvsgirl= "SMboyvsgirl: School Mean Boy=0, Girl=1"
       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 grade10; MERGE grade10 SchoolMeans; BY schoolID;
       * Selecting only schools with data from at least 30 students;
       IF SchoolN < 31 THEN DELETE; RUN;
* Outputting table of descriptives to rtf document;
ODS RTF FILE="&example.\Descriptive Stats.rtf";
TITLE "Getting means to center predictors with";
PROC MEANS N MEAN STDDEV MIN MAX DATA=grade10;
       VAR math boyvsgirl frlunch SMmath SMboyvsgirl SMfrlunch SchoolN;
RUN; TITLE; ODS RTF CLOSE;
```

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
math	math: Math Test Score Outcome	13082	48.1185599	17.2590473	0	83.0000000
boyvsgirl	boyvsgirl: Boy=0, Girl=1	13082	0.4981654	0.5000157	0	1.0000000
frlunch	frlunch: 0=No, 1=Free/Reduced Lunch	13082	0.3075218	0.4614850	0	1.0000000
SMmath	SMmath: School Mean Math Outcome	13082	48.1185599	6.8181301	29.4509804	61.6136364
SMboyvsgirl	SMboyvsgirl: School Mean Boy=0, Girl=1	13082	0.4981654	0.0422383	0.3333333	0.6842105
SMfrlunch	SMfrlunch: School Mean 0=No, 1=Free/Reduced Lunch	13082	0.3075218	0.2220852	0	0.8032787
SchoolN	SchoolN: # Students Contributing Data	13082	274.9501605	155.3319041	31.0000000	515.0000000

```
* Centering school mean predictors;
DATA grade10; SET grade10;
       SMboyvsgirl50 = SMboyvsgirl - .50; LABEL SMboyvsgirl50= "SMboyvsgirl50: 0=.50";
       SMfrlunch30 = SMfrlunch - .30;      LABEL SMfrlunch30=   "SMfrlunch30: 0=.30"; RUN;
```

```
* Location for .sas file that holds macro programs;
%LET macropath=F:\Example Data;
%INCLUDE "&macropath.\MLM_Macros.sas";
```

Model 1a: Single-Level Empty Means, Residual Variance Only Model

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

```
TITLE1 "Model 1a: Single Level Empty Means, Residual Variance Only";
PROC MIXED DATA=example10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
      CLASS schoolID studentID;
      MODEL math = / SOLUTION DDFM=Satterthwaite; RUN;
```

Dimensions

Covariance Parameters	1
Columns in X	1
Columns in Z	0
Subjects	1
Max Obs Per Subject	13082

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr > Z
Residual	297.85	3.6828	80.88	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
111648	2	111652	111652	111657	111667	111669

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	48.1186	0.1509	13E3	318.90	<.0001

Model 1b: Empty Means, Random Intercept Model

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

```
TITLE1 "Model 1b: 2-Level Empty Means, Random Intercept";
PROC MIXED DATA=example10 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;
```

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z Per Subject	1
Subjects	94
Max Obs Per Subject	515

Null Model Likelihood Ratio Test

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

Calculate the ICC (correlation of students in same school in math): ICC = 45 / (45+253) = **.15**

Calculate the design effect: = 1 + [(n – 1) * ICC]
DE = 1 + [(275-1)*.15] = **42.1**

Calculate the effective N: $N_{effective}$ = (#Total Obs) / DE
13,082 / 42.1 = **311!!!**

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

Calculate the 95% random effects CI for the intercept across schools:
Fixed effect ± 1.96*SQRT(variance)
48 ± 1.96*SQRT(45) = **35 to 61**

Model 2a: Adding a Fixed Effect of Student Gender

$$\begin{array}{l} \text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyvsGirl}_{ij}) + e_{ij} \\ \text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + U_{0j} \\ \quad \text{Student Gender: } \beta_{1j} = \gamma_{10} \end{array}$$

```
TITLE1 "Model 2a: Adding a Fixed Effect of Student Gender";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
      CLASS schoolID studentID;
      MODEL math = boyvsgirl / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
      ODS OUTPUT CovParms=CovGen1 InfoCrit=FitGen1;
RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	schoolID	44.8203	7.0210	6.38	<.0001
Residual		253.00	3.1394	80.59	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
109782	4	109790	109790	109794	109800	109804

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	47.3300	0.7319	102	64.66	<.0001
boyvsgirl	0.8462	0.2791	13E3	3.03	0.0024

What does the main effect of student gender represent in this model?

