Scientific Software International

Constraining fixed effects using the HSB data

1. Description of the data

High School and Beyond (**HS&B**) is a national longitudinal study originally funded by the United States Department of Education's National Center for Education Statistics (NCES) as a part of their longitudinal studies program. Its purpose was to document the "educational, vocational, and personal development of young people following them over time as they begin to take on adult roles and responsibilities". Here a subset of the data representing 160 schools and a total of 7,185 students is used to illustrate the fitting and interpretation of a two-level hierarchical linear model.

This example is the sixth in a set of seven examples based on these data described on the HLM Support page.

Level-1 file. For our HS&B example data, the level-1 file (HSB1.SAV) has 7,185 cases and four variables (not including the SCHOOL ID). The variables are:

- MINORITY, an indicator for student ethnicity (1 = minority, 0 = other)
- FEMALE, an indicator for student gender (1 = female, 0 = male)
- SES, a standardized scale constructed from variables measuring parental education, occupation, and income
- MATHACH, a measure of mathematics achievement

Data for the first ten cases in HSB1.SAV are shown in Fig. 2.1.

Note: level-1 cases must be grouped together by their respective level-2 unit ID. To assure this, sort the level-1 file by the level-2 unit ID field prior to entering the data into HLM2.

	id	minority	female	ses	mathach
1	1224	0	1	-1.528	5.876
2	1224	0	1	-.588	19.708
3	1224	0	0	-.528	20.349
4	1224	0	0	-.668	8.781
5	1224	0	0	-.158	17.898
6	1224	0	0	.022	4.583
7	1224	0	1	-.618	-2.832
8	1224	0	0	-.998	.523
9	1224	0	1	-.888	1.527
10	1224	0	0	-.458	21.521

Figure 2.1   **First ten cases in HSB1.SAV**

Level-2 file. At level 2, the illustrative data set HSB2.SAV consists of 160 schools with 6 variables per school. The variables are:

- SIZE (school enrollment)
- SECTOR (1 = Catholic, 0 = public)
- PRACAD (proportion of students in the academic track)
- DISCLIM (a scale measuring disciplinary climate)
- HIMNTY (1 = more than 40% minority enrollment, 0 = less than 40%)
- MEANSES (mean of the SES values for the students in this school who are included in the level-1 file)

The data for the first ten schools are displayed in Fig 2.2.

	id	size	sector	pracad	disclim	himinty	meanses
1	1224	842	0	.350	1.597	0	-.428
2	1288	1855	0	.270	.174	0	.128
3	1296	1719	0	.320	-.137	1	-.420
4	1308	716	1	.960	-.622	0	.534
5	1317	455	1	.950	-1.694	1	.351
6	1358	1430	0	.250	1.535	0	-.014
7	1374	2400	0	.500	2.016	0	-.007
8	1433	899	1	.960	-.321	0	.718
9	1436	185	1	1.000	-1.141	0	.569
10	1461	1672	0	.780	2.096	0	.683

Figure 2.2   **First ten cases in HSB2.SAV**

2. Description of the model

Of interest in this example is the relationship between a student's mathematical achievement (MATHACH) and socio-economic status (SES). Socio-economic status was measured at both levels: at the student level, a measure of the individual's socio-economic status is represented by the variable SES. For each school, a measure of the mean SES for the school is represented by the variable MEANSES.

The 7185 level-1 units are the students, and at this level the social class (denoted by the variable SES) is used to model the math achievement. At level-2, each of the 160 schools' intercept and slope are predicted by school sector and school mean social class.
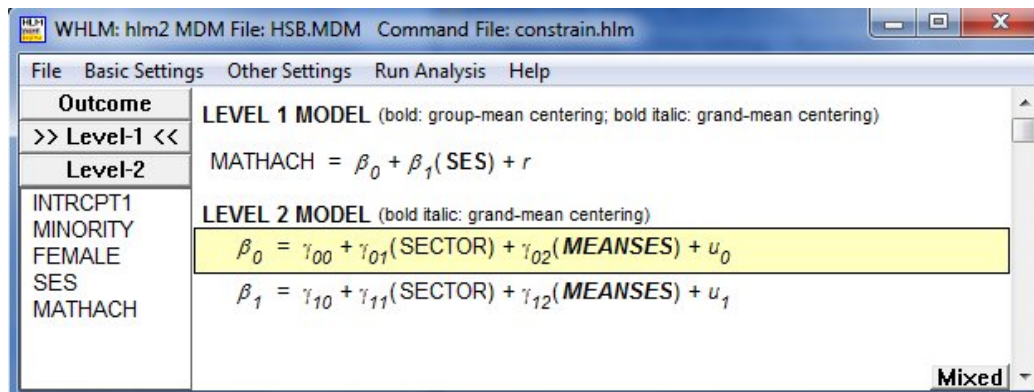

