WIP: Exploring STEM students' enrollment in entrepreneurship education programs: a binary logistic regression approach

Carlos Felipe Rodriguez-Hernandez
Postdoctoral Research Associate
New Jersey Institute of Technology
NJ, USA
carlosfelipe.rodriguezhernandez@njit.edu

Prateek Shekhar
Assistant Professor
New Jersey Institute of Technology
NJ, USA
prateek.shekhar@njit.edu

***Abstract*—This research WIP paper investigates STEM students' enrollment in entrepreneurship education programs (EEPs). Previous research has suggested that STEM careers are important for the economic and technological development of today's society. To prepare students for STEM careers, a growing body of literature indicates that EEPs allow STEM students to develop skills to be competitive in a wide range of STEM careers and industries. Despite the benefits described above, enrollment of STEM students in EEPs has received less attention among researchers. This lack of research requires attention because it limits the understanding of what STEM students are gaining from these elective EEPs. The objective of this study is to enhance this understanding by exploring the predictive relationship between STEM students' background (i.e., major, undergraduate GPA, socioeconomic status, race/ethnicity, and sex) and their enrollment in EEPs through a binary logistic regression model. Binary logistic regression is commonly used for predicting the probability of the outcome in binary classification tasks (i.e., enrolled or not enrolled in EEPs). The data set included records from 23,411 undergraduate STEM students enrolled at a public research university in the United States. The major distribution is as follows: Science (14%), Technology (23.9%), Engineering (60.3%), and Mathematics (1.8%). The binary logistic regression results identify notable classification patterns. First, when compared to Engineering students, Science and Mathematics students are less likely to enroll in EEPs, but Technology students are more likely. Second, students' GPA is negatively related to their enrolment in EEPs. Third, students coming from low SES are more likely to enroll in EEPs. Finally, students' race and sex seem to be non-significantly related to their enrolment in EEPs. We provide implications of these findings, highlighting key preliminary results and the future steps we intend to perform in this research project.**

Keywords—entrepreneurship education, logistic regression

I. INTRODUCTION

STEM education initiatives aim to develop the skills in students that are needed to be competitive in today's world [1]. While the need for STEM students to acquire technical and scientific knowledge seems obvious, STEM students also need to develop professional skills such as proficiency in communication, business, creativity, and leadership [2]. In other words, the STEM workforce demands that professionals demonstrate additional sets of skills beyond the technical ones [3]. Entrepreneurial skills are one of these additional sets of skills that are gaining more attention among STEM students, educators, and employers [4].

Entrepreneurship education and training are critical to creating and transferring entrepreneurial skills to students and trainees [5]. In consequence, entrepreneurship education has received increased attention in universities across the United States and worldwide, with STEM and entrepreneurial programs sharing similar objectives for the students [2]. In this way, entrepreneurship education contributes STEM students to be innovative and entrepreneurial citizens [4].

Previous research on entrepreneurship education for science and technology students has been limited [6]. However, recent findings have indicated that entrepreneurship education programs had a positive effect on the entrepreneurial mindset change of STEM university students [3]. Furthermore, participating in an entrepreneurship course has been noted to positively influence feasibility, desirability, and intentions for entrepreneurship among STEM students [7] However, the factors influencing STEM students' enrollment in entrepreneurship education programs remain significantly underexplored. As the offering of entrepreneurship education programs continues to increase, it is important to identify what type of students benefit the most from these courses [8]. We argue that this understanding may contribute to maximizing the intended impact of entrepreneurship education programs by increasing STEM students' enrollment in entrepreneurship education courses.

Astin's theory of involvement [9] frames this work-in-progress study from a theoretical perspective. Students' involvement can be understood as the physical and psychological energy students invest in their university experience. In particular, students' involvement is observed through their academic work, participation in activities, and interaction with faculty members and other students. The theory states that greater involvement is likely to impact student's learning and development [9]. In the presented context, this implies that involvement in entrepreneurship education programs (EEPs) is likely to contribute to the development of entrepreneurship-related skills and mindsets among STEM students as evident in the literature [2], [10].

This paper studies involvement by exploring how STEM students' backgrounds (i.e., major, undergraduate GPA, socioeconomic status, race/ethnicity, and sex) relate to their enrollment in entrepreneurship education programs. The results derived from this preliminary analysis would allow us to gain more insights into several STEM students' characteristics interaction when predicting their enrollment in EEPs. From the methodological perspective, we have framed the exploratory analysis to fit the predictive model based on binary logistic regression, before using more robust models such as decision trees and random forest in the latter phases of our project.

