Exercise Sheet 6 — Mixed-Effects Model

Problem 1. The Study of Instructional Improvement Project
The SII[1] was carried out to assess the math achievement scores of first and third-grade pupils in randomly selected classrooms from a national US sample of elementary schools. The data set includes results for 1190 first-grade pupils sampled from 312 classrooms in 107 schools.
The data set `school.csv` contains the school level variables

schoolid School id number

housepov % of households in the neighbourhood of the school below the poverty level.

The data set `class.csv` includes the classroom level variables

classid Classroom id number

yearstea Years of teacher's experience in teaching in the first grade

mathprep Number of preparatory courses on the first-grade math contents and methods followed by the teacher (a numeric value from 1 to 6)

mathknow Teacher's knowledge of the first-grade math content (the higher the better)

The data set `pupil.csv` is concerned with the pupil level variables

childid Pupil's id number

mathgain Pupil's gain in the math achievement score from the spring of kindergarten to the spring of first grade

mathkind Pupil's math score in the spring of kindergarten

sex Gender

minority Indicator variable for the minority status

The outcome of interest in contained in the variable `mathgain`.

(a) The variables `schoolid` and `classid` are also present in the data sets `class.csv` and `pupil.csv`. Merge the three data sets by these variables.

(b) How many nested levels is there in the data? How many random effects would you consider?

(c) Find a good model for the data. Explain the effect of the predictors on the gain math achievement score.

[1] Hill, H., Rowan, B., and Ball, D. (2005). Effect of teachers mathematical knowledge for teaching on student achievement. American Educational Research Journal, 42, 371-406.