The information above implies that the level-1 model will have two coefficients for each student: the intercept (β_{0j}) and the SES slope (β_{1j}). At level-2 the intercept and SES slope are modeled as functions of sector and mean SES. Note that both the intercept and the slope are modeled as having randomly varying residuals. The assumption is that the intercept and slope vary not only as a function of the two predictors SECTOR and MEANSES, but also as a function of a unique school effect.
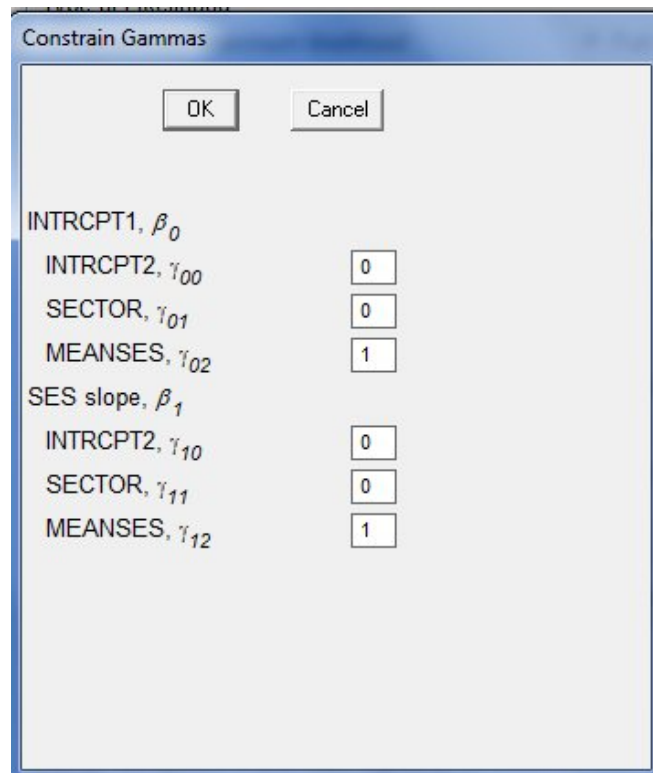
3. Creating the command file

Now suppose we wish to constrain the MEANSES effect to be the same on both the intercept and SES slope equation. In other words, we want to specify

$$\gamma_{02} = \gamma_{12}$$

As a first step, the model specification should be completed. Here we use the model created in the first of this set of examples (**HSB1.HLM**) as starting point. Select the **Estimation Settings** option from the **Other Settings** menu on the main menu bar.
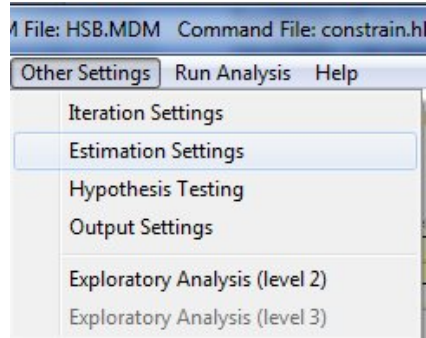
Click the **Constraint of fixed effects** button on the **Estimation Settings** dialog box to open the **Constrain Gammas** dialog box. Enter the value "1" next to MEANSES, γ_{02} and MEANSES, γ_{12}. Note that up to 5 constraints may be imposed in this manner.

After saving the command file, clicking the **Run Analysis** option at the top of the main window initiates the analysis. Once the iterative procedure has converged, the output will automatically be displayed in the format set on the **Preferences** dialog box.

4. Interpreting the output

The final estimation of fixed effects table shows the effect of the imposed constraint. Instead of two MEANSES coefficients, only one is reported as the constraint set these to be equal to each other.

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}	11.741110	0.208949	56.191	157	<0.001
SECTOR, γ_{01}	2.016592	0.331788	6.078	157	<0.001
MEANSES, γ_{02}*	2.654259	0.253568	10.468	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	3.159138	0.158565	19.923	158	<0.001
SECTOR, γ_{11}	-2.096300	0.240992	-8.699	158	<0.001

The "*" gammas have been constrained. See the table on the header page.

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	1.84024	3.38650	157	787.30255	<0.001
SES slope, u_1	0.60746	0.36901	157	192.72983	0.027
level-1, r	6.06228	36.75129			

Statistics for the current model

Deviance = 46567.491225
Number of estimated parameters = 9

When these results are compared with those for the unconstrained model (given below), it is clear from the table of variance component results that imposing the constraint had a significant impact on the statistical significance of the SES slope. In the unconstrained model, there was no evidence of significant random variation in the SES slope. In the constrained model, there is significant random variation in the SES slope. Results suggest that an unconstrained MEANSES effect on the slope equation plays an important role in explaining the random variation in the SES slope.

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}	12.095005	0.173546	69.694	157	<0.001
SECTOR, γ_{01}	1.226775	0.308341	3.979	157	<0.001
MEANSES, γ_{02}	5.331626	0.334568	15.936	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.939314	0.147363	19.946	157	<0.001
SECTOR, γ_{11}	-1.644087	0.237434	-6.924	157	<0.001
MEANSES, γ_{12}	1.042828	0.332814	3.133	157	0.002

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	1.52199	2.31646	157	604.95809	<0.001
SES slope, u_1	0.24925	0.06212	157	162.22131	0.371
level-1, r	6.05992	36.72266			

Statistics for the current model

Deviance = 46494.592685
Number of estimated parameters = 10