II. Methodology

A. Data description

The dataset for this study was retrieved from institutional records at a public research university in the United States and comprised registers from 23,411 STEM students. Table 1 displays the descriptive statistics of enrollment in EEPs and the variables describing students' backgrounds. The percentage of students who enrolled in EEPs (4.5%) was lower than that of those who did not enroll (95.5%). The students' mean undergraduate GPA was 3.16 (S.D.=.46). The distribution of students' STEM majors was as follows: Science (14%), Technology (23.9%), Engineering (60.3%), and Mathematics (1.8%). Most of the students (70.5%) were from non-underrepresented race/ethnic groups including White and Asian, while the rest of the students (29.5%) were from underrepresented race/ethnic groups such as American Indian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander. There were more male students (80.3%) than female students (19.7%). Most students were from high SES (89.8%) while the rest were from low SES (10.2%). In this study, SES was operationalized as participation in the Educational Opportunity Program (EOP), meaning that participating students were classified as students from low SES, while non-participating students were categorized as high SES students.

TABLE I. DESCRIPTIVE STATISTICS

Variable	Category	Frequency	Percentage
Enrollment in EEPs	No	22360	95.5%
	Yes	1051	4.5%
STEM area	Science	3289	14%
	Technology	5591	23.9%
	Engineering	14116	60.3%
	Mathematics	415	1.8%
Underrepresented minority	No	16507	70.5%
	Yes	6904	29.5%
SES	High	21017	89.8%
	Low	2394	10.2%
Sex	Male	18788	80.3%
	Female	4623	19.7%

A binary logistic regression of STEM students' enrollment in EEPs was performed using statistical software (through SPSS version 29). Binary logistic regression is used for analyzing the relationship between multiple independent variables (categorical or continuous) and a binary dependent variable that takes values of 1 and 0 [11],[12]. The values of 1 and 0 indicate the presence and absence of the condition being investigated, respectively. In the present study, the dependent variable *Enrolled* is 1 for students who enrolled in EEPs and 0 for students who did not enroll in an EEP. We ensured that the assumptions for binary logistic regression (independence of errors, absence of multicollinearity, and lack of influential outliers) were fulfilled in the predictive model [13],[14].

The binary logistic regression models the odds ratio of the dependent variable based on the interaction of the independent variables. The odds ratio is the ratio between the probability of an event happening and the probability of that same event not happening [15]. Moreover, the odds ratio is an indicator of the effect size of the analyzed variables in a binary logistic regression model [16]. Therefore, the relationship between the dependent variable and the independent variables is expressed in binary logistic regression as given in Equation 1.

$$\text{Log } (p/1\text{-}p) = B_0+B_1x_1+B_2X_2 + B_3X_3 \quad (1)$$

In Equation 1, p indicates the probability of the dependent variable happening, p/1-p indicates the odds ratio, and the regression coefficients (B) represent the odds associated with each independent variable being investigated [15]. For ease of interpretation, we calculated the exponential values of the regression coefficients (exp(B)) to understand the binary logistic regression results [12]. Specifically, we used the exp(B) to examine the influence of independent variables pertinent to student background on the dependent variable (enrollment in EEPs). An exp(B) value larger than one, represents a positive impact of the independent variable, with an increase in the dependent variable related to a greater probability of enrollment in EEPs, and vice-versa.

III. Results

We begin this section by commenting on the assumptions of the binary logistic regression model [13],[14]. First, the scatterplot between predicted probabilities and residuals indicated that errors were independent (i.e., there was no association pattern). Second, there was no indication of severe multicollinearity among the predictors in the dataset. Third, Cook's distance indicated that there were no influential outliers.

TABLE II. BINARY LOGISTIC REGRESSION RESULTS

Variable	B	S.E.	p-value	Exp (B)
Underrepresented minority (Yes)	.050	.071	.483	1.051
Gender (Female)	.085	.083	.305	1.088
SES (Low)	.193	.1	.053	1.213
GPA	-.148	.069	.033	.863
Major (Science)	-1.201	.151	<.001	.301
Major (Technology)	.226	.070	.001	1.253
Major (Mathematics)	-3.027	1.002	.003	.048

Table II summarizes the results of the binary logistic regression model. While GPA, SES, and major were significant predictors of students' enrollment in EEPs, race and gender were not. A further inspection of the significant results in Table 2 noted three key findings. First, a lower GPA was related to an increased likelihood of enrolling in EEPs. Similarly, the students classified as coming from low SES were 1.213 times more likely to enroll in EEPs than their high SES peers. Finally, a comparison across majors revealed that engineering students were respectively 3.32 (1/.301) and 20.83 (1/.048) times more likely to enroll in EEPs than science and mathematics students. Conversely, technology students were 1.253 times more likely to enroll in EEPs than engineering students.

