Cross-level Interaction Example: Predicting Random Slopes

I used the HSB data set to examine a cross-level interaction between sector (the school-level variable indicating whether the school was public or private: 0=public, 1=Catholic) and individual student SES. A significant interaction test would indicate that the relationship between SES and math achievement depends on whether the school is public or private. I left the default Sattherthwaite degrees of freedom in both SPSS and R `lme4`, because there should be sufficient sample size with 160 schools and there should be no difference from the Kenward-Roger approach.

SPSS

```
get file='c:\jason\spsswin\mlrclass\hsbmerged.sav'.

* Following recommendations of Curran & Bauer (2006) this analysis group-centered SES.[1]
*easiest in this case to just use the existing meanses variable.
compute cses=ses - meanses.

*but aggregate command would normally be needed.
* AGGREGATE creates group means for ses variable if you use the group id on the break subcommand.
*AGGREGATE
   /BREAK schoolid
   /mnses=MEAN(ses).
*compute cses=ses - mnses.

** Test the SES and sector cross-level interaction.
MIXED mathach WITH cses sector
  /CRITERIA=MXITER(1000) SCORING(1)
  /METHOD = REML
  /PRINT = SOLUTION TESTCOV HISTORY
  /FIXED =  cses sector cses*sector| SSTYPE(3)
  /RANDOM = INTERCEPT  cses | SUBJECT(schoolid) COVTYPE(UN).
```

Information Criteria[a]

-2 Restricted Log Likelihood	46638.605
Akaike's Information Criterion (AIC)	46646.605
Hurvich and Tsai's Criterion (AICC)	46646.610
Bozdogan's Criterion (CAIC)	46678.122
Schwarz's Bayesian Criterion (BIC)	46674.122

The information criteria are displayed in smaller-is-better forms.

Fixed Effects

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	11.41064	.2929978	157.820	38.944	.000	10.8319397	11.9893453
cses	2.8028147	.1549603	132.895	18.087	.000	2.4963069	3.1093224
sector	2.7995407	.4395387	153.252	6.369	.000	1.9312037	3.6678778
cses * sector	-1.34109	.2337667	144.133	-5.737	.000	-1.8031459	-.8790385

a. Dependent Variable: mathach.

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	157.820	1516.671	.000
cses	1	132.895	327.151	.000
sector	1	153.252	40.568	.000
cses * sector	1	144.133	32.912	.000

a. Dependent Variable: mathach.

Covariance Parameters

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Residual		36.70560	.6257696	58.657	.000	35.4993786	37.9528069
Intercept + cses [subject = schoolid]	UN (1,1)	6.7504155	.8668911	7.787	.000	5.2483024	8.6824473
	UN (2,1)	1.0507234	.3425159	3.068	.002	.3794047	1.7220422
	UN (2,2)	.2656916	.2288151	1.161	.246	.0491268	1.4369356

a. Dependent Variable: mathach.

---

[1] These authors also recommend centering the level-2 variable, which I did not do here, but that would be perfectly acceptable to do. The interaction and the sector main effect will not change but the main effect for SES will change. This is because the SES slope represents the relationship between SES and math achievement when sector equal 0 without centering (and is thus one of the simple effects). With sector centered the slope for SES is for the average of the sample with private and public schools combined. The AGGREGATE command can be used to derive the full sample mean if the /BREAK subcommand is left off.

```
AGGREGATE
  /mnsector=mean(sector).
compute gndsector=sector - mnsector.
```

R

```
> #Model with group-centered SES

> # group-mean centering of original ses variable called cses
> mydata$cses <- mydata$ses - mydata$meanses

> # group average already in data set with meanses, normally would need to compute for group-mean centering
> #mydata$cses <- mydata$ses - ave(mydata$ses,mydata$schoolid)

> # grand-mean centering of sector (not used in the example)
> #mydata$gndsector <- mydata$sector - mean(mydata$sector)
> #Model with group-centered SES (necessary in this case--more later)
> library(lme4)
> model1 <- lmer(mathach ~ cses + sector + cses*sector + (cses|schoolid), data = mydata, REML = TRUE)
> summary(model1)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: mathach ~ cses + sector + cses * sector + (cses | schoolid)
   Data: mydata

REML criterion at convergence: 46638.6

Scaled residuals:
     Min       1Q   Median       3Q      Max
-3.06490 -0.73237  0.01565  0.75370  2.94195

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 schoolid (Intercept)  6.7504  2.5982
          cses         0.2657  0.5154   0.78
 Residual             36.7056  6.0585
Number of obs: 7185, groups:  schoolid, 160

Fixed effects:
            Estimate Std. Error       df t value             Pr(>|t|)
(Intercept)  11.4106     0.2930 158.4267  38.944 < 0.0000000000000002 ***
cses          2.8028     0.1550 141.6607  18.087 < 0.0000000000000002 ***
sector        2.7995     0.4395 153.7010   6.369        0.00000000209 ***
cses:sector  -1.3411     0.2338 151.5366  -5.737        0.00000005077 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) cses   sector
cses         0.262
sector      -0.667 -0.175
cses:sector -0.174 -0.663  0.264

> confint(model1)
Computing profile confidence intervals ...
                 2.5 %     97.5 %
.sig01       2.2811696  2.9319119
.sig02       0.3049392  1.0000000
.sig03       0.1775862  0.8726071
.sigma       5.9586835  6.1611600
(Intercept) 10.8365900 11.9846946
cses         2.4993178  3.1070957
sector       1.9384001  3.6608755
cses:sector -1.7987809 -0.8826888
```

HLM

In HLM, testing an interaction is done simply by using the group-level variable to predict slopes at level 2. Notice in the output that the mixed (single, multilevel) equation shows an interaction term. HLM provides options for group or grand centering variables as they are added to the equation. SES was group-mean centered in this example and HLM prints a note about that. Sector was entered uncentered. HLM also provides convenient graphing, which I will illustrate in more detail later.

```
The maximum number of level-1 units = 7185
The maximum number of level-2 units = 160
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

The outcome variable is MATHACH
```

Summary of the model specified

Level-1 Model

$$MATHACH_{ij} = \beta_{0j} + \beta_{1j}*(SES_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(SECTOR_j) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}*(SECTOR_j) + u_{1j}$$

SES has been centered around the group mean.

Mixed Model

$$MATHACH_{ij} = \gamma_{00} + \gamma_{01}*SECTOR_j + \gamma_{10}*SES_{ij} + \gamma_{11}*SECTOR_j*SES_{ij} + u_{0j} + u_{1j}*SES_{ij} + r_{ij}$$

Final Results - Iteration 43

Iterations stopped due to small change in likelihood function

τ (as correlations)

INTRCPT1, β_0	1.000	0.725
SES, β_1	0.725	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, β_0	0.884
SES, β_1	0.138

The value of the log-likelihood function at iteration 43 = -2.331840E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	11.393836	0.292784	38.915	158	<0.001
SECTOR, γ_{01}	2.807465	0.439216	6.392	158	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.802449	0.156523	17.904	158	<0.001
SECTOR, γ_{11}	-1.340634	0.236028	-5.680	158	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	11.393836	0.292348	38.974	158	<0.001
SECTOR, γ_{01}	2.807465	0.435634	6.445	158	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.802449	0.157937	17.744	158	<0.001
SECTOR, γ_{11}	-1.340634	0.230324	-5.821	158	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	2.59609	6.73966	158	1383.78477	<0.001
SES slope, u_1	0.55141	0.30405	158	175.31196	0.164
level-1, *r*	6.05722	36.68995			

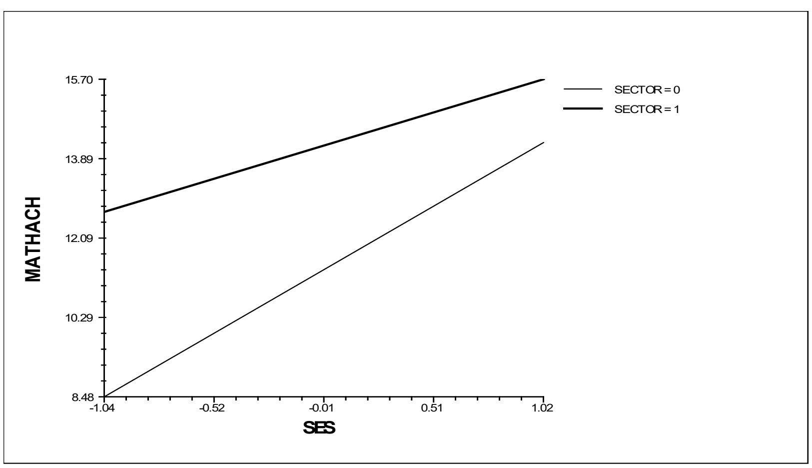
Statistics for current covariance components model

```
Deviance = 46636.802657
Number of estimated parameters = 4
```

Write Up

A multilevel regression model was tested to investigate whether student socioeconomic status (SES) and school type (public vs. private) were predictive of a student's math achievement. Both intercepts and slopes were allowed to vary across schools. The cross-level interaction between school type and student SES was included in order to investigate whether the relationship between SES and math achievement depended on the type of school the student attended. In accordance with several recommendations about independent variable scaling with cross-level interaction tests (e.g., Bauer & Curran, 2005; Enders & Tofighi, 2007), the SES variable was group-mean centered in this analysis in order to improve the interpretation of the main effect. Sector was left uncentered in this analysis. The fixed effects intercept was 11.39, which represents the average math achievement score in public schools, because school type was coded 0 for public schools. The test of the random effect indicated that the average math achievement within the public sector varied significantly across schools when SES was at the school average, τ_0^2 = 6.74, $\chi^2(158)$ = 1383.78, $p < .001$. SES was significantly related to math achievement, with an increase of 2.80 points in the math achievement test for each unit increase in SES, γ_{10} = 2.80, SE = .16, $p < .001$. The slopes for SES did not vary significantly across schools, τ_1^2 = .30, $\chi^2(158)$ = 175.311, ns, however. The type of school was also significantly related to math achievement, γ_{01} = 2.81, SE = .44, $p < .001$. Students attending a private school had a math achievement score approximately 2.81 points higher than those attending a public school. These "main effects" should be interpreted in light of the significant cross-level, interaction, however. The effects of SES on math achievement depended on whether the student attended a private or public school, γ_{11} = -1.34, SE = .23, $p < .001$. Figure 1, which plots the interaction, shows that SES was more strongly related to math achievement for students in the public sector than the private sector. *[I include the following text, because it is along the lines of what I would recommend for a write-up of a cross-level interaction in a journal article. Values are taken from the handout "Simple Slope Tests of Cross-level Interactions" available on the course website].* Simple slope tests (e.g., Preacher, Curran, & Bauer, 2006) indicated that SES was significantly related to math achievement for both school types. Students in public schools had math achievement scores that were 2.8 points higher for each unit increase in SES, $\gamma_{SES|pub}$ = 2.802, $p < .001$, whereas, students in private schools had math achievement scores only approximately one and one-half points higher for every unit increase in SES, $\gamma_{SES|pri}$ = 1.462, $p < .01$.

Figure 1



References

Bauer, D.J., & Curran, P.J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40*, 373-400.

Enders, C.K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121-138.