Girls score an average of 0.8462 higher than boys.

What are we assuming about the main effect of student gender?

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

Proportion reduction in each variance relative to the 2-level empty model:

```
* Calculate PseudoR2 relative to empty means model;
%PseudoR2(NCov=2, CovFewer=CovEmpty, NameFewer=Empty, CovMore=CovGen1, NameMore=GenderL1);
```

PsuedoR2 (% Reduction) for Empty vs. GenderL1

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
Empty	UN(1,1)	schoolID	44.9335	7.0391	6.38	<.0001	.
Empty	Residual		253.18	3.1415	80.59	<.0001	.
GenderL1	UN(1,1)	schoolID	44.8203	7.0210	6.38	<.0001	.002520558
GenderL1	Residual		253.00	3.1394	80.59	<.0001	.000690265

Why were both variances reduced when gender is a level-1 predictor?

The smushed effect of gender contains part of the level-2 gender effect, too.

Model 2b: Adding a Fixed Effect of School Proportion of Girls

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

Level 2: Intercept: $\beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_j) + U_{0j}$

Student Gender: $\beta_{1j} = \gamma_{10}$

```
TITLE1 "Model 2b: Adding Fixed Effect of School Proportion Girls";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
      CLASS schoolID studentID;
      MODEL math = boyvsgirl SMboyvsgirl50 / SOLUTION DDFM=Satterthwaite OUTPM=PredGen2;
      RANDOM INTERCEPT  / TYPE=UN SUBJECT=schoolID;
      ODS OUTPUT CovParms=CovGen2 InfoCrit=FitGen2;
      ESTIMATE "Gender Between-School Effect"  boyvsgirl 1 SMboyvsgirl50 1;
RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	schoolID	43.4390	6.7859	6.40	<.0001
Residual		253.00	3.1392	80.59	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
109779	5	109789	109789	109794	109801	109806

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	47.2605	0.7229	103	65.37	<.0001
boyvsgirl	0.8352	0.2791	13E3	2.99	0.0028
SMboyvsgirl50	20.8313	11.9611	103	1.74	0.0846

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Gender Between-School Effect	21.6665	11.9578	103	1.81	0.0729

What does the main effect of school mean gender represent in this model?

After controlling for kid gender, there is no contextual (incremental between-school) effect of gender; the additional increment to school mean math scores for a one-unit difference in proportion of students who are girls of 20.83 (the difference between 0 and 100% girls) is nonsignificant. If we don't control for kid gender, the between-school gender effect of 21.67 is still nonsignificant.

Proportion reduction in each variance due to effect of school mean gender:

```
* Calculate PseudoR2 relative to level-1 gender only model;
%PseudoR2(NCov=2, CovFewer=CovGen1, NameFewer=GenderL1, CovMore=CovGen2, NameMore=GenderL2);
```

PsuedoR2 (% Reduction) for GenderL2 vs. GenderL2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
GenderL1	UN(1,1)	schoolID	44.8203	7.0210	6.38	<.0001	.
GenderL1	Residual		253.00	3.1394	80.59	<.0001	.
GenderL2	UN(1,1)	schoolID	43.4390	6.7859	6.40	<.0001	0.030817
GenderL2	Residual		253.00	3.1392	80.59	<.0001	0.000021

```
* Calculate Total R2; PROC CORR NOSIMPLE DATA=PredGen2; VAR math pred; RUN;
```

	math	Pred
math	1.00000	0.03016
math: Math Test Score Outcome		0.0006

R = .03016, so total R^2 = .009

Model 2c: Adding a Random Effect of Student Gender

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyVsGirl}_{ij}) + e_{ij}$$
$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_{j}) + U_{0j}$$
$$\text{Student Gender: } \beta_{1j} = \gamma_{10} + U_{1j}$$

