International Journal of Engineering Education Vol. 37, No. 4, pp. 1013–1023, 2021
Printed in Great Britain
0949-149X/91 $3.00+0.00

Time to Graduate for Latinos/Hispanics in Comparison to Other Diverse Student Groups: A Multi-Institutional/ Multilevel MIDFIELD Study*

NOEMI V. MENDOZA DIAZ
College of Education and Human Development and College of Engineering, Texas A&M University, College Station, TX, USA.
E-mail: nmendoza@tamu.edu

CIJY ELIZABETH SUNNY
Department of Information Systems and Business Analytics, Hankamer School of Business Baylor University, TX, USA.
E-mail: cijy_sunny@baylor.edu

TRINIDAD SOTOMAYOR
College of Engineering, Pontificia Universidad Católica de Chile, Santiago, CHILE. E-mail: tssotomayor@uc.cl

JACQUES RICHARD
College of Engineering, Texas A&M Engineering, College Station, TX, USA. E-mail: richard@tamu.edu

Higher institutions of education represent the social mobility mechanisms that create more just societies. The STEM fields are particularly critical in the development of these modern, more just societies. In the United States, the social and racial justice debates are ever more relevant and present in academia. Studies focused on under-represented and under-served groups in education, especially STEM fields, are timely and of paramount importance. This is a study that analyzed student data of 19 institutions, concentrated in what is known as the Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD). It utilized multilevel (HLM) analysis focused on the Time to Graduate outcome of under-represented populations, emphasizing the Latino/Hispanics group. Multilevel analysis is a powerful tool to evaluate differences in groups such as institutions and races/ethnicities, which is the type of data MIDFIELD affords. Results show that depending on the multilevel model, either fixed or random slope, there is a significant difference between the number of terms taken to graduate for under-represented groups, including Latinos/ Hispanics, compared to White groups and for Black compared to White groups. This suggests that Black students tend to be more impacted by their institution than other racial/ethnical groups. Since the emphasis was Latinos/Hispanics, the question remaining is if these results transfer to a sample with more Latino/Hispanic representation.

Keywords: Multiple-Institution database; STEM, MIDFIELD; Latinos; Hispanics

1. Introduction

Latest racial events in the United States have captured the attention of the world and have ignited a scholarly debate comparable to the civil rights era. African Americans, Native Americans, and Latinos/Hispanics, the minority groups underrepresented in STEM fields, have engaged in conversations and research conducive to highlight systemic disparities and unveil all forms of discrimination.

In 2019, Latinos/Hispanics reached nearly 61 million U.S. inhabitants and constitute 18% of the country's population, the second largest racial or ethnic group only behind White non-Hispanics. During the last decade, this group had the fastest growth in the south with a median age of 30 years, positioning itself as the youngest group [1]. In 2016, 10% of undergraduate engineering and computer science degrees were awarded to Latinos/Hispanics. Ten years prior, in 2006, the ratio was 7% and 20 years prior in 1996, the ratio was 5–6% [2].

Hispanic serving institutions keep producing the majority of Hispanic engineers and scientists in the nation [3]. Attrition, however, is still very high in this group that is more likely to be constituted of first time in college (FTIC) individuals [4]. For 4-year postsecondary institutions, in the cohort entry year of 2010, 54% of Hispanic students graduated within 6 years, behind the 74% of Asian and 64% of White students [5]. In terms of STEM programs, for the 2003–2004 entry cohort, 23% of Hispanics left college without a degree and 26% switched to a non-STEM field [6]. The growth in degrees awarded is evident but the extent and time to which education catches up with the population growth is still debatable.

Current Latinos/Hispanics in engineering research involve recruitment/attraction, retention, graduation rates and a multitude of factors affecting them [7–10]. The International Journal of Engineering Education has recently published work addressing issues of Latinos/Hispanics in the manner of undergraduate attrition and contribut-

* Accepted 15 February 2021.

ing factors [11] and the social capital deficit of first generation college students [12]. These IJEE works have been published in an aggregated form, sharing findings with other demographic groups, without a concentration on Latinos/Hispanics

In 2004, a partnership of higher education institutions collected information of their students and shared it in a de-identified databased called Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD). The richness and amount of information provided by this database, including the number of terms each student take to graduate, affords opportunities to investigate issues of Latinos/Hispanics not considered before. The purpose of this study is to investigate the time Latinos/Hispanics take to graduate in engineering and computer science and to compare it against other demographic groups in the context of the institutions participating in MIDFIELD.

1.1 MIDFIELD Database

The study utilizes the Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD) which encompasses whole population data of degree-seeking students at 19 member institutions, including students of all disciplines, transfer students, part-time students, and students who first enroll at any time of year [13]. The MIDFIELD member institutions include seven of the 50 largest U.S. engineering programs in terms of engineering bachelor's degrees awarded, which means that the MIDFIELD population includes 10% of all engineering graduates of U.S. engineering programs [13]. The MIDFIELD dataset contains records for 1,606,962 students from 1987 through 2019. This high number of students enables the MIDFIELD database to represent a substantial portion of U.S. students who attend a wide variety of the typical large continental U.S. institutions. The MIDFIELD database covers a broad demographic of gender, race, and ethnicities for a wide range of studies [14–19]. MIDFIELD data fields are divided in 4 tables, the student table, the term table, the course table, and the grade table. These four tables contain approximately 50 fields that include detailed information of each student. From the term the student enrolled and graduated in the institution, to gender, race/ethnicity, courses enrolled, courses dropped, and grades, including SAT scores. The records are only identified by a number which connects the four tables.

Utilization of the MIDFIELD dataset will enable the study to incorporate many institutions, programs, and cohorts where students may have sufficient numbers representative of persistence, inter-institutional transfers, distinct from migration into and out of majors (switching/changing majors, intra-institutional transfers), etc.

1.2 Literature Review

Prior work with MIDFIELD data has helped to address several questions regarding engineering education. Studies of demographic variables such as gender or race/ethnicity in the context of engineering have been a major concern. One of the main scopes of MIDFIELD researchers has been to study student persistence, where the metric "stickiness" has been widely used [20]. Stickiness is the probability of remaining in a major once it has been declared. This percentage considers the number of students who graduate in a major (in four or six years) divided by the number of students who ever declared that major (transfers are considered). Stickiness has helped to analyze with detail which major makes the students "stick" more, to determine whether the program is doing well. With the creation of the variable, a 2014 study compared the persistence of students filtering by gender or transfer condition (FTIC or transfer) in majors such as Industrial Engineering (IE), Civil Engineering (CvE), Mechanical Engineering (ME), Chemical Engineering (ChE), and Electrical Engineering (EE) among institutions. A later work added race in the comparison [21]. In 2015, another study used stickiness to compare persistence in ChE by gender and race. In this mixed-method work, the authors aimed to discover which institution was making a better effort in helping their students persist [22]. In 2017, another work studied gender persistence differences in majors and compared EE with other majors such as CvE, ChE, IE, and ME [23].

Another important metric using MIDFIELD data is "trajectories" which calculates the number of students who stay within the major they initially declared (at four and six years of study). Trajectories analysis describes which major loses fewer students over the years. In Lord, Layton and Ohland [23], this information was disaggregated by race with added demographics and other outcomes for EE and Computer Engineering (CE). Some work that included stickiness also added trajectories in the analysis to study retention [22, 23].

Prior to the creation of stickiness and trajectory metrics, four and six year graduation rates were the measurements utilized. Orr, Ngambeki, Long & Ohland used the six year graduation rate in relation to gender, race, and major [25]. Previous studies focused on measures of success in engineering education and used both, four and six graduation rates [16, 26, 27].

We consider these metrics valuable because they measure student persistence, engagement, and migration in majors. With the purpose of providing more information about student outcomes, we propose to measure how long the students take to graduate, without the limitation to achieve it in 4 or 6 years. We prefer not to consider persistence as binary (graduate in a certain period or not) but more related to the time devoted to graduate. We are interested in the analysis of the particular path each student had and how that might be related to his/her race/ethnicity and institution.

In terms of institutions, there have been two publications that have utilized multilevel analysis, authored by Ricco and Salzman, Long, and Ohland [28, 29]. The first study looked into the grades of core chemistry, calculus, and physics courses and their relationship with sections with wide variations among institutions and subjects, being calculus the least variating subject (probably due to its 'right answer' type of assessment). The second study looked into the class sizes of the sections affecting grades on the same subjects and found a low variability between them.

1.3 Time to Degree Metric

Time to degree has been a metric of performance and success for colleges and universities. However, the interpretation differs according to the context and interest of analysis. Tinto created a theoretical framework that guided research on this metric with the Interactionalist Theory [30]. This theory establishes two dimensions of commitment that affect students' progress to degree completion, the student commitment and the institutional commitment [31]. To persist, the students need to engage in both formal and informal features of higher education. In the formal feature, the student must integrate academic performance and extracurricular activities. In the informal feature, the student should interact with faculty and staff in the academic systems and also interact with peers. The theory principles are based on the dynamic interaction between the academic and social systems of an institution with its students and their background characteristics.

Although with MIDFIELD data we have the opportunity to explore different variables in engineering education, this project focuses on studying the time to degree metric among FTIC students of different races/ethnicities and institutions utilizing multilevel analysis.

2. Method

2.1 The Data

This research performed a secondary data analysis on the MIDFIELD data. As mentioned before, the MIDFIELD database includes institutional data from all undergraduate degree-seeking students at 19 different institutions (n = 1606962) in the U.S. For the purposes of this study, *students* and *degree* were merged by a unique anonymized MIDFIELD student identifier (n = 809468). Fig. 1 provides a visual representation of the dataset with sample sizes. Only three programs were considered using the Classification of Instructional Programs (CIP) IDs from the MIDFIELD dataset, as this research focused on undergraduate engineering programs where students matriculated. The selected engineering programs from the dataset were, (1) computer and information sciences and support services (CIP 11), (2) engineering (CIP 14), and (3) engineering technologies and engineering-related fields (CIP 15).

This study focused on First time in College

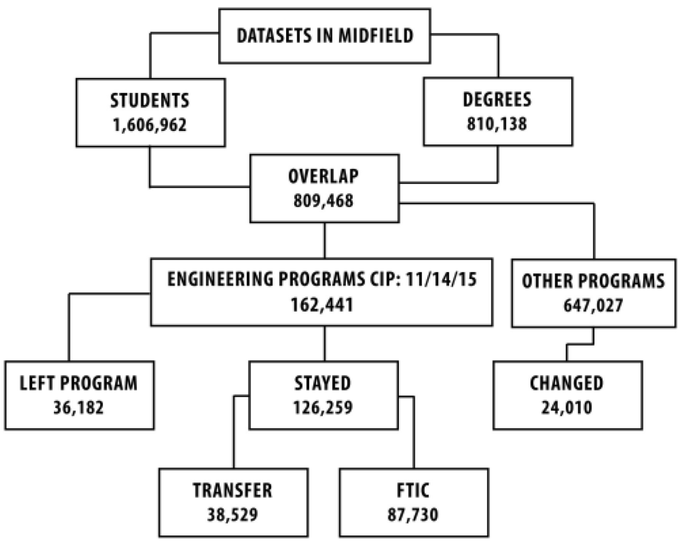

Fig. 1. Research specific MIDFIELD data classification.

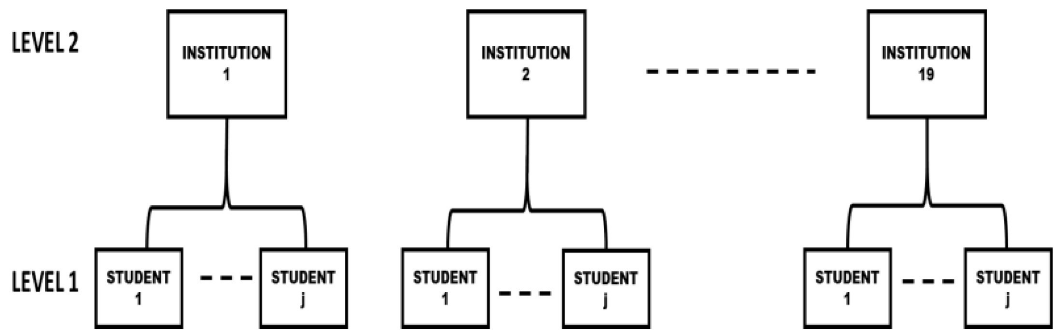

Fig. 2. Two-level multilevel resilient FTIC data model.

(FTIC) students as these students in terms of persistence are different from transfer students [32–33]. The dataset that was analyzed consisted of FTIC students who were resilient or persistent and had earned their bachelors in engineering. Resilient or persistent are terms used to define students who were admitted into engineering and remained in the program until graduation. The missing percent was less than 3% and the statistical analysis would not be biased as such for this already collected data [34]. Without compromising on power by treating the data through a listwise deletion, a multilevel analysis was performed on the final sample of FTIC resilient undergraduate bachelors in engineering students (n = 87589). The analysis was performed using the programming language R with the lme4 package [35, 36].

The outcome or dependent variable for this study is 'time to graduate' which is operationalized as the number of terms taken by a student to eventually complete the engineering program of study from the date of entry into the program. The predictor or independent variable of interest was the race of the students in the sample. This variable was coded as 0 = White, 1 = Hispanic/Latino, 2 = Black, 3 = Asian, 4 = Native American, 5 = International, 6 = Other/Unknown. The programming language R uses the lowest level as the reference level so the interpretations of our analysis will be made with reference to this level (0, White).

2.2 Research Questions

The research questions that primarily guided this study are (1) How do the FTIC resilient Hispanic/ Latino students differ, in terms of the number of terms taken to graduate, in comparison to White students at different institutions? And (2) How do FTIC resilient students from other racial backgrounds compare to White on the same outcome at different institutions?

For quantitative multilevel analysis purposes, these two research questions can be translated into the following specific sub-questions:

(a) How much of the variation in the number of terms taken to graduate is attributed to institutions?
(b) What is the estimate of the within and between institution variance in the number of terms taken to graduate when race is considered?
(c) What is the relationship between the number of terms taken to graduate and the student's race?

2.3 Multilevel Model

The multilevel model considered for this study is a two-level model as shown in Fig. 2. In this model, students are nested within institutions. Students are in the first level (level 1) also called the individual level and institutions are in the second level (level 2) also called the group level.

The above model is analytically represented as,

(1) Null or empty model or random intercept model:

A null model is a model with no independent variables. This model is used to calculate the intraclass correlation coefficient and is the basis for all multilevel models. Mathematically, the model is represented as,

$$\text{Level 1:} \quad Y_{ij} = \beta_{0j} + r_{ij} \quad r_{ij} \sim N(0, \sigma_{0|}^2) \quad (1)$$

$$\text{Level 2:} \quad \beta_{0j} = \gamma_{00} + \mu_{0j} \quad \mu_{0j} \sim N(0, \tau_{0|}^2) \quad (2)$$

Where Y_{ij} is the number of terms taken to graduate for student 'i' in institution 'j'. The element β_{0j} is the intercept (average of the outcome) specific to institution 'j'. The element r_{ij} is the variability of student 'i' around the respective institutional average or the student specific deviation and is distributed as $N(0, \sigma_0^2)$. γ_{00} is the overall grand average across all institutions and μ_{0j} is the variability of institutional (j) average across the overall grand average or the school specific deviation and is distributed as $N(0, \tau_0^2)$. For further clarity, the above is represented visually in Fig. 3 and it helps see the parameters in Equation (1) and Equation (2). Because it is the random intercept model, the intercepts are different for each institution and the slopes are a constant. For simplicity, in Fig. 3 only two institutions are shown yet for this study data comes from 19 institutions.

The intraclass correlation coefficient (ICC) as shown in Equation (3) estimates the percent of variance in the outcome or dependent variable (number of terms taken to graduate) that is attributable to the differences among institutions. In other words, it is the proportion of variance in the dependent variable accounted by the group level [37]. Larger ICC values are indicative of a greater impact of clustering [38]. In simpler terms it shows how much differences between 19 institutions contribute to explaining the changes in the outcome variable. The ICC value for educational performance often lies between 0.10 to 0.25 [39].

$$\mathrm{ICC} = \frac{\tau_0^2}{(\tau_0^2+\sigma_0^2)} \quad (3)$$

Where τ_0^2 is the variance in the outcome between institutions also called the variability of the institutional mean around the overall mean and σ_0^2 is the variability in the outcome within institutions also called the variability of students around the respective institutional mean. Simple stated, it indicates how much the student varies in the outcome within each institution.

(2) Model with student level variable-race (conditional model):

A conditional model is a model with independent variables. In this study, the variable of interest is race. We are interested in seeing the relationship between race and the outcome variable which is the number of terms taken to graduate. We analyzed two models, first a model with no change in slopes or regression coefficients for the level 1 variable (race) across the institutions. Second, we varied the slopes (random slope) for the race variable across the institutions. Both models were compared to arrive at the final results. Mathematically the model with fixed slope is shown below,

$$\text{Level 1:}\quad Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + r_{ij} \quad r_{ij} \sim N\left(0, \sigma_{0|}^2\right) \quad (4)$$

$$\text{Level 2:}\quad \beta_{0j} = \gamma_{00} + \mu_{0j} \quad \mu_{0j} \sim N\left(0, \tau_{0|}^2\right) \quad (5)$$

Where β_1 in (4) is the regression coefficient of the student level variable race (X_{ij}) for student 'i' at institution 'j'.

The random slope model is mathematically shown below,

$$\text{Level 1:}\quad Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + r_{ij} \quad r_{ij} \sim N\left(0, \sigma_{0|}^2\right) \quad (6)$$

$$\text{Level 2:}\quad \beta_{0j} = \gamma_{00} + \mu_{0j} \quad (7)$$

$$\vec{\mu} \sim MVN(0, \vec{\tau})$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \quad \mu_{1j} \sim N\left(0, \tau_{1|}^2\right) \quad (8)$$

Where β_{1j} in (6) and (8) is the regression coefficient for the variable race and it represents the relation-

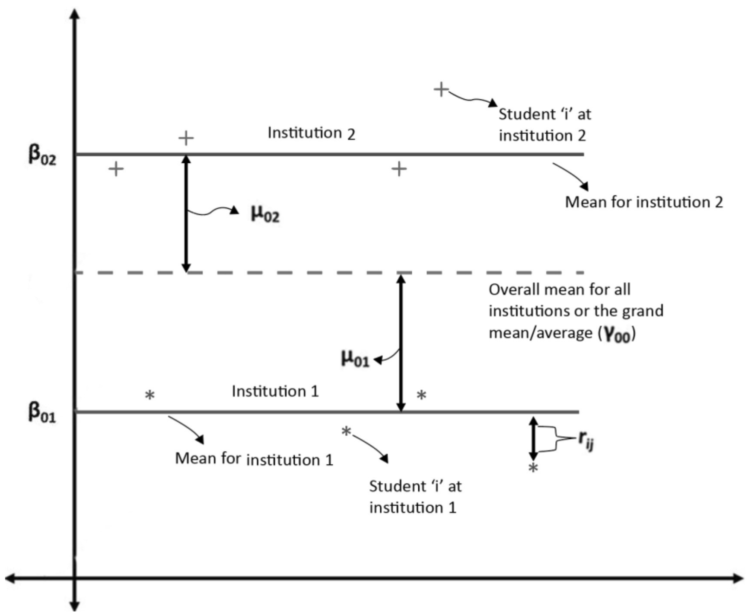

Fig. 3. Visual representation of the null model.

ship between race and number of terms taken to graduate and is changing for different institutions. The element μ_{1j} in (6) represents the institution specific deviation in terms of how race is related to the number of terms taken to graduate. These unobserved effects have a multivariate normal distribution (MVN). In addition to capturing the variance in institution specific mean and the variance in the slope, it also now represents the covariance between these random effects [37].

2.4 Model Comparison

Models were compared using two criterions namely the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion in R. The model with the smaller AIC and BIC values indicates a better fit and was selected as the best model to explain the relationship between race and the number of terms taken to graduate for this sample of FTIC resilient bachelors in engineering [39].

3. Results

The demographic characteristics and descriptive statistics for FTIC resilient undergraduate bachelors in engineering, sample of interest in terms of race and gender from the MIDFIELD data, is shown in Table 1. The students were under 25 years of age.

Table 1. Demographics and descriptives for FTIC resilient bachelors in engineering sample

Gender/Race	**% (n)**
Female	19 (16280)
Male	81 (71309)
White	77 (67162)
Hispanic/Latino	3 (2684)
Black	8 (7497)
Asian	6 (5684)
Native American	1 (389)
International (non-U.S.)	3 (2381)
Other/Unknown	2 (1933)
Variable	**Terms Mean (SD)**
Overall	24.85 (6.98)
Female	24.10 (5.77)
Male	25.03 (7.22)
White	24.80 (6.88)
Hispanic/Latino	26.50 (7.18)
Black	26.00 (7.63)
Asian	24.24 (6.77)
Native American	25.73 (9.53)
International (non-U.S.)	22.68 (6.17)
Other/Unknown	24.48 (7.41)
No. of institutions (j)	19
No. of students (i)	87589

In order to answer the first sub-research question, the ICC was calculated for an empty model with no independent variables. Table 2 shows the estimates for this model which yielded an ICC of twelve percent of the variation in the number of terms taken to graduate to be attributed to differences between institutions. The standard errors were calculated using parametric bootstrap (Residuals are simulated using multivariate normal distribution. In bootstrapping, 'n' samples are drawn with replacement from the observed sample) in R for 100 iterations [40].

For the overall distribution of the number of terms taken to graduate, the estimates in the empty model provide a mean of 24.67 and a standard deviation of 7.23. This mean is the expected value of the number of terms taken to graduate for a random student from a randomly drawn institution. This is slightly different from the raw mean shown in Table 1 as the estimation of the model implies a weighting of the various institutions which is not considered for the raw calculation of mean [37].

Race was added as a variable in level-1 (student level) in order to answer the second and third sub-research questions; which look at the relationship between race and number of terms taken to graduate. Table 3 shows the parameter estimates for this fixed slope model.

It can be observed from Table 3 that after accounting for the impact of race, the estimate of variance in intercepts between institutions is 6.48 while the within-institution (or between student)

Table 2. Estimates for empty model

Fixed Effect	**Coefficient**	**S.E.**[a]
γ_{00} = Intercept	24.67	0.59
Random Part	**Variance Component**	**S.E.**
Level-two variance: τ_0^2	6.50	2.20
Level-one variance: σ_0^2	45.99	0.25

[a] S.E. is the standard error.

Table 3. Estimates for the conditional model/fixed slope

Fixed Effect	**Coefficient**	**S.E.**
γ_{00} = Intercept	24.40*	0.59
β **= Coefficient of Race**		
Hispanic/Latino	1.43*	0.13
Black	2.17*	0.11
Asian	–0.55*	0.09
Native American	1.10*	0.34
International	–1.97*	0.14
Other/Unknown	–0.30	0.16
Random Part	**Variance Component**	**S.E.**
Level-two variance: τ_0^2	6.48	2.17
Level-one variance: σ_0^2	45.6	0.23

* Statistically significant relationship, $p < 0.05$.

variation is 45.60. These variances are a little lower than the empty model as some differences are now partially explained by the explanatory variable, race. Also, there is a significant difference in the number of terms taken to graduate for Hispanic/Latino in comparison to White (t = 10.62, $p < 0.05$) with number of terms taken to graduate for Hispanics increasing by 1.43. Similarly, for Blacks in comparison to White (t = 19.24, $p < 0.05$) it increases by 2.17, and for Native Americans (t = 3.13, $p < 0.05$) it increases by 1.08 conditional on other races respectively. In comparison to White students, Asians (t = –5.86, $p < 0.05$) and International students (t = –13.65, $p < 0.05$) differ significantly in that the number of terms taken to graduate decrease by 0.56 and 1.95 respectively.

The relationship between race and number of terms taken to graduate, third sub-research question, was also analyzed by changing the slope for the race variable across the institutions as recorded in Table 4.

As seen in Table 4, there was only significant main effect of race for Black students in comparison to White students (t = 10.84, p < 0.05) for the number of terms taken to graduate in this model. The number of terms taken to graduate for Black students increases by 2.39 compared to White students. Hispanics/Latino and Native Americans students also take longer than White students to graduate, however it is not significant. This random slope model with the regression coefficients for race changing across the 19 institutions was a better model as it had a lower AIC and BIC value as shown in Table 5.

Table 4. Estimates for the conditional model/random slope

Fixed Effect	**Coefficient**	**S.E.**
γ_{00}= Intercept	24.30*	0.63
β = Coefficient of Race		
Hispanic/Latino	0.88	0.59
Black	2.39*	0.22
Asian	–0.15	0.25
Native American	1.07	0.69
International	–0.70	0.56
Other/Unknown	–0.20	0.25
Random Part	**Variance Component**	**S.E.**
Level-two variance: τ_0^2	7.09	2.23
Slope Variance		
Hispanic/Latino	5.10	–
Black	0.60	–
Asian	0.65	–
Native American	5.42	–
International	4.55	–
Other/Unknown	0.54	–
Level-one variance: σ_0^2	45.13	0.21

* Statistically significant relationship, $p < 0.05$.

Table 5. Model comparison

Model Name	**AIC**	**BIC**
Fixed slope	583259	583344
Random slope	582519	582857

It can be observed that while 12% of the variance for the number of terms taken to graduate is attributable to differences between institutions, in the analysis between the institutions in MIDFIELD, only FTIC resilient bachelors in engineering students who identified themselves as Black significantly differed from White in terms of the number of terms taken to graduate. Statistically this model was also the best model. We also observe that while not significant, consistent with theory, students who identified as Hispanic/Latino and Native American took longer to graduate in comparison to White students. Students who identified as Asians, Internationals and other/unknown took lesser number of terms to graduate compared to White students as is evidenced by the respective negative slope coefficients in Tables 3 and 4. This was consistent to both models.

4. Discussion

The variation in the number of terms taken to graduate attributed to institutions (first sub-research question) was 12% which validated the subsequent analysis based on race. The multilevel empty model showed a slight difference in the number of terms taken compared to the average number of terms in the descriptive statistics, something we expected. From the descriptive statistics, it is worth noting that the representation of Latinos/Hispanics in the sample is of 3% which by no account represents the participation of this group in Engineering or in the U.S. population. Nevertheless, we noticed this fact early in our study, we decided to continue with our analysis.

In terms of the estimate of the within and between institutions variance in the number of terms taken to graduate when race is considered (second sub-research question), we considered the random intercept model and also the random slope model. In this model we found a significant difference in terms of the numbers taken to graduate for Hispanic/Latino in comparison to White, with number of terms taken to graduate significantly . This is something we expected given the disadvantages this group experience, well documented already in the literature. The same can be said about other under-represented groups such as Black and Native Americans as well as the 'over-represented' groups, Asians and International students.

The third model, the random slope model, produced interesting results. The only significant main effect of race was for Black students in comparison with White students. The differences between White students and Hispanic/Latino and Native Americans were not significant. We attributed this result to the lack of representation of Latinos/Hispanics in the sample and the closer-to-reality representation of Black students in the sample, 8% of participants. As for the Native American representation, we consider the model appropriately representing the institution's variance. As Table 5 shows, the random slope model is the best model of all. The basic assumption made from this model is that Black students have it 'harder' to graduate among the 19 participant institutions than Hispanics and Native Americans. This invites the IJEE readership to contribute to the conversation on plausible causes, such causes could include campus climate, identity development, systemic discrimination/critical race theory, currently being areas of relevant research in engineering education.

As a follow-up it would be interesting to study the effects of adding other student level variables like gender, SES, and academic achievement scores like ACT, SAT, GPA and interactions including cross-level to look at the effect on the number of terms to graduate for Hispanic/Latino students in comparison to students from other races for the MIDFIELD sample. Another follow-up is in terms of qualitative analysis. This type of analysis has the potential of unveiling the root causes at participant institutions that might explain these results.

The limitations of this study include the sampled institutions and the associated sample sizes, specifically for Latinos/Hispanics. With a more representative sample, results may vary. While this study is also limited in its type of analysis in that it uses a basic multilevel model, we would like to progress into other models for the next phases of this study by adding other variables of interest as mentioned previously both at the student level and institutional level and look at cross level interactions for example institutional SES and gender and race for number of terms to graduate. Another limitation of the study is that with this being secondary data, the only information we have regarding the age of the students is that they are under 25. This also is a limitation as we cannot study how the maturity of the students with age would predict the number of terms to graduate for Hispanic/Latino students and look at interactions of age and gender.

5. Conclusions

The purpose of this study was to investigate the time Latinos/Hispanics take to graduate in engineering and computer science and to compare it against other demographic groups in the context of the institutions participating in MIDFIELD. In this context, the findings suggest that underrepresented groups, including Latinos/Hispanics have an overall significant disadvantage compared to White, Asian and International students.

When analyzed with multilevel analysis, the disadvantage across all institutions remained significant but when analyzed between institutions, the African American resulted as the only group that significantly remained at a disadvantage.

Multilevel analyses are very powerful tools when characterizing large groups of data. In this case, the groups of data correspond to institutions. The results reported in this study provide a useful snapshot of the challenges minorities face in the United States. For the specific case of Latinos/Hispanics, the 19 MIDFIELD participant institutions could not provide the desired representation of Latinos/Hispanics in the USA (18%) or the Colleges of the USA awarding engineering and computer science degrees to Latinos/Hispanics (10%). That in itself constitutes another challenge faced by Latinos/Hispanics; without data reporting aspects of this group of the population, there is no way to understand or improve their representation in STEM. Authors expect this study to provide a basis for further studies in terms of African American, Latinos/Hispanics and Native American social justice.

Acknowledgments – The authors would like to express their gratitude for the support of Dr. Matthew Ohland, Dr. Russell Long, and Mr. Hossein Ebrahiminejad at Purdue University. MIDFIELD data was obtained under the auspices of the NSF #1545667-Expanding Access to Participation in the Multiple Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD) award.

References

1. Pew Research Center, https://www.pewresearch.org/fact-tank/2020/07/07/u-s-hispanic-population-surpassed-60-million-in-2019-but-growth-has-slowed, Accessed 5 August 2020.
2. National Science Foundation, https://ncses.nsf.gov/pubs/nsf19304/digest/field-of-degree-minorities#hispanic-or-latino-graduates, Accessed 5 August 2020.
3. National Science Foundation, https://www.nsf.gov/statistics/2018/nsb20181/report/sections/higher-education-in-science-and-engineering/the-u-s-higher-education-system,Accessed 5 August 2020.

4. D. Santiago, J. Laurel, J. Martinez, C. Bonilla and E. Labandera, Latinos in higher education: compilation of fast facts, *Latinos in Higher Education: Compilation of Facts*, Excelencia in Education, Washington, DC. https://files.eric.ed.gov/fulltext/ED595093.pdf.
5. National Center for Education Statistics, https://nces.ed.gov/programs/raceindicators/indicator_RED.asp, Accessed 5 August 2020.
6. National Center for Education Statistics, https://nces.ed.gov/pubs2014/2014001rev.pdf, Accessed 5 August 2020.
7. E. M. Bensimon, A. C. Dowd, R. Stanton-Salazar and B. A. Dávila, The role of institutional agents in providing institutional support to Latinx students in STEM, *The Review of Higher Education*, **42**(4), pp. 1689–1721, 2019.
8. C. Cruz, G. Rajpal, M. Lecocke, I. Martines and A. Lurie, Peer coaching program development: a framework of first-year latina/o student persistence pursuing STEM pathways at a Hispanic serving institution, *Journal of Hispanic Higher Education*, https://doi.org/10.1177/1538192719867096, 2019.
9. M. R. Kendall, M. Denton, N. H. Choe, L. M. Procter and M. Borrego, Factors influencing engineering identity development of Latinx students, *IEEE Transactions on Education*, **62**(3), pp. 173–180, 2019.
10. L. Pawley, Learning from small numbers: studying ruling relations that gender and race the structure of U.S. engineering education, *Journal of Engineering Education*, **108**(1), pp. 13–31, 2019.
11. S. Haag, N. Hubele, A. Garcia and K. McBeath. Engineering undergraduate attrition and contributing factors, *International Journal of Engineering Education*, **23**(5), pp. 929–940, 2007.
12. J. P. Martin, M. K. Miller and D. R. Simmons. Exploring the theoretical social capital "deficit" of first generation college students: Implications for engineering education, *International Journal of Engineering Education*, **30**(4), pp. 822–836, 2014.
13. M. W. Ohland and R. A. Long, The multiple-institution database for investigating engineering longitudinal development: an experiential case study of data sharing and reuse, *Advances in Engineering Education*, **5**(2), pp. 1–25, 2016.
14. S. M. Lord, M. W. Ohland, R. A. Layton and M. M. Camacho, Beyond pipeline and pathways: Ecosystem metrics, *Journal of Engineering Education*, **108**(1), pp. 32–56, 2019.
15. M. W. Ohland, S. M. Lord and R. A. Layton, Student demographics and outcomes in civil engineering in the United States, *Journal of Professional Issues in Engineering Education and Practice*, **141**(4), 04015003, 2015.
16. M. W. Ohland, C. E. Brawner, M. M. Camacho, R. A. Layton, R. A. Long, S. M. Lord and M. H. Wasburn, Race, gender, and measures of success in engineering education, *Journal of Engineering Education*, **100**(2), pp. 225–252, 2011.
17. E. G. Shealy, C. E. Brawner, C. Mobley and R. Layton, A descriptive study of engineering transfer students at four institutions: Comparing lateral and vertical transfer pathways, In *120th ASEE Annual Conference and Exposition*, Atlanta, GA, 2013.
18. C. E. Brawner, M. M. Camacho, R. A. Long, S. M. Lord, M. W. Ohland and M. H. Wasburn, Work in progress-the effect of engineering matriculation status on major selection, *39th IEEE Frontiers in Education Conference*, 18–21 October 2009, San Antonio, TX, pp. 1–2, 2009.
19. S. M. Lord, C. E. Brawner, M. M. Camacho, R. A. Layton, R. A. Long, M. W. Ohland and M. H. Wasburn, Work in progress-effect of climate and pedagogy on persistence of women in engineering programs, *38th Annual Frontiers in Education Conference*, 2–25 October 2008, Saratoga Springs, NY, pp. S4F-21, 2008.
20. M. W. Ohland, M. K. Orr, R. A. Layton, S. M. Lord and R. A. Long, Introducing "stickiness" as a versatile metric of engineering persistence, *2012 Frontiers in Education Conference Proceedings*, 3–6 October 2012, Seattle, WA, pp. 1–5, 2012.
21. S. M. Lord, R. A. Layton and M. W. Ohland, A disciplinary comparison of trajectories of USA engineering students, *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, 22–25 October 2014, Madrid, Spain, pp. 1–4, 2014.
22. C. E. Brawner, S. M. Lord, R. A. Layton, M. W. Ohland and R. A. Long, Factors affecting women's persistence in chemical engineering. *International Journal of Engineering Education*, **31**(6), pp. 1431–47, 2015
23. S. M. Lord, R. A. Layton and M. W. Ohland, Multi-institution study of student demographics and outcomes in electrical and computer engineering in the USA, *IEEE Transactions on Education*, **58**(3), pp. 141–150, 2014.
24. R. Layton, M. W. Ohland and S. M. Lord, Comparing student outcomes for women and men in electrical engineering to civil, chemical, industrial, and mechanical engineering in the USA, 2017 27th *EAEEIE Annual Conference (EAEEIE)*, Grenoble, France, pp. 1–5, 2017.
25. M. K. Orr, I. Ngambeki, R. A. Long and M. W. Ohland, Performance trajectory of students in the engineering disciplines. 2011 *Frontiers in Education Conference (FIE)*, Rapid City, SD, pp. S3H-1–S3H-4, 2011.
26. S. M. Lord, M. M. Camacho, R. A. Layton, R. A. Long, M. W. Ohland and M. H. Wasburn, Who's persisting in engineering? A comparative analysis of female and male Asian, black, Hispanic, Native American, and white students, *Journal of Women and Minorities in Science and Engineering*, **15**(2), pp. 167–190, 2009.
27. M. W. Ohland, S. D. Sheppard, G. Lichtenstein, O. Eris, D. Chachra and R. A. Layton, Persistence, engagement, and migration in engineering programs, *Journal of Engineering Education*, **97**(3), 259–278, 2008.
28. G. D. Ricco. Beyond Sectionality and into Sizeness or How Course Size Effects Grades: An Exploration of the MultipleInstitution Database for Investigating Engineering Longitudinal Development through Hierarchal Linear Models. *age*, *26*, 1. *Proceedings ASEE Seattle*. Washington, 2015.
29. G. D. Ricco, N. Salzman, R. A. Long and M. Ohland. Sectionality or Why Section Determines Grades: an Exploration of Engineering Core Course Section Grades using a Hierarchical Linear Model and the Multiple-Institution Database for Investigating Engineering Longitudinal Development, *Proceedings ASEE*, San Antonio, 2012.
30. V. Tinto. Dropout from higher education: A theoretical synthesis of recent research, *Review of Educational Research*, **45**, pp. 89–125, 1975.
31. J. M. Braxton and A. S. Hirschy, Theoretical developments in the study of college student departure, *College student retention: Formula for student success*, **3**, pp. 61–87, 2005.
32. L. V. Tuttle and G. D. Musoba, Transfer student persistence at a Hispanic-serving university, *Journal of Latinos and Education*, **12**(1), pp. 38–58, 2013.
33. G. N. Mitchell and C. L. Grafton, Comparative study of reverse transfer, lateral transfer, and first-time community college students, *Community Junior College Research Quarterly of Research and Practice*, **9**(3), pp. 273–280, 1985.
34. D. A. Bennett, How can I deal with missing data in my study, *Australian and New Zealand Journal of Public Health*, **25**(5), pp. 464–469, 2001.

35. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
36. D. Bates, M. Mächler, B. Bolker and S. Walker, Fitting linear mixed-effects models using lme4, *Journal of Statistical Software*, **67**(1), pp. 1–48, 2015.
37. T. A. B. Snijders and R. J. Bosker, *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.), London: Sage, pp. 17–18, 2012.
38. W. H. Finch, J. E. Bolin and K. Kelle, *Multilevel modeling using R*, London: CRC Press, pp. 24–28, 2014.
39. L. V. Hedges and E. C. Hedberg, Intraclass correlation values for planning group-randomized trials in education, *Educational Evaluation and Policy Analysis*, **29**(1), pp. 60–87, 2007.
40. J. Hox, *Multilevel analysis, techniques and applications*, New Jersey: Lawrence Erlbaum Assoc, Inc., p. 208, 2002.

Noemi V. Mendoza Diaz is Assistant Professor at the College of Education and Human Development with a courtesy appointment in the College of Engineering at Texas A&M University. She obtained her PhD from Texas A&M University in Educational Administration and Human Resource Development, her Master of Science and her Bachelors in Electronics and Communications from the National Polytechnic Institute in Mexico City. She worked as a Postdoctoral Researcher with the Institute for P-12 Engineering Research and Learning-INSPIRE at the School of Engineering Education-Purdue University where she researched and published in K-12 Engineering Education. She was a recipient of the Apprentice Faculty Grant from the Educational Research Methods ASEE Division in 2009. She also has been an Electrical Engineering Professor for two Mexican universities and first year-engineering instructional professor at Texas A&M. She is co-advisor of the Society for Hispanic and Professional Engineers at Texas A&M. Dr. Mendoza is interested in computational thinking, sTEm education, socioeconomically disadvantaged students, and Latino studies in engineering.

Cijy Elizabeth Sunny is currently a PD Research Associate in the Department of Information Systems and Business Analytics, Hankamer School of Business at Baylor University. She is a research methodologist and psychometrician who has applied her skills in quantitative and mixed methods research methodology in the substantive areas of STEM education research, medical education, and more recently in engineering education. Her research interests include secondary data analysis on large scale datasets like MIDFIELD to study trends longitudinally using structural equation modeling and multilevel models to understand underrepresented minority student enrollment, retention, and trajectory in engineering by applying varied theoretical frameworks, design, develop, and validate cognitive and non-cognitive (attitude, persistence, etc.) assessments using the psychometric frameworks and investigate relationship(s) between latent factors that foster engagement, learning, and retention in engineering education research, and investigating bias via measurement invariance across subgroups like race, gender, social identities, etc. to ensure the attributes measured are constant across and not a function of sub-group membership. Additionally, she has been an educator and has taught primarily physics and also research methodology on three different continents. In addition to research, she has also conducted workshops on using concept mapping methodology for scale development, mixed methods research methodology for standardized patient educators, and standard-setting for physician educators. Dr. Sunny continues to invest her skills in engineering education research through her collaborations. She is the 2020 FIE New Faculty Fellow and recipient of the 2021 ASEE ERM Apprentice Faculty Grant. As part of her new undertaking at Baylor University, she is investing her skills as a research methodologist and data analyst to fight human trafficking through the use of Information Technology working alongside the research team there in collaboration with a diverse group of stakeholders.

Trinidad Sotomayor is a graduate student at Pontificia Universidad Católica de Chile (UC) where she is currently working towards her master's degree in engineering. She also received her bachelor's degree from UC in mechanical engineering with a major in design and innovation. Trinidad is part of the DILAB (UC) engineering design initiative, where she works as a researcher on issues within engineering education such as minorities, epistemologies, entrepreneurship, and more. Her research work centers on qualitative and quantitative methods like data analysis, statistical inference, and psychometrics. Trinidad received a scholarship and had the opportunity to work as a researcher in the School of Engineering at Texas A&M. At A&M, she worked as a data analyst and studied the educational experience of racial, gender, and socioeconomic minorities with the MIDFIELD database. In addition to her research interests, Trinidad enjoys teaching; hence, she has been a university teaching assistant in design, innovation, and research courses for three years.

Jacques Richard obtained his PhD from Rensselaer Polytechnic Institute, 1989 & a BS at Boston University, 1984. He was at NASA Glenn, 1989–1995, taught at Northwestern for Fall 1995, worked at Argonne National Lab, 1996–1997, Chicago State, 1997–2002. Dr. Richard is a Instructional Associate Professor in Engineering @ Texas A&M since 1/03. His research is focused on computational plasma modeling using spectral and lattice Boltzmann methods for studying plasma turbulence and plasma jets. His research has also included fluid physics and electric propulsion using Lattice-Boltzmann methods, spectral element methods, Weighted Essentially Non-Oscillatory (WENO), etc. Past research includes modeling single and multi-species plasma flows through ion thruster optics and the discharge cathode assembly; computer simulations of blood flow interacting with blood vessels; modeling ocean-air interaction; reacting flow systems; modeling jet engine turbomachinery going unstable at NASA for 6 years (received NASA Performance Cash awards). Dr. Richard also conducts engineering education research. Dr. Richard also studies how emerging technology can impact engineering education (e.g., eTextbooks with embedded simulations) and the complex correlation between instructional material and student development. Dr. Richard is involved in many outreach activities: e.g., tutoring, mentoring, directing related grants (for example, a grant for an NSF REU site). Dr, Richard is active in professional societies (American Physical Society (APS), American Institute for Aeronautics and Astronautics (AIAA), ASEE, ASME. Dr. Richard has authored or co-authored about 25 technical articles (21 of which are refereed publications). Dr. Richard teaches courses ranging from first-year introductory engineering project design, fluid mechanics, to space plasma propulsion.