EPSY 8268 Assignment 4

To complete this assignment, identify a longitudinal data set with at least three data points within person – although you can have missing data such that not all persons are measured at each time point and the time-points of measurement can be different across persons (not all persons measured the same number of times or at the same time points). However, this should be a growth model – such that you have measures nested within person to estimate growth over time. It doesn’t necessarily need to be academic growth; it could include change over time in any performance domain (learning-based achievement, productivity on the job, skill development, etc.). Remember, there are lots of data sets online – there are a few links to data at the class website. You may use the data set simulated for this assignment, described below.

To help our review of your work, please use the outline structure below in your written assignment – feel free to simply write and/or paste your responses under each element in the following list. Keep your responses brief as a direct response to each element.

 You can elect to use longitudinal data from another source (your own data or data available from another study), or you may elect to use the data available at the class website which is simulated data (based on the *US Sustaining Effects Study*, an example data set used in the HLM Users’ Guide and provided in the example data accompanying the HLM software):

***Growth in Mathematics Study - Modified***

A longitudinal study of children’s growth in mathematics achievement during primary school, from grades 1 to 5. This study focuses on children with multiple at-risk indicators. Level-1 provides an opportunity to model growth curves per student. Level-2 provides an opportunity examine variability in growth curves given student characteristics.

Level-1 Data: EG1s.sav

 1479 observations for 350 children, beginning in grade 1, annually through grade 5. The study occurred over 6 years, enlisting students over time, and including students that repeated a grade because of retention. On average, there are 4 observations per student (lots of missing data). Some students were measured multiple times in the same grade because of being retained (repeating grades), but only once per year.

* Year (year of the study minus 3.5, 0 = the midpoint of the study)
* Grade (the grade of the student minus 1, so that 0 = grade 1, initial status)
* Math (the math score on an IRT scale score metric; *M* = 0, *SD* = 1.5)
* Retained (1 = retained in grade during that year, 0 = not retained; 4% were retained)

Level-2 Data: EG2s.sav

 350 children, with full information (no missing data at level-2).

* Female (1 = female, 0 = male; 56% female)
* Black (1 = African American, 0 = Not AA; 67% AA)
* Latino (1 = Latino, 0 = not Latino; 11% Latino)
* SOC (1 = Student of Color, 0 = White; 78% SOC)
* posid (Positive Identify, a positive outlook and belief in oneself: *M* = 50, *SD* = 10)
* ses (social-economic status, composite of several variables: *M* = 0, SD = 1)

Assignment Tasks:

1. Establish a baseline.
	1. Briefly describe the data set being used, the variables included in the model (as reported above) for the assignment and report the software being used.
	2. Estimate a baseline model (includes the time-point variable). Report a table of the fixed and random effects, interpreting each coefficient. Try to do this in a list format – rather than a paragraph.
	3. Compute a 95% confidence interval around the average intercept (β00) and slope (β10).
	4. Estimate the range of plausible values for the fixed effect (the range of 95% of the intercepts and slopes).
	5. What are the reliabilities of the intercepts and slopes?
	6. What is the correlation between the intercepts and slopes? What is your interpretation of this correlation – what does it mean?
2. Specify a model with no more than one explanatory variable at level 1 (a potential time-varying covariate; use ‘retention’ in the EG data, or no variables at the level-1 model) and one variable at level 2 (in the intercept model and the slope model). If you include a time-varying covariate at level 1, fix the effect at level 2. Clearly state whether and how each variables is centered.
	1. Write out the model in hierarchical notation.
	2. Report a table of the fixed and random effects.
	3. Define and interpret the meaning of each fixed effect and random variance component. [Again, create a list for this.]
	4. Compute 95% confidence intervals for the fixed effects.
	5. Estimate the variance explained at level-1 due to the inclusion of the level-1 explanatory variable(s) [if you included a level-1 time-varying covariate].
	6. Estimate the variance explained at level-2 for the intercept, due to the inclusion of the level-2 explanatory variable in the intercepts model.
	7. Estimate the variance explained at level-2 for the slopes, due to the inclusion of the level-2 explanatory variable in the slopes model.
	8. Provide a general statement about growth and the variables included in the model: What is the overall conclusion regarding growth as measured and modeled in your selected study?

Follow the overall guidance for model building, especially when monitoring variance explained.

1. Specify the level-1 model. This will be the base model, with the time-point variable, but it may also include a time-varying covariate.
2. Specify the intercept model, including the level-2 explanatory variable for the intercepts.
3. Specify the slopes model, including the same level-2 explanatory variable for the slopes.