```
TITLE1 "Model 2c: Adding Random Effect of Student Gender";
PROC MIXED DATA=work.grade10 NOCLPRINT COVTEST NAMELEN=100 IC METHOD=ML;
      CLASS schoolID studentID;
      MODEL math = boyvsgirl SMboyvsgirl50 / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT boyvsgirl / G TYPE=UN SUBJECT=schoolID;
      ODS OUTPUT CovParms=CovGenRand InfoCrit=FitGenRand;
RUN;
```

Estimated G Matrix

Row	Effect	School ID number	Col1	Col2
1	Intercept	125	44.0992	-0.7533
2	boyvsgirl	125	-0.7533	0.6203

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	schoolID	44.0992	7.1632	6.16	<.0001
UN(2,1)	schoolID	-0.7533	2.1120	-0.36	0.7213
UN(2,2)	schoolID	0.6203	0.8957	0.69	0.2443
Residual		252.84	3.1432	80.44	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
109778	7	109792	109792	109799	109810	109817

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	47.2521	0.7289	94.9	64.82	<.0001
boyvsgirl	0.8553	0.3005	55.1	2.85	0.0062
SMboyvsgirl50	20.5425	11.9666	103	1.72	0.0890

Does the random slope for student gender help the model? *No, model fit is not better.*

```
* Calculate difference in model fit relative to both fixed gender effects model;
%FitTest(FitFewer=FitGen2, NameFewer=FixedGender, FitMore=FitGenRand, NameMore=RandomGender);
```

Likelihood Ratio Test for FixedGender vs. RandomGender

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FixedGender	109779	5	109789	109801	.	.	.
RandomGender	109778	7	109792	109810	0.63311	2	0.72865

So what does this mean about the effect of student gender?

The 0.86 advantage for girls is approximately the same across schools.

Given the non-significant improvement in model fit, we will remove the random slope for gender. However, we will continue to control for gender at both levels of the model.

Model 3a: Adding a Fixed Effect of Student Free/Reduced Lunch

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyVsGirl}_{ij}) + \beta_{2j}(\text{FRlunch}_{ij}) + e_{ij}$$
$$\text{Level 2: } \quad \text{Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_j) + U_{0j}$$
$$\text{Student Gender: } \beta_{1j} = \gamma_{10}$$
$$\text{Free/Reduced Lunch: } \beta_{2j} = \gamma_{20}$$

```
TITLE1 "Model 3a: 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 = boyvsgirl SMboyvsgirl50 frlunch / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovFR1 InfoCrit=FitFR1;
RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	schoolID	26.1636	4.3105	6.07	<.0001
Residual		239.14	2.9675	80.59	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
109003	6	109015	109015	109022	109031	109037

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.1325	0.5868	110	85.44	<.0001
boyvsgirl	0.8555	0.2714	13E3	3.15	0.0016
SMboyvsgirl50	15.0796	9.5669	104	1.58	0.1180
frlunch	-9.4298	0.3316	13E3	-28.43	<.0001

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

Kids who get free/reduced lunch score 9.4 points lower than kids who don't.

What are we assuming about the main effect of student free/reduced lunch?

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

Proportion reduction in each variance relative to Model 2b with both fixed gender effects:

```
* Calculate PseudoR2 relative to both fixed gender effects model;
%PseudoR2(NCov=2, CovFewer=CovGen2, NameFewer=Gender, CovMore=CovFR1, NameMore=FRLunch1);
```

PsuedoR2 (% Reduction) for Gender vs. FRLunch1

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
Gender	UN(1,1)	schoolID	43.4390	6.7859	6.40	<.0001	.
Gender	Residual		253.00	3.1392	80.59	<.0001	.
FRLunch1	UN(1,1)	schoolID	26.1636	4.3105	6.07	<.0001	0.39769
FRLunch1	Residual		239.14	2.9675	80.59	<.0001	0.05478

Why were both variances reduced when FRlunch is a level-1 predictor?

The smushed effect of free/reduced lunch contains part of the free/reduced lunch level-2 effect, too.

Model 3b: Adding a Fixed Effect of School Proportion Free/Reduced Lunch

Level 1: $\text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyVsGirl}_{ij}) + \beta_{2j}(\text{FRlunch}_{ij}) + e_{ij}$

Level 2: Intercept: $\beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_j) + \gamma_{02}(\overline{\text{SchoolFRLunch}}_j) + U_{0j}$

Student Gender: $\beta_{1j} = \gamma_{10}$

Free/Reduced Lunch: $\beta_{2j} = \gamma_{20}$

```
TITLE1 "Model 3b: 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 = boyvsgirl SMboyvsgirl50 frlunch SMfrlunch30
              / SOLUTION DDFM=Satterthwaite OUTPM=PredLunch2;
       RANDOM INTERCEPT / TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovFR2 InfoCrit=FitFR2;
    ESTIMATE "FR Lunch Between-School Effect" frlunch 1 SMfrlunch30 1;
RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	schoolID	13.3767	2.5214	5.31	<.0001
Residual		239.21	2.9693	80.56	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108954	7	108968	108968	108975	108986	108993

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.1574	0.4542	113	110.44	<.0001
boyvsgirl	0.8549	0.2714	13E3	3.15	0.0016
SMboyvsgirl50	5.6834	7.4083	101	0.77	0.4448
frlunch	-9.1757	0.3343	13E3	-27.45	<.0001
SMfrlunch30	-16.5983	2.0143	83.8	-8.24	<.0001

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
FR Lunch Between-School Effect	-25.7740	1.9863	79.2	-12.98	<.0001

What does the main effect of school proportion free/reduced lunch represent in this model?

This is the contextual (incremental between-school) effect for FRlunch: after controlling for kid free/reduced lunch status, for every 10% more kids in their school who receive free/reduced lunch, school mean math is lower by 1.66. If we don't control for kid FRlunch, the between-school effect of FRlunch of −2.5.77 per 10% is still significant.

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

This is the pure within-school effect: holding school free/reduced lunch status constant (i.e., within the same school), kids who receive free/reduced lunch score 9.2 points lower than kids who don't.

Proportion reduction in each variance due to effect of school mean FRlunch:

```
* Calculate PseudoR2 relative to level-1 FRlunch only model;
%PseudoR2(NCov=2, CovFewer=CovFR1, NameFewer=FRLunch1, CovMore=CovFR2, NameMore=FRLunch2);
```

PsuedoR2 (% Reduction) for FRLunch1 vs. FRLunch2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
FRLunch1	UN(1,1)	schoolID	26.1636	4.3105	6.07	<.0001	.
FRLunch1	Residual		239.14	2.9675	80.59	<.0001	.
FRLunch2	UN(1,1)	schoolID	13.3767	2.5214	5.31	<.0001	0.48873
FRLunch2	Residual		239.21	2.9693	80.56	<.0001	-0.00031

Proportion reduction in each variance due to **both** effects of FRlunch:

```
* Calculate PseudoR2 for both FRLunch effects relative to both gender effects;
%PseudoR2(NCov=2, CovFewer=CovGen2, NameFewer=Gender, CovMore=CovFR2, NameMore=FRLunch2);
```

PsuedoR2 (% Reduction) for Gender vs. FRLunch2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
Gender	UN(1,1)	schoolID	43.4390	6.7859	6.40	<.0001	.
Gender	Residual		253.00	3.1392	80.59	<.0001	.
FRLunch2	UN(1,1)	schoolID	13.3767	2.5214	5.31	<.0001	**0.69206**
FRLunch2	Residual		239.21	2.9693	80.56	<.0001	0.05448

Total reduction in math variance due to both effects of FRlunch:

```
* Calculate Total R2 change relative to both gender effects only model;
%TotalR2(DV=math, PredFewer=PredGen2, NameFewer=Gender, PredMore=PredLunch2,
NameMore=FRlunch2);
```

Total R2 (% Reduction) for Gender vs. FRlunch2

Name	Pred Corr	TotalR2	Total R2Diff
Gender	0.03016	0.00091	.
FRlunch2	0.40455	0.16366	0.1627

Model 3c: Adding a Random Effect of Student Free/Reduced Lunch

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyVsGirl}_{ij}) + \beta_{2j}(\text{FRlunch}_{ij}) + e_{ij}$$

$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_j) + \gamma_{02}(\overline{\text{SchoolFRLunch}}_j) + U_{0j}$$

$$\text{Student Gender: } \beta_{1j} = \gamma_{10}$$

$$\text{Free/Reduced Lunch: } \beta_{2j} = \gamma_{20} + U_{2j}$$

```
TITLE1 "Model 3c: 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 = boyvsgirl SMboyvsgirl50 frlunch SMfrlunch30 / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT frlunch / G TYPE=UN SUBJECT=schoolID;
      ODS OUTPUT CovParms=CovFRRand InfoCrit=FitFRRand;
RUN;
```

Estimated G Matrix

Row	Effect	School ID number	Col1	Col2
1	Intercept	125	19.7443	-11.7978
2	frlunch	125	-11.7978	12.7448

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	schoolID	19.7443	3.7094	5.32	<.0001
UN(2,1)	schoolID	-11.7978	3.1478	-3.75	0.0002
UN(2,2)	schoolID	12.7448	3.3242	3.83	<.0001
Residual		236.63	2.9443	80.37	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108866	9	108884	108884	108893	108907	108916

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	49.7910	0.5309	97.4	93.79	<.0001
boyvsgirl	0.8961	0.2703	13E3	3.31	0.0009
SMboyvsgirl50	4.3223	7.2337	104	0.60	0.5515
frlunch	-8.4552	0.5621	98.2	-15.04	<.0001
SMfrlunch30	-16.8499	1.9441	77.7	-8.67	<.0001

Does the random slope for student gender help the model? *Yes, model fit is better.*

```
* Calculate difference in model fit relative to both fixed FRlunch effects model;
%FitTest(FitFewer=FitFR2, NameFewer=FixedFRlunch, FitMore=FitFRRand, NameMore=RandomFRLunch);
```

Likelihood Ratio Test for FixedFRlunch vs. RandomFRLunch

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FixedFRlunch	108954	7	108968	108986	.	.	.
RandomFRLunch	108866	9	108884	108907	88.2542	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.

Calculate and interpret a 95% random effects confidence interval for the random slope:

$-8.4552 \pm 1.96*SQRT(12.7488) = -15.45 \text{ to } -1.46$

On average, the gap in math performance related to kid free/reduced lunch status is 8.4 points, but across 95% of the schools, that gap is predicted to be anywhere from 1.46 to 15.45 points.

Model 3d: Adding a Cross-Level Interaction of Student by School Free/Reduced Lunch

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyVsGirl}_{ij}) + \beta_{2j}(\text{FRlunch}_{ij}) + e_{ij}$$

$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_j) + \gamma_{02}(\overline{\text{SchoolFRLunch}}_j) + U_{0j}$$

$$\text{Student Gender: } \beta_{1j} = \gamma_{10}$$

$$\text{Free/Reduced Lunch: } \beta_{2j} = \gamma_{20} + \gamma_{22}(\overline{\text{SchoolFRLunch}}_j) + U_{2j}$$

```
TITLE1 "Model 3d: 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 = boyvsgirl SMboyvsgirl50 frlunch SMfrlunch30 frlunch*SMfrlunch30
                              / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT frlunch / G TYPE=UN SUBJECT=schoolID;
       ODS OUTPUT CovParms=CovInt1 InfoCrit=FitInt1; RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	schoolID	19.5909	3.6558	5.36	<.0001
UN(2,1)	schoolID	-11.1933	3.0624	-3.66	0.0003
UN(2,2)	schoolID	11.8021	3.1679	3.73	<.0001
Residual		236.62	2.9439	80.37	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108863	10	108883	108883	108894	108909	108919

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	49.7497	0.5300	99.7	93.87	<.0001
boyvsgirl	0.8996	0.2703	13E3	3.33	0.0009
SMboyvsgirl50	4.8849	7.2788	103	0.67	0.5037
frlunch	-8.7042	0.5679	112	-15.33	<.0001
SMfrlunch30	-19.2722	2.4760	89.2	-7.78	<.0001
frlunch*SMfrlunch30	4.3141	2.6386	103	1.64	0.1051

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

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 lower by 8.70.

What does the effect of school proportion free/reduced lunch NOW represent in this model?

This is the 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 lower by 1.9.

What does the cross-level interaction of student by school free/reduced lunch represent?

The effect of being a kid who receives free/reduced lunch is reduced nonsignificantly by 0.4 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.

Proportion reduction in each variance relative to Model 3c with random FRLunch:

```
* Calculate PseudoR2 for both FRlunch effects relative to random FRlunch;
%PseudoR2(NCov=4, CovFewer=CovFRRand, NameFewer=LunchMain, CovMore=CovInt1,
NameMore=LunchInt1);
```

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
LunchMain	UN(1,1)	schoolID	19.7443	3.7094	5.32	<.0001	.
LunchMain	UN(2,2)	schoolID	12.7448	3.3242	3.83	<.0001	.
LunchMain	Residual		236.63	2.9443	80.37	<.0001	.
LunchInt1	UN(1,1)	schoolID	19.5909	3.6558	5.36	<.0001	0.007768
LunchInt1	UN(2,2)	schoolID	11.8021	3.1679	3.73	<.0001	0.073968
LunchInt1	Residual		236.62	2.9439	80.37	<.0001	0.000060

Which variance component should be reduced (non-significantly)? *Random slope for FRlunch*

Model 3e: Adding a Level-2 Interaction of Quadratic School Free/Reduced Lunch

$$\text{Level 1: } \text{Math}_{ij} = \beta_{0j} + \beta_{1j}(\text{BoyVsGirl}_{ij}) + \beta_{2j}(\text{FRlunch}_{ij}) + e_{ij}$$

$$\text{Level 2: Intercept: } \beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{\text{SchoolGender}}_j) + \gamma_{02}(\overline{\text{SchoolFRLunch}}_j) + \gamma_{03}(\overline{\text{SchoolFRLunch}}_j)^2 + U_{0j}$$

$$\text{Student Gender: } \beta_{1j} = \gamma_{10}$$

$$\text{Free/Reduced Lunch: } \beta_{2j} = \gamma_{20} + \gamma_{22}(\overline{\text{SchoolFRLunch}}_j) + U_{2j}$$

```
TITLE1 "Model 3e: 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 = boyvsgirl SMboyvsgirl50 frlunch SMfrlunch30 frlunch*SMfrlunch30
                   SMfrlunch30*SMfrlunch30 / SOLUTION DDFM=Satterthwaite OUTPM=PredInt2;
      RANDOM INTERCEPT frlunch / G 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 Interact" frlunch*SMfrlunch30 1 SMfrlunch30*SMfrlunch30 1;
RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	schoolID	18.8449	3.5480	5.31	<.0001
UN(2,1)	schoolID	-10.8805	3.0207	-3.60	0.0003
UN(2,2)	schoolID	11.8613	3.1882	3.72	<.0001
Residual		236.61	2.9439	80.37	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
108861	11	108883	108883	108894	108911	108922

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	50.3622	0.6594	114	76.37	<.0001
boyvsgirl	0.9016	0.2703	13E3	3.34	0.0009
SMboyvsgirl50	3.0898	7.2988	100	0.42	0.6730
frlunch	-8.8417	0.5775	114	-15.31	<.0001
SMfrlunch30	-17.9370	2.5918	88.1	-6.92	<.0001
frlunch*SMfrlunch30	5.5170	2.7679	108	1.99	0.0488
SMfrlunch30*SMfrlunch30	-13.5458	8.9313	86.4	-1.52	0.1330

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
FR Lunch Between-School Main Effect	-26.7787	2.6001	90.5	-10.30	<.0001
FR Lunch Between-School Interaction	-8.0288	8.5365	74.4	-0.94	0.3500

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 by 0.55 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 becomes nonsignificantly more negative by 2*1.3 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 becomes nonsignificantly more negative by 2*0.80 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.

Proportion reduction in each variance due to the quadratic school free/reduced lunch effect:

```
* Calculate PseudoR2 relative to level-1 lunch interaction only model;
%PseudoR2(NCov=4, CovFewer=CovInt1, NameFewer=LunchInt1, CovMore=CovInt2, NameMore=LunchInt2);
```

PsuedoR2 (% Reduction) for LunchInt1 vs. LunchInt2

Name	CovParm	Subject	Estimate	StdErr	ZValue	ProbZ	PseudoR2
LunchInt1	UN(1,1)	schoolID	19.5909	3.6558	5.36	<.0001	.
LunchInt1	UN(2,2)	schoolID	11.8021	3.1679	3.73	<.0001	.
LunchInt1	Residual		236.62	2.9439	80.37	<.0001	.
LunchInt2	UN(1,1)	schoolID	18.8449	3.5480	5.31	<.0001	0.038080
LunchInt2	UN(2,2)	schoolID	11.8613	3.1882	3.72	<.0001	-0.005017
LunchInt2	Residual		236.61	2.9439	80.37	<.0001	0.000015

Total reduction in math variance due to both intra-variable interactions of FRlunch:

```
* Calculate Total R2 change relative to both gender effects only model;
%TotalR2(DV=math, PredFewer=PredLunch2, NameFewer=LunchMain, PredMore=PredInt2,
NameMore=LunchInt2);
```

Total R2 (% Reduction) for LunchMain vs. LunchInt2

Name	Pred Corr	TotalR2	Total R2Diff
LunchMain	0.40455	0.16366	.
LunchInt2	0.40589	0.16475	.001088340

Plot of model-predicted math by free/reduced lunch status to illustrate interactions: