

## **AC 2008-2383: A METHOD FOR PREDICTING POST-SECONDARY EDUCATIONAL OUTCOMES**

### **Gillian Nicholls, University of Pittsburgh**

Gillian Nicholls is a Ph.D. candidate in Industrial Engineering at the University of Pittsburgh. Her research interests are in applying statistical analysis and optimization to engineering education and transportation management. She holds the B.S. in Industrial Engineering (Lehigh University), Masters in Business Administration (Penn State University), and M.S. in Industrial Engineering (University of Pittsburgh.) Address: 1048 Benedum Hall, University of Pittsburgh, Pittsburgh, PA 15261; telephone 412.400.8631; fax: 412.624.9831; e-mail: gmn3@pitt.edu.

### **Harvey Wolfe, University of Pittsburgh**

Harvey Wolfe is the William Kepler Whiteford Professor of Industrial Engineering at the University of Pittsburgh. After many years working in the area of applying operations research methods to the health field, he is now active in the development of models for assessing engineering education. He is a co-author of *Engineering Ethics: Balancing Cost Schedule and Risk - Lessons Learned from the Space Shuttle* (Cambridge University Press, 1997). He holds the B.E.S. in Industrial Engineering, M.S.E. in Operations Research, and Ph.D. in Operations Research (Johns Hopkins University).

### **Mary Besterfield-Sacre, University of Pittsburgh**

Mary Besterfield-Sacre is an Associate Professor of Industrial Engineering and the Fulton C. Noss Faculty Fellow at the University of Pittsburgh. Her research interests are in engineering education evaluation and empirical and cost modeling applications for quality improvement in manufacturing and service organizations. She holds the B.S. in Engineering Management (University of Missouri Rolla), M.S. in Industrial Engineering (Purdue University), and Ph.D. in Industrial Engineering (University of Pittsburgh).

### **Larry Shuman, University of Pittsburgh**

Larry J. Shuman is Associate Dean for Academic Affairs, School of Engineering, University of Pittsburgh and Professor of Industrial Engineering. His research includes studies directed at the undergraduate engineering learning experience, assessment and the ethical behavior of engineers. Dr. Shuman has published widely in the engineering education literature and is a co-author of *Engineering Ethics: Balancing Cost Schedule and Risk - Lessons Learned from the Space Shuttle* (Cambridge University Press, 1997). He holds the B.S. in Electrical Engineering (University of Cincinnati) and Ph.D. in Operations Research (Johns Hopkins University).

# A Method for Predicting Post-Secondary Educational Outcomes

## Abstract

Identifying potential engineering students and understanding what affects their choice of college major is critical to engineering educational research. Insufficient numbers of students are majoring in Science, Technology, Engineering, or Mathematics (STEM) topics. Understanding the factors that affect students' interest in studying STEM, capability of succeeding in STEM, and likelihood of persisting to achieve a STEM degree is of vital concern to educators.

This study used an extensive national longitudinal dataset of over 12,000 students to develop a set of logistic regression models for predicting which students ultimately achieve a STEM degree vs. another educational outcome. The potential educational outcomes included no college degree, a less than four year college degree, a Non-STEM college degree, a STEM college degree, and a newly proposed category of STEM-Related college degree. Another model comparing the probability of STEM vs. all the other possible outcomes combined was also constructed. The resulting models demonstrated strong predictive accuracy in discriminating between a STEM degree and an alternative educational outcome. The predictive accuracy of the models was examined with Receiver Operating Characteristic (ROC) Curves. Several measures of student academic capability, prior academic performance, attitudes, experiences, and family influences were consistently found to be statistically significant predictors of STEM.

## Introduction

The progression of students through the American educational system from kindergarten to acquisition of a college degree is a lengthy process. At present, the quantity of students that complete degrees in Science, Technology, Engineering, and Math (STEM) is not sufficient<sup>1</sup>. The volume of undergraduates enrolled from 1992 to 2004 increased steadily<sup>2</sup>. However, the pattern of degrees earned from 1996 through 2005 indicates that only small increases have occurred in the bachelors, masters, and doctoral degrees achieved by U.S. citizens and permanent residents. The number of mathematics degrees earned during this period also exhibited little growth. The volume of full-time graduate students increased but the number earning advanced degrees in other science topics exhibited slight increases or declines through 2001<sup>3</sup>.

Producing STEM degree-holders is a process that depends upon the students, educators, and the means by which students are educated. The students are a vital portion of the raw materials to this process and issues that affect their quantity and quality also affect the resulting number of degree-holders. Studying this process in order to identify significant factors that affect the production of degree-holders could provide a guide towards improving the process. A methodology to test the effect of these factors could aid in designing an intervention program to encourage and assist more students in pursuing a college degree in STEM.

Developing such a methodology starts with examining the work of education researchers who have explored the motivations of students and the predictors of student success in school. Variables found by other education researchers to have been significant predictors of STEM interest, persistence, and successes are a natural starting point for this analysis.

Prior research into the degree acquisition process has found that previous academic performance and personal attitudes are good predictors of success in earning a degree. Astin<sup>4</sup> used data from the Cooperative Institutional research Program (CIRP)<sup>5</sup> for a set of 36,581 students at many colleges. He identified high school academic performance and standardized test scores like those of the SAT and ACT as strong predictors of academic performance in college. Zhang, Anderson, Ohland, and Thorndyke<sup>6</sup> reported similar results after examining the records at nine different schools of 87,167 engineering students who attended from 1987 to 1996. The longitudinal study tracked graduation rates and used multiple logistic regression to estimate the time to graduation. Verbal SAT scores, gender, ethnicity, and citizenship were also found to be significant predictors for college graduation.

Smythe and McArdle<sup>7</sup> examined CIRP data for average high school grades, anticipated majors, and SAT scores. The students' college transcripts were obtained separately to test for significance of racial/ethnic and gender variables in STEM degree acquisition. The colleges were ranked by their selectivity, and a series of models were constructed to estimate the effects on the log-transformed odds of STEM graduation. They found that Asian students were the most likely to obtain a STEM degree followed by Caucasian students and then minority students. Female students were less likely to graduate with a STEM degree than male students. SAT math scores, high school grades, and other quantitative measures of academic performance were significant in predicting degree acquisition. Another finding was that the risk of not graduating with a STEM degree rose when students chose selective colleges where their own standardized test scores fell into a low percentile compared to their college cohort. The authors concluded that students were disadvantaged by choosing schools where their individual performance measures compared unfavorably to those of the overall student body.

Besterfield-Sacre, Atman, and Shuman<sup>8</sup> explored student attitudes to determine if they predicted academic performance and persistence in engineering. They designed a methodology for measuring the expectations, self-confidence, and attitudes of freshmen engineering students. They found that students who performed well but chose to leave engineering were often those that expressed less interest in the subject. These students were capable of succeeding but had less motivation to pursue engineering after identifying other subjects that were more personally appealing. In contrast, the students that departed the engineering track after poor performance had tended to have high expectations of the subject.

Huang, Taddese, and Walter<sup>9</sup> studied students' pursuit of college science and engineering degrees to identify factors affecting minority and female students. They found that persistence to degree was more likely in students with greater academic strength in mathematics/science, greater personal interest in STEM, higher parental expectations, and a more supportive family. Racial/ethnic and gender gaps in STEM degree acquisition shrank when these factors were controlled. Interestingly, although smaller numbers of female students pursued a STEM degree compared to their male cohort, they tended to perform well after choosing that course of study.

Nicholls, Wolfe, Besterfield-Sacre, Shuman, and Larпкиattaworn<sup>10</sup> examined CIRP data to identify variables which exhibited significant differences between college freshmen with STEM and Non-STEM majors. Qualitative variables found to have significant differences included

students' self-ratings of academic ability, computer skills, social interests, and future goals. Significant quantitative variables included high school grade point average and mathematics SAT scores. The verbal SAT score was found to be significant but this was less consistent across the students.

## Data Collection

The means of obtaining data for this analysis was one of the key initial decisions. Designing and conducting a study to collect the volume of data required was judged to be impractical. Instead, existing data sources that could be adapted for the purposes of this analysis were sought. Extensive research has been performed about students' educational experiences in high school and college so several sources were considered. Among them were a series of longitudinal studies conducted by the Department of Education's National Center for Education Statistics (NCES) collecting information about students as they progress through secondary education and college. The most recent of these studies to be completed is the National Education Longitudinal Study of 1988 (NELS:88)<sup>11</sup> which collected extensive demographic, experiential, attitudinal, educational, and career data from students at set points in their lives.

The NELS:88 sample was chosen to be a representative cohort of students nationally. The study began in 1988 with students in the 8<sup>th</sup> grade and continued through four follow up waves of data collection in 1990 (10<sup>th</sup> grade), 1992 (12<sup>th</sup> grade), 1994, and 2000. Academic performance was assessed through cognitive tests administered by NCES, high school grades, standardized test scores, high school transcripts, and college transcripts if available. The students' parents, teachers, and school administrators were asked to answer survey questions about themselves, the students, and the schools. The dataset's breadth was impressive with over 7,000 variables from just the students' high school years. Special focus was placed by the NELS:88 study designers on retaining sample members who were minority students or who had dropped out of school at some point. The final follow up sample included 12,144 students of which 11,328 had participated during all five waves of data collection. The size of the NELS dataset, its lengthy time span, and the volume of prior educational research based upon it were factors in favor of utilizing it for this research.

Most of the NELS:88 data was categorical with many variables possessing an ordinal scale while other were strictly nominal. Potential answers to the NELS survey questions were often numerical and represented either single values or ranges of values. Some of the questions had only two possible answers. Very few variables were continuous in scale. The categorical qualities of the data led to a decision to use a nonlinear statistical technique in order to analyze it. Logistic regression<sup>12</sup> was chosen because of its ability to handle ordinal, nominal and continuous variables.

A process of recoding the variables was undertaken to resolve problems in variables with disjoint values that disrupted the modeling. For example, the variable measuring a student's overall reading proficiency quartile from the cognitive test in the 1988 base year (BY) data collection, "BY2XRQ," had seven potential values. These values were members of the set [1, 2, 3, 4, 6, 8, 9] and represented "Quartile 1 Low," "Quartile 2," "Quartile 3," "Quartile 4 High," "Legitimate Skip/Not in wave," "Missing", and "Test Not Completed," respectively. There were 1,180

records out of the original 12,144 that had a value other than 1 through 4 for the BY2XRQ variable. Since the potential values were not purely ordinal the variable's utility for model-fitting was hampered.

Other categorical variables had dichotomous responses but were not truly binary since the potential values were [1, 2] as opposed to [0, 1]. For example, the 2000 fourth follow up (F4) variable for students' sex, "F4SEX," was originally coded as 1 for male and 2 for female. Still other variables possessed a purely nominal set of potential values. The variable "F4Race2" had potential values within the set [1, 2, 3, 4, 5, -9] representing American Indian/Alaska Native, Asian/Pacific Islander, Black – not Hispanic, White – not Hispanic, Hispanic/Latino, or Missing, respectively.

The variables measuring standardized test scores for the SAT and ACT were categorical but resembled a series of mostly ordinal integer values. For example, the 1992 second follow up (F2) variable "F2RACTE" provided a student's ACT English score and included integer values of 6 through 36, 98, and 99. The first range of values represented an actual point score on the English portion of the ACT test while "98" indicated "Missing Data," and "99" represented "Legitimate Skip/Not in wave."

The difficulties caused by the categorical nature of the data were addressed by reviewing each variable's potential values and creating companion recoded variables that conveyed the information in a strictly ordinal or binary fashion. This required a process of examining the potential values and recoding non-ordinal values such as those for "Missing," "Legitimate Skip," etc. to a value of "0" indicating that no useful information was provided by that variable for that individual record. Other responses were grouped to create stronger delineations between answers. An example of this was a recoded variable for the father's highest educational level that categorized the answer as either "College and above" or "No college degree."

Categorical variables with dichotomous potential values were recoded as needed to make them truly binary. Thus, a recoded version of the variable F4SEX variable, "F4SEXro," was created with 0 for male and 1 for female. In such a case, the reference or base case was set to 0. Similarly, the nominal values for the variable measuring students' race, "F4Race2" was recoded into a set of binary dummy variables in which the base case was Caucasian with a value of 0. For these dummy variables "F4RACE2AI," "F4RACE2As," "F4RACE2BI," and "F4RACE2Hi" the value was 1 if the student's race was American Indian/Pacific Islander, Asian, African-American, or Hispanic, respectively. Thus Caucasian students were represented by having each of these four dummy variables equal to 0.

The complexity of reviewing each potential variable to develop a recoded version was deemed too time-consuming and not practical for future applications. In order to apply the findings of this research to future school settings, the data collection would have to be limited to a quantity and scope that would not be onerous to busy educators. Thus a strategic decision was made to limit the set of potential variables to a more manageable size. The BY data from 8<sup>th</sup> grade was the earliest data collected about the students and represented the earliest point in the NELS study at which academic assessments could be made. Prior research findings in the literature were used to select a smaller set of variables to be tested. A set of 66 variables was selected. These

variables reflected aspects of students for which prior educational research had found significant differences existed between outcomes<sup>13</sup>. These variables included basic demographic measures of sex, race, socioeconomic status, and family structure. Performance variables indicating standardized test results, NELS cognitive testing measures, subject competency ratings, and average grades were also included. Several attitudinal/behavioral variables were also selected. These included measures of student and parental attitudes about education, individual subjects, degree aspirations, and student capabilities. Behavioral variables examined how students spent time on homework, social activities, television watching, etc. Once the recoding process was completed including creation of several dummy variables a set of 76 potential predictors was available for model development.

The final step in data preparation was to classify the students' educational outcomes to create a dependent variable for the logistic regression models. Students with degrees in the "hard" Sciences, Engineering, and Mathematics were categorized as having a "STEM" outcome. Students with majors that did not generally involve extensive coursework in quantitative, technical subjects were classified as "Non-STEM." This category included the Fine Arts, English, and Other Humanities. A third class was created for four year college degrees that involved extensive quantitative coursework and represented a potential "gray" area between STEM and Non-STEM. This included the Health professions (medicine, dentistry, veterinary, pharmacy, nursing, and clinical therapies), Agriculture, Forestry, Social Sciences, Psychology, Business (Accounting, Business Administration, Finance, Marketing, and Management), and technical fields such as Computer Programming. Students who achieved a less than four year college degree were classified as "Sub 4-Yr Degree," and those who did not earn a college degree were classified as "No Degree." Another category was created by grouping the four outcomes other than STEM together into "All Else" to predict STEM vs All Else.

## Analysis

Models were fitted using SAS® statistical analysis software with the set of 76 potential variables and the Stepwise selection method with an alpha ( $\alpha$ ) error level threshold of 0.05 to enter or leave. This variable selection method ensured that SAS constructed the model by first identifying the most valuable predictor with a chi-square  $p$ -value of at least  $\leq 0.05$ . Once this variable had been entered into the model SAS continued choosing potential variables in the same fashion with the provision that if a variable's entry caused a prior entrant's individual  $p$ -value to increase above 0.05 it was automatically removed from the model. SAS stopped after considering all the potential variables for inclusion or reaching a user-defined limitation on the number of potential cycles. Using the Stepwise selection method to test potential variables for inclusion in the model was far more efficient in considering a large set of variables. It would have been impractical to test such a large group of potential variables by constructing separate models with different fixed combinations of variables. The records used in model fitting were selected from the 11,328 students that had participated in each of the five waves of data collection.

Stepwise logistic regression was used to create models predicting a STEM outcome between two possibilities such as STEM vs. All Else, STEM vs. STEM-Related, STEM vs. Non-STEM, etc. The models were fitted with randomly selected samples of students that were constructed to

proportionally represent their numbers in the population with those outcomes. Thus, a model to predict between STEM and STEM-Related was fitted using records randomly drawn from the sets of STEM and STEM-Related students. The samples were stratified by the outcome of interest with 70% of the total records from each stratum randomly selected for the model fitting. For example, in modeling STEM vs. All Else the total number of student records was 11,328 of which 738 obtained a STEM degree. These records were stratified by STEM = 1 for STEM students and STEM = 0 for all other outcomes. Of the 11,328 records approximately 70% of the STEM (517 records) and All Else students (7,414 records) were randomly selected for the model fitting.

Subsequently, each of the logistic regression models were validated by taking the model developed and applying it to predicting the outcome for the remaining 30% of the records from each stratum. Thus in the STEM vs. All Else case, the records of the remaining 221 STEM students and 3,176 of the All Else students were used to validate the model created with the original 70% of the total records.

Prediction models were evaluated by how well they employed data to choose between potential results. Each model was validated by constructing a Receiver Operating Characteristics (ROC) curve<sup>14</sup> to assess its predictive ability. The curves visually depicted the impact on correct/incorrect predictions of STEM based on the probability of STEM estimated. ROC curves were a valuable tool in assessing the sensitivity of the discrimination between the two potential educational outcomes. These curves plotted the probability of a correct STEM prediction (sensitivity) vs. the probability of an incorrect STEM prediction (1 – specificity). The prediction accuracy is function of the sensitivity and its corresponding specificity. They are balanced with sensitivity increasing as specificity decreases.

Predicting a student's educational outcome as STEM vs. All Else was done using a threshold value for the probability of a STEM outcome. The estimated probability of STEM lies in the range [0, 1] and the threshold value is the dividing line or "cutpoint" within that range. The model predicted the educational outcome depending on whether the estimated probability of STEM was greater than or equal to the cutpoint value. The potential cutpoints considered ran from 0.01 to 0.99. If the model estimated a student's probability of earning a STEM degree at 0.3 and the cutpoint was equal to 0.2, then the outcome was predicted to be STEM. Conversely, if a student's probability of earning a STEM degree was estimated by the model to be 0.1, then the model predicted the outcome was not STEM. The lower the cutpoint value, the more student records were classified by the model as STEM outcome predictions. This resulted in the model correctly predicting more of the actual STEM students to have a STEM outcome, but it also resulted in more actual All Else students being incorrectly predicted to have a STEM outcome. Low cutpoint values produced high sensitivity in the model but correspondingly low specificity (probability of correct All Else predictions). The higher the cutpoint value, the fewer true STEM students were correctly classified as having a STEM outcome, but the more true All Else students were correctly classified as All Else. High cutpoint values produced low sensitivity but high specificity. The selection of the threshold value governed the model's ability to accurately discriminate between the students' potential educational outcomes.

The ROC curves visually conveyed the predictive abilities of the models. Those models that had strong predictive ability produced ROC curves with a curve that rose sharply showing a high probability of correct STEM predictions and an associated low probability of incorrect predictions of All Else students as STEM. The area under the ROC curve (“AUC”) was the primary method of assessing a model’s predictive strength. The AUC value lay within [0, 1] with a high AUC value suggesting greater predictive accuracy. Hosmer and Lemeshow<sup>15</sup> rated ROC curves with AUC values of  $0.7 \leq \text{AUC} < 0.8$  as