

Drawing Valid Inferences from the Nested Structure of Engineering Education Data: Application of a Hierarchical Linear Model to the SUCCEED Longitudinal Database

Miguel A. Padilla, Guili Zhang, and Timothy J. Anderson
Educational Psychology and Chemical Engineering, University of Florida

Matthew W. Ohland, General Engineering, Clemson University

Abstract

Although hierarchical linear models are seldom used in engineering educational research, the nested structure of students in various colleges of engineering and the longitudinal nature of student records supports the use of such models. Hierarchical linear models account for the nested structure and can test hypotheses on both the schools and the students within the schools simultaneously, thereby eliminating aggregation bias and misestimated standard errors that result when the nested structure is ignored. In the present study, a hierarchical linear model is fitted to the SUCCEED longitudinal database using only students that graduated. As an example, cumulative GPA is regressed on Carnegie school classification, school setting, degree received, gender gap, and citizenship gap with SAT total score and number of terms attended as covariates. The results indicate that there is significant cumulative GPA variance between schools, accounting for 19% of the variance. Additionally, the gender gap and citizenship gap accounted for 6% of the within school cumulative GPA variance, but school setting accounted for 61% of the between school citizenship gap variance. In particular, students that receive their degree in engineering had the highest cumulative GPA. Non-citizens tended to have higher cumulative GPAs than citizens. Another finding is total SAT score is more predictive of cumulative GPA in urban schools than suburban schools. Finally, urban and/or research schools had the strongest relationship between number of terms until graduation with cumulative GPA in that longer times to graduation are associated with lower cumulative GPA.

Introduction

The Southeastern University and College Coalition for Engineering Education (SUCCEED) compiled a student database to help evaluate the impact of its various experiments in undergraduate engineering education. This comprehensive longitudinal database contains the academic records of all students enrolled in the nine SUCCEED universities during the period 1987 to 2002. The extent of the database in terms of the number of students, length of time, and number of universities enables the exploration of a variety of educational questions with statistical significance. Perhaps the most important use of such an extensive database is to understand the relationship between a specific outcome (e.g. cumulative GPA) on various factors (e.g., preparation – SAT scores, gender, discipline).

For purposes of quantitative analysis and generalizability, it is common to represent this relationship with a mathematical model, with linear models being most common. It is important to realize, however, that the data in the SUCCEED database does not result from an experimental design. That is to say, students were not randomly selected from the population and then

randomly assigned to one of the nine schools. Thus internal validity¹ and the independence assumption required by standard linear models such as multiple regression and analysis of variance are not assured. It should also be realized that the estimated model parameters may vary depending on higher level factors. Taking the problem of predicting cumulative GPA as an illustration, one might expect to see differences in model parameters depending on the engineering discipline, or on whether they are an engineering or non-engineering student, or whether they attended an urban or rural institution.

With this in mind, the appropriateness of a hierarchical linear model (HLM) to predict cumulative GPA is fitted to the SUCCEED longitudinal database. Cumulative GPA was selected because it can be a good indicator of graduation. Previous work has looked at various factors influencing graduation and retention of engineering students.^{2,3} The results indicated that SAT math scores, SAT verbal scores, high school GPA, gender, ethnicity, and citizenship all play a role in both graduation and retention rates. Additional work has compared engineering students to other non-engineering majors with respect to similar factors previously mentioned.⁴ The results of this work demonstrated ways that engineering student characteristics differ from those of non-engineering majors. Other work has shown that there are freshman attitudinal differences with respect to school setting (e.g., rural and urban) and whether the school is research oriented or not.⁵ Attitudinal differences between schools are important because they can help explain performance across schools.

In general the assumption of a linear model is that the errors are normally and independently distributed with a mean of zero and a constant variance, $\varepsilon \sim NID(0, \sigma^2)$. Research on the robustness of linear models indicates the limits of each of the following assumptions:

- **Normality.** The linear model is generally robust to the normality assumption (in that the Type I error is little affected) except when the sample size is quite small and the departure from normality is extreme.⁶⁻⁸
- **Constant Variance.** With respect to the constant variance assumption, when the group with the larger sample size also has the largest variance (positive pairing) then the true Type I error will be below the theoretical Type I error (conservative Type I error), and when the group with the smaller sample size has the largest variance (negative pairing) then the true Type I error is inflated (liberal Type I error). If the group sample sizes are equal and sufficiently large, however, the linear model is robust to the constant variance assumption.^{9,10}
- **Independence.** The linear model is not robust to violations of the independence assumption. When the errors are positively correlated, this will produce a liberal Type I error because the standard error will be decreased. This generally occurs when subjects are nested within treatments or schools because the errors for subjects under the same treatment or in the same school are related and not independent. Hierarchical linear models can take into account the nesting of subjects in treatments or schools and thus produce correct standard errors.

In general, HLM can alleviate aggregation bias, misestimated standard errors, and heterogeneity of regression.¹¹ Aggregation bias occurs when an explanatory variable can take on different meanings. For example, social class at the student level (level-1) measures an individual

experience in a particular social class, but at the school level (level-2) it can mean a school's general perception of a particular class. Misestimated standard errors occur when the dependence among individual responses within the same school are not taken into account. This violates the independence assumption just described. Finally, heterogeneity of regression is when, at the student level (level-1), the relationship between the explanatory variables with the outcome variable varies across schools or organizations (level-2).

Methods

The data for this study is the SUCCEED longitudinal database (LDB). The SUCCEED LDB consists of student academic records for nine universities: Clemson University, Florida A&M University, Florida State University, University of Florida, Georgia Institute of Technology, North Carolina A&T State University, North Carolina State University, University of North Carolina at Charlotte, and Virginia Polytechnic Institute and State University. As described in greater detail in previous work, this sample represents a sufficiently large cross-section of engineering graduates (1/12 of all U.S. engineering degrees and higher percentages of women and African-Americans) so the findings should be generalizable to medium and large public universities, comprising a very large fraction of each year's class of engineering graduates.

Although the LDB has data from 1987 to 2002, only data for students matriculating from 1987 to 1996 was extracted to allow for the study of six-year graduation rates. The total sample size extracted from the LDB is $N = 128,814$. There were a total of 124,773 citizens and 4,041 non-citizens. Additionally, the data extracted had 59,179 males and 69,635 females.

The students were categorized by the degree they received at graduation. The three broad categories of degrees received are engineering, other science, or non-science. The sample sizes for gender by citizenship by degree are presented in Table 1. It should be noted that these are the sample sizes before the data are analyzed through a hierarchical linear model.

Table 1. Sample Sizes by Gender by Citizen by Degree

Gender	Citizen	Degree		
		Engineering	Other Science	Non-Science
Male	Yes	21,164	7,912	38,052
	No	1,145	376	986
Female	Yes	5,957	6,060	45,628
	No	229	202	1,103

There was interest in building a valid model for predicting cumulative GPA (CGPA) for undergraduate students that graduated within six years. In particular, cumulative GPA was regressed on the on the following variables: Carnegie school classification,¹² school setting, degree received, total SAT score, number of terms attended, gender gap, and citizenship gap. Level-2 or school level variables are Carnegie school classification and school setting. All other variables are level-1 or student level. The research questions are:

- 1) How much do the schools vary in their cumulative GPA?
- 2) What is the average intercept and slope of the schools; that is, what is the average cumulative GPA across the schools and what is the average relationship between cumulative GPA and level-1 variables across the schools?
- 3) How much variation exists between schools in the regression equations; that is, how much variation is there in the intercepts and slopes or in this case cumulative GPA and the relationship between cumulative GPA and level-1 variables between the schools?

Results

The first model estimated is a one-way random ANOVA to see how much variation exists in cumulative GPA between and within the schools. The results of the random ANOVA are presented in Table 2. The results indicate that the average cumulative GPA across the schools is $\gamma_{00} = 2.72$. Additionally, there is significant cumulative GPA variation between ($\tau_{00} = .07$) and within ($\sigma^2 = .30$) the schools. Additionally, the interclass correlation expressed as:

$$\rho = \frac{\tau_{00}}{(\tau_{00} + \sigma)} = \frac{.069}{.069 + .30} = .19$$

indicates that 19% of the variance in cumulative GPA is between schools.

Table 2. One-Way Random ANOVA for Cumulative GPA

Fixed Effect	Estimate	SE	df	t	p-value
Average School CGPA, γ_{00}	2.7230	.08752	8	31.11	.0001
Random Effect	Variance Component	SE		z	p-value
School CGPA, τ_{00}	.06889	.034310		2.01	.0223
Level-1 Effect, σ^2	.30020	.001183		253.84	.0001

The second model estimated is a random-coefficient model. The results of this model are presented in Table 3. The average cumulative GPA remains unchanged. The average gender gap indicates that, in a typical school, males are scoring .081 points behind females in cumulative GPA. Also, the average citizenship gap indicates that, in a typical school, citizens are scoring .084 points behind non-citizen in cumulative GPA. From the variance component portion of the table, one can see that there is marginal significant variation between the schools with respect to cumulative GPA, significant variation between schools with respect to the gender gap, non-significant variation between schools with respect to the citizenship gap, and significant variation within the schools with respect to cumulative GPA. Additionally, as expressed by:

$$\frac{.30 - .28}{.30} = .06,$$

the gender gap and citizenship gap account for 6% of the within school or level-1 cumulative GPA variance.

Table 3. Random-Coefficient Regression Model

Fixed Effect	Estimate	SE	df	t	p-value
Average School CGPA Mean, γ_{00}	2.7230	.09024	8	30.17	.0001
Average School Gender Gap, γ_{10}	-.08108	.01117	130000	7.26	.0001
Average School Citizenship Gap, γ_{20}	-.08407	.03341	130000	2.52	.0119
Random Effect	Variance Component	SE		z	p-value
School CGPA Mean, τ_{00}	.073240	.050790		1.44	.0747
Gender Gap, τ_{10}	.000953	.000361		2.64	.0041
Citizenship Gap, τ_{20}	.007887	.007094		1.11	.1331
Level-1 Effect, σ^2	.281100	.001108		253.74	.0001

The last model fitted is a random-coefficient model with predictors at level-1 and level-2 added. This type of model is sometimes referred to as an intercept- and slopes-as-outcomes model and is presented in Table 4. In terms of cumulative GPA, students in rural schools have higher cumulative GPA than student in urban schools and students in suburban schools have lower cumulative GPA than students in urban schools. In general, students in rural schools have the highest cumulative GPA and students in suburban schools have the lowest cumulative GPA. Also, students in research schools, as by categorized by the Carnegie classification system, tend to have higher cumulative GPA than students in schools categorized as M.S. institutions. With regard to type of degree received, students receiving their degree in engineering tend to have the highest cumulative GPA and students receiving their degree in the non-sciences tend to have the lowest cumulative GPA. Lastly, African-Americans have a higher cumulative GPA than Whites and Others, Hispanics have higher cumulative GPA than African-Americans, Whites, and Others. Overall, Hispanics and African-Americans have the highest cumulative GPA with Others having the lowest cumulative GPA.

The average school relationship indicates that total SAT score is positively related to cumulative GPA. The relationship between total SAT scores and cumulative GPA is weaker for both rural and suburban schools than urban schools and the relationship is stronger at research institutions than M.S. institutions.

The average school relationship indicates that the number of terms attended is negatively related to cumulative GPA—students attending more terms are likely to have a lower cumulative GPA. The relationship between number of terms to graduation and cumulative GPA is stronger for urban schools than either rural or suburban schools, and stronger at M.S. institutions than research institutions. On average for all students, females tend to have higher cumulative GPA than males.

With regard to the citizenship gap, on average non-citizens tend to have higher cumulative GPA than citizens. Non-citizens at rural and suburban schools tend to have higher cumulative GPA than those in urban schools.

The random effects portion of the analysis indicates that there is significant variation in the gender gap and non-significant variation in the citizenship gap, indicating that there is still unexplained cumulative GPA variability in the gender gap across the schools. Additionally, there is marginal within school or student level cumulative GPA variation, signifying that there may still be unexplained cumulative GPA variance at the student level. Although the between school variation for the citizenship gap was non-significant, adding school setting accounted for 61% of the citizenship gap variation.

Conclusions

Some of the results of this study remained consistent with previous work.^{2,9,12,13} In particular, students that graduate in engineering majors tended to have higher cumulative GPA than students that receive their degree in non-engineering majors. Additionally, students in research and/or rural schools tended to have the highest cumulative GPA. One interesting finding was that Hispanic students had the highest cumulative GPA, which is similar to the finding by Besterfield, Atman, and Shuman¹³ with respect to attitudes toward engineering. In their finding, Besterfield and her colleagues found that Hispanic and African-American freshmen enter engineering programs with considerably higher “impressions about engineering.” The higher impressions about engineering could be translated into more enthusiasm for engineering and help explain why Hispanic students have higher cumulative GPA.

With respect to the relationship between total SAT score and cumulative GPA, urban schools had the strongest total SAT score relationship with cumulative GPA than did rural and suburban schools. Additionally, research schools had a stronger total SAT score relationship with cumulative GPA. This indicates that total SAT score is more predictive of cumulative GPA in urban or research schools. Because research institutions are under heavy pressure to produce publishable research,¹⁴ it is likely that such institutions select highly qualified students that require very little attention and hence the strong relationship between total SAT score and cumulative GPA. Furthermore, Felder and his colleagues¹⁵ found that engineering students that came from an urban background did much better than rural students on almost every measure of academic performance they investigated, suggesting that the academic success that may be attributable to an urban setting in high school may also be attributable to the urban setting in college.

As in previous work, number of terms until graduation is negatively related to cumulative GPA, indicating that more required time till graduation is related to a lower cumulative GPA. This could be a result of students repeating courses or lack of interest in school activities manifesting in enrolling in an insufficient number of hours per term and/or taking free electives.¹⁶ Either or both of these cases can result in students taking longer to graduate and can subsequently decrease cumulative GPA. Furthermore, urban schools had the strongest number of terms until graduation and cumulative GPA relationship. Also, non-research schools have a stronger number of terms

until graduation and cumulative GPA relationship. Hence, students in urban or non-research school that take longer to graduate are more likely have lower cumulative GPA on average.

Although males tend to be behind females in terms of cumulative GPA and there is variation in this gender gap, none of the level-2 or school variables in this study was able to account for some of this variation. In addition, non-citizen students in rural schools had the highest cumulative GPA.

With regard to total cumulative GPA variance, 19% of the variance is between schools. This is a sizable portion of cumulative GPA variance, given that there are only nine schools, suggesting that future work should look for additional level-2 or school level variables. Also, the gender gap and citizenship gap account for 6% percent of the cumulative GPA variance within schools or at the level-1. As before, this suggests that future work should attempt to look for more randomly varying level-1 variables. Finally, school setting accounted for 61% of the between school variation for citizenship gap. As before, given that there are only nine schools, this accounts for a sizeable portion in citizenship gap variation.

The study described here can be seen as the start of an in-depth analysis to determine the critical student and school characteristics determining engineering student graduation as it is related to cumulative GPA. The analysis provided a new approach on how to analyze nested engineering data via a hierarchical model. It is clear that hierarchical models can provide additional information in that one can model both student and school level variables simultaneously and partition the model variance into the appropriate within school and between school components.

Author biographies

MIGUEL A. PADILLA

is a Ph.D. candidate in Educational Research and Statistics, Department of Educational Psychology, University of Florida and previously received an M.A. from that program. He received a B.A. in Psychology from California State University, Domingus Hills.

GUILI ZHANG

is a Ph.D. candidate in Educational Research and Statistics, Department of Educational Psychology, University of Florida. She received a B.A. in British and American Language and Literature at Shandong University, China, and a M.Ed. in English Education at Georgia Southern University. She has published extensively and has won numerous awards at the national and regional level in the area of educational research in China. She teaches Measurement and Assessment in Education at the University of Florida. Her research interests involve applied quantitative research designs, categorical data analysis, and structural equation modeling.

TIMOTHY J. ANDERSON

is Associate Dean for Research and Graduate Programs in the College of Engineering, and Professor in the Department of Chemical Engineering, University of Florida. He received a Ph.D. at the University of California-Berkeley in 1979. His research interests include electronic materials processing, thermochemistry and phase diagrams, chemical vapor deposition, bulk crystal growth and advanced composite materials.

MATTHEW W. OHLAND

is Assistant Professor in General Engineering at Clemson University and is the President of Tau Beta Pi, the national engineering honor society. He received a Ph.D. in Civil Engineering with a minor in Education from the University of Florida in 1996. He previously served as Assistant Director of the NSF-sponsored SUCCEED Engineering Education Coalition. His research is in freshman programs and educational assessment.

Table 4. Intercepts- and Slopes-as-Outcomes Model

Fixed Effect	Estimate	SE	df	t	p-value
School CGPA Mean					
Base, γ_{00}	2.5459	.009137	5	278.63	.0001
Rural, γ_{01}	.15290	.004774	5	32.02	.0001
Suburban, γ_{02}	-.07266	.004716	5	15.41	.0001
Research, γ_{03}	.21120	.005667	5	37.27	.0001
Engineer, γ_{04}	.02863	.005442	16	5.26	.0001
None-Science, γ_{05}	-.08545	.004772	16	17.91	.0001
White, γ_{06}	-.00016	.007103	24	.0200	.9819
Black, γ_{07}	.07945	.008688	24	9.15	.0001
Hispanic, γ_{08}	.11050	.010660	24	10.37	.0001
SAT Total Score					
Base, γ_{10}	.001321	.000032	130000	41.71	.0001
Rural, γ_{11}	-.00099	.000032	130000	30.98	.0001
Suburban, γ_{12}	-.00104	.000031	130000	33.57	.0001
Research, γ_{13}	.00021	.000021	130000	9.96	.0001
Number of Terms					
Base, γ_{20}	-.01060	.000336	130000	31.55	.0001
Rural, γ_{21}	-.00306	.000399	130000	7.65	.0001
Suburban, γ_{22}	-.00145	.000359	130000	4.03	.0001
Research, γ_{23}	-.00203	.000334	130000	6.08	.0001
Gender Gap					
Base, γ_{30}	-.06858	.01107	130000	6.20	.0001
Citizenship Gap					
Base, γ_{40}	-.1793	.04134	130000	4.34	.0001
Rural, γ_{41}	.1703	.05428	130000	3.14	.0017
Suburban, γ_{42}	.1380	.05402	130000	2.55	.0106
Random Effect	Variance Component	SE	Z	p-value	
Gender Gap, τ_{30}	.000928	.000464	2.00	.0227	
Citizenship Gap, τ_{40}	.003066	.003841	.80	.2124	
Level-1 Effect, σ^2	.27430	.001081	253.75	.0001	

References

- [1] Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Casual Inference*. Boston, MA: Houghton-Mifflin.
- [2] Anderson, T., Carter, R., Ohland, M., Thorndyke, B., & Zhang, G. (2002). "Identifying Factors Influencing Engineering Student Retention Through a Longitudinal and Cross-institutional Study Using Quantitative and Qualitative Methods," *Proceedings of the American Society of Engineering Education*, Montreal, Canada, June 2002.
- [3] Zhang, G., Anderson, T., Ohland, M., Carter, R., & Thorndyke, B. (2002). "Identifying Factors Influencing Engineering Student Graduation and Retention: A Longitudinal and Cross-institutional Study," *Proceedings of the American Society of Engineering Education in the Southeast*, Gainesville, FL, April 2002, Session 2793.
- [4] Zhang, G., Thorndyke, B., Carter, R., Ohland, M., and Anderson, T. (2003) "A Comparison of Demographic Factors and Academic Performances Between Students Graduated in Engineering and Other Disciplines." *Proceedings of the American Society of Engineering Education*, Nashville, Tennessee, June 2003.
- [5] Besterfield, M., Mareno, M., Shuman, L. J., & Atman, C. J. (1999). "Comparing Entering Freshman Engineers: Institutional Differences in Student Attitudes," *American Society of Engineering Education Conference Proceedings*, Charlotte, NC, June 1999.
- [6] Donaldson, T. S. (1968). "Robustness of the F Test to Errors of Both Kinds and the Correlation Between the Numerator and Denominator of the F Ratio," *Journal of the American Statistical Association*, 63, 660-676.
- [7] Lindquist, E. F. (1996). *Design and Analysis of Experiments in Education and Psychology*. Boston, MA: Houghton-Mifflin.
- [8] Scheffe, H. (1959). *The analysis of Variance*. New York: Wiley.
- [9] Clinch, J. J., & Keselman, H. J. (1982). "Parametric Alternatives to the Analysis of Variance," *Journal of Educational Statistics*, 7, 207-214.
- [10] Tomarken, A. J., & Serlin, R. C. (1986). "Comparison of ANOVA Alternatives Under Variance Heterogeneity and Specific Noncentrality Structures." *Psychological Bulletin*, 99, 90-99.
- [11] Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and data Analysis Methods*. Thousand Oaks, CA: Sage Publications.
- [12] The Carnegie Foundation for the Advancement of Teaching (1994). *A Classification on Institutions of Higher Education*. <http://www.carnegiefoundation.org/cihe/>.
- [13] Besterfield, M., Atman, C. J., & Shuman, L. J. (1998). "Engineering Student Attitudes Assessment," *Journal of Engineering Education*, 133-141.
- [14] Pister, K. S. (1993). "A Context for Change in Engineering," *Journal of Engineering Education*, 82, 66-69.
- [15] Felder, R. M., Mohr, P. H., Dietz, E. J., & Baker, L. (1994). "A Longitudinal Study of Engineering Student Performance and Retention II. Rural/Urban Student Differences," *Journal of Engineering Research*, 209-217.
- [16] Giesey, J. J., & Manhire, B. (2003). "An Analysis of B.S.E.E Degree Completion Time at Ohio University." *Journal of Engineering Education*, 275-280.