Multilevel Modeling and Ordinary Least Squares Regression: How Comparable Are They? APPLIED EXAMPLE

A large, publicly-available secondary dataset was analyzed using several regression procedures to compare and contrast results (see main article). The objective of the current example is to illustrate how regression results may differ based on the technique used. A specific level-one question was investigated: Did speaking another language at home have an association with reading achievement, while controlling for socioeconomic status?

Dataset

Data came from the Programme for International Student Assessment (PISA) 2012, a large-scale international assessment conducted in 65 countries and economies and organized by the Organisation of Economic Co-operation and Development (OECD). Worldwide, PISA assessed the competencies of approximately 510,000 15-year-olds in reading, mathematics, and science. For the current study, we used the student and school data files for Thailand and data were downloaded from the PISA website (https://www.oecd.org/pisa/).

PISA used a two-stage stratified sampling design where in the first stage, schools were sampled and 15-year-old students within the school were randomly sampled. In total, the Thai data file contained responses from 6,606 students (57% female) from 239 schools. For demonstration purposes, weights and plausible values were not used though we recommend using both the weights and the full set of plausible values when investigating research questions of substantive interest (OECD, 2009). The current analyses focused on 5,607 students attending 210 public schools (average of 26.7 students per school).

Measures

Reading. Reading was assessed using 45 literacy-related items and the assessment was administered in Thai only. For the current analyses, only the first set of plausible value was used (PV1READ; *M* = 459.49, *SD* = 81.90, range = 173.20 – 719.01).

Other language. Respondents indicated the language often spoken at home (LANGN) and was coded as Thai or another language or dialect. Students who spoke a language other than Thai (40%) were coded as a 1 (OTHL = 1) and students who spoke Thai (60%) were coded as 0.

Socioeconomic status. At the student level, the number of books at home (ST28Q01, 1 = 0 – 10 books to 6 = more than 500 books) was used as a measure of socioeconomic status (SES, *M* = 2.58, *SD* = 1.22).

Analytic Strategy

As an initial step, an overall pooled correlation table was examined and then disaggregated to within- and between-group correlation tables. Covariance theorem allows a pooled correlation matrix to be broken down into two separate and orthogonal components (Bliese, 2000; Pedhazur, 1997; Robinson, 1950).

Using the complete dataset, we then performed a standard OLS regression, a multilevel random-intercept model, and a school fixed effects model. For each model, we used reading score as the dependent variable and if the child spoke another language at home and the number of books at home, a proxy for SES, as the predictors of interest. As is customary with MLMs, we ran an unconditional model which only included the outcome variable with no predictors to compute the intraclass correlation or ICC.

Additional models were specified to investigate bias reducing strategies. We then performed three more regressions but included the group mean of both language spoken at home

and number of books at home (see main article for explanations). The fourth regression model used OLS regression with regular standard errors. The fifth model used OLS regression with cluster robust standard errors. Finally, the last model used MLM including the group means at level two. The SCHOOLID variable was used as the grouping variable in the cluster robust and multilevel models.

All data management and analyses was performed using R (R Core Team, 2016) and multilevel modeling was done using the *lme4* (Bates, Mächler, Bolker, & Walker, 2015) package. All syntax and the reduced data files can be downloaded from the first author's website.

Results

An overall correlation table using the level-one and the aggregated level-two variables is presented in Table 1. All correlations are statistically significant. In general, students who spoke another language at home (OTHL = 1) had lower reading scores on average (r = -.13). The correlation of the individual level other language variable and the aggregated other language variable was also high (r = .72).

Based on guidelines proposed by Pedhazur (1997) and Robinson (1950), between- and within-group correlations are shown in Table 2. In the upper diagonal of the Table 2, the pooled-within group correlations with reading scores are generally positive and small (rs < .08, ps < .001). The within-group correlation between OTHL and the reading score is small, positive (r = .07), and statistically significant (p < .001) indicating that individuals who spoke another language at home had slightly higher reading scores. In comparison, the between-group correlations (also referred to as ecological correlations) are generally much larger as seen in the lower diagonal of Table 2. The between-group correlation table indicates that the greater the

percent of students speaking another language at the school, the lower the school average reading scores (r = -.30).

Regression Results without Including the Group Mean

Examination of a null multilevel model indicated that the intraclass correlation coefficient was high, ICC = .49 indicating that a large proportion of the outcome variable can be attributable to factors at the group level. Models using standard OLS, MLM, and fixed effects models are then shown in the first three columns of Table 3. Ignoring differences in standard errors, reviewing the first three columns of Table 3 indicates how different the point estimates can be using the OLS, MLM, and FE models. Standard OLS results indicate that those who spoke another language at home other than Thai had lower reading scores (B = -10.85) and that SES had a positive association with reading scores (B = 21.05). On the other hand, MLM results indicated that students who spoke another language at home had higher reading scores (B = 8.12) and also a smaller, but still statistically significant association of SES and reading scores (B = 5.14). The fixed effects model results, which are unbiased by missing level-two variables, indicated similar results to the multilevel model. Results from FE models are considered pure within-group effects and account for all variability at level two.

Regression Results of Including-the-group-mean Models

A disadvantage of the FE model is that as a result of dummy coding the school variables, no other level-two predictors may be included in the model. However, adding the group averages of each level-one predictor will have the same result as running a FE model but allows for the addition of other level-two predictors (see Table 3). Including the group means in models 4 to 6 show that level-one estimates are exactly the same as those derived using a FE model (in model 3). Models 4 to 6 show that speaking another language at home other than Thai was positively

associated with reading scores (B = 9.740). The association of books in the home with reading achievement was also positive (B = 4.303) but not as large as when estimated using OLS without the group mean or using standard MLM (Models 1 and 2).

As a result, the models indicate how different OLS and MLM results can be using a real-life dataset. We also show how a model can be specified such that OLS and multilevel models can have the same point estimates by merely including the group means in the models. Also of note is that at times, when level-one variables are examined, OLS standard errors may not necessarily be smaller when comparing OLS (Model 4) and MLM (Model 6) such as in the case of both level-one predictors.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. http://doi.org/10.18637/jss.v067.i01

Bliese, P. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. Bollen & J. Long (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

OECD. (2009). *PISA data analysis manual: SAS, second edition*. OECD Publishing. Retrieved from http://www.oecd-ilibrary.org/education/low-performing-students_9789264250246-en

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. USA: Wadsworth.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*, 351–357. http://doi.org/10.2307/2087176

Table 1.

Correlation Table (*n* = 5,607).

	Reading	Othl	Books	M.Othl
Othl	-.13			
Books	.33	-.20		
M.Othl	-.21	.72	-.21	
M.Books	.55	-.30	.51	-.42

Notes. Othl = spoke another language other than Thai at home. All correlations statistically significant (*ps* < .001). Variables with M. prefix are the group means of level-one variables aggregated at level two.

Table 2.

Between- and Within-group Correlation Tables (*n* = 5,607).

	Reading	Othl	Books
Reading	1.00	.05	.07
Othl	-.30	1.00	-.07
Books	.77	-.42	1.00

Notes. Lower diagonal data represent between-group correlations. Upper diagonal data represent within-group correlations. Othl = spoke another language other than Thai at home. All correlations statistically significant (*ps* < .001).

Table 3.

Comparison of Results Using Ordinary Least Squares (OLS), Fixed Effects (FE), and Multilevel Modeling (MLM).

	Without IGM			*IGM*		
	OLS (1)	MLM (2)	FE (3)	OLS (4)	OLS[1] (5)	MLM (6)
Othl	-10.853***	8.122***	9.740***	9.740***	9.740***	9.740***
	(2.146)	(2.225)	(2.269)	(2.664)	(2.770)	(2.270)
Books	21.048***	5.140***	4.303***	4.303***	4.303***	4.303***
	(0.865)	(0.739)	(0.743)	(0.873)	(0.809)	(0.743)
M.othl				-4.744	-4.744	-1.445
				(3.906)	(8.919)	(8.775)
M.books				69.516***	69.516***	66.243***
				(1.836)	(4.837)	(4.908)
Constant	409.863***	435.515***	[a]	267.440***	267.440***	272.626***
	(2.748)	(4.499)		(4.849)	(13.917)	(14.201)

Notes. n = 5,607. IGM = including the group mean. othl = spoke another language other than Thai at home (1 = yes, 0 = no). Standard errors within parentheses. [1]Cluster robust standard errors within parentheses. Variables with the M. prefix are the group means of level-one variables aggregated at level two. [a]Constant not meaningful for FE models. ***$p < .001$.