IV. Discussion

The objective of the present work-in-progress study was to explore how STEM students' backgrounds were related to their enrollment in entrepreneurship education programs. To do so, a predictive model of students' enrollment was fitted based on binary logistic regression analysis. Findings revealed that undergraduate GPA, major, and SES were significant predictors of students' enrollment in EEPs, while race and gender were not. The preliminary implications of these findings are presented next.

First, it was found that STEM students' undergraduate GPAs and their enrollment in EEPs were statistically significantly related. This finding is aligned with previous literature that has indicated that undergraduate GPA is a relevant predictor of students' enrollment decisions [17]. However, the negative regression coefficient reported in Table II deserves further comments. Although it seems that better academic performance reduces the likelihood of enrollment, the reasons why low-, middle-, and high-performing students enroll in entrepreneurship courses remain unexplored. Thus, more research is needed to achieve a better understanding of these nuances.

Second, there were differences in the likelihood of students' enrollment according to the STEM major. These differences could be related to different levels of alignment between entrepreneurship and each area. For example, it could be the case that entrepreneurship education courses are more aligned with engineering and technology, in contrast with science and mathematics fields. Undoubtedly, differences in disciplinary perspectives that students might resonate with could also influence students' enrollment in EEPs. Therefore, a comparison across disciplines should be further explored, leading to strengthened entrepreneurship education programs across all STEM areas.

Third, our findings suggest that students belonging to low SES were more likely to enroll in EEPs. In this respect, previous research has indicated that high and low-SES families have different reasons when pursuing entrepreneurship opportunities [18]. However, it is critical to further explore how SES and enrollment are related, by operationalizing SES through multiple indicators such as parents' educational level, parents' occupation, and family income. This would facilitate a deeper understanding of SES background and its relationship with students' enrollment in EEPs.

Fourth, our result suggesting that gender and race do not significantly influence students' enrollment should be interpreted with caution. Previous literature on entrepreneurship education has pointed out that female and male students enroll in entrepreneurship education programs for different reasons [19], [20] and that enrollment of students from underrepresented backgrounds tends to be lower [21]. Considering that the participation of women and ethnic minorities in STEM fields has been limited [22], it seems necessary to continue investigating how gender and race relate to enrollment in entrepreneurship courses. For example, future work could examine the data sets separately for men and women, and contrast differences (if any) in the predictive importance of different academic and socioeconomic variables between the two groups.

This work-in-progress study has some limitations that should be acknowledged. First, the analyzed data set was collected in only one institution, which implies that our results could be less generalizable to other STEM institutions. We suggest increasing the efforts to retrieve data sets at different universities, such that institution-wise comparisons can be conducted. Second, information regarding students' prior academic achievement was not available, thus, this key predictor could be considered in future analysis. Finally, it is to be noted that the dataset was imbalanced with the percentage of students enrolled in EEPs considerably lower than the percentage of non-enrollment. In the upcoming stages of our research, we will consider alternate statistical measures for handling the imbalance issue such as oversampling and undersampling.

V. Conclusion

The present study is part of a broader research effort to understand how students' backgrounds relate to their enrollment in EEPs. Even though the results from this study are preliminary, they contribute to the education research field both at the conceptual and the methodological levels. On the one hand, we hope that our initial results will help educators, leaders, and policymakers to propose better and more effective strategies to increase the participation of STEM students in EEPs. On the other hand, these first results are the basis for further research on the enrollment of STEM students in EEPs, using more complex algorithmic methods such as decision trees and random forests.

VI. Acknowledgment

This work is supported by the US National Science Foundation through grant number 2321175. The opinions are those of the authors and do not necessarily represent the National Science Foundation.

References

[1] A. Sheffield, H. Morgan, and C. Blackmore, "Lessons learned from the STEM Entrepreneurship Academy," *Journal of Higher Education Outreach and Engagement*, vol. 22, no. 3, pp. 185-200, 2018.

[2] S. E. Zappe, S. L. Cutler, and L. Gase, "A systematic review of the impacts of entrepreneurial support programs in Science, Technology, Engineering, and Math fields," *Entrepreneurship Education and Pedagogy*, vol.6, pp. 6-31, 2023.

[3] M. Ruiz González, M. De Hoyos-Ruperto, C. Pomales-García, and M. Amador-Dumois, "Entrepreneurial education program for STEM teachers and students." In *2019 ASEE Southeastern Section Annual Conference*, Raleigh, NC, March 10-12, 2019.

[4] İsa Deveci and Jaana Seikkula-Leino, "The link between entrepreneurship and STEM education," in *Enhancing Entrepreneurial Mindsets Through STEM Education. Integrated Science*, S. Kaya-Capocci and E. Peters-Burton, Eds. Springer, pp. 3-23, 2023.

[5] Y.A. Gangi, "The role of entrepreneurship education and training on creation of the knowledge economy: Qatar leap to the future", *World Journal of Entrepreneurship, Management and Sustainable Development*, vol. 13, no.4, pp. 375-388, 2017.

[6] J. Ferreira, A. Paço, M. Raposo, C. Hadjichristodoulou, and D. Marouchou, "International entrepreneurship education: Barriers versus support mechanisms to STEM students," *Journal of International Entrepreneurship*, vol. 19, no. 1, pp. 130-147, 2021.

[7] A. Pergelova, F. Angulo-Ruiz, T.S. Manolova, and D. Yordanova, "Entrepreneurship education and its gendered effects on feasibility, desirability, and intentions for technology entrepreneurship among STEM students," *International Journal of Gender and Entrepreneurship*, vol. 15, no. 2, 191-228.
[8] E. Lyons and L. Zhang, "The impact of entrepreneurship programs on minorities," *American Economic Review*, vol. 107, no. 5, pp. 303-307, 2017.
[9] A.W. Astin, "Student involvement: A developmental theory for higher education," *Journal of College Student Development*, vol. 40, no. 5, pp. 518-529, 1999.
[10] C.S. Woodcock, P. Shekhar, and A. Huang-Saad, "Examining project-based entrepreneurship and engineering design course professional skills outcomes," *International Journal of Engineering Education*, vol. 35, no.2, pp. 631-644, 2019.
[11] V. Bewick, L. Cheek, and J. Ball, "Statistics review 14: Logistic regression," *Critical Care*, vol. 9, no. 1, pp. 112-118, 2005.
[12] P. Ranganathan, C. S. Pramesh, and R. Aggarwal, "Common pitfalls in statistical analysis: Logistic regression," *Perspectives in Clinical Research*, vol. 8, no. 3, pp. 148-151, 2017.
[13] J. C. Stoltzfus, "Logistic Regression: A brief primer," *Academic Emergency Medicine*, vol. 18, no. 10, pp. 1099-1104, 2011.
[14] J. K. Harris, "Primer on binary logistic regression," *Family Medicine and Community Health*, vol. 9, 2021.
[15] S. Sperandei, "Understanding Logistic Regression Analysis," *Biochemia Medica*, vol. 24, no. 1, pp. 12-18, 2014.
[16] C. Ialongo, "Understanding the effect size and its measures," *Biochemia Medica*, vol. 26, no. 2, pp. 150-163, 2016.
[17] M. E. Engberg and G. C. Wolniak, "College student pathways to the STEM disciplines.," *Teachers College Record*, vol. 115, no. 1, pp. 1-27, 2013.
[18] J. A. Molina, "Family and entrepreneurship: New empirical and theoretical results," *Journal of Family and Economic Issues*, vol. 41, no. 1, pp. 1-3, 2020.
[19] N. Duval-Couetil and Z. Long, "Career impacts of entrepreneurship education: How and when students intend to utilize entrepreneurship in their professional lives," *Journal of Business and Entrepreneurship*, vol. 26, no. 1, pp. 63-87, 2014.
[20] S. Yi and N. Duval-Couetil, "Interdisciplinary entrepreneurship education: Exploring 10-Year trends in student enrollment, interest, and motivation," *Entrepreneurship Education and Pedagogy*, vol. 4, no. 2, pp. 100-118, 2021.
[21] G. A. Cadenas, E. A. Cantú, N. Lynn, T. Spence, and A. Ruth, "A programmatic intervention to promote entrepreneurial self-efficacy, critical behavior, and technology readiness among underrepresented college students," *Journal of Vocational Behavior*, p. 103350, 2019.
[22] G. Dost, "Students' perspectives on the 'STEM belonging' concept at A-level, undergraduate, and postgraduate levels: an examination of gender and ethnicity in student descriptions," *International Journal of STEM Education*, vol. 11, no. 1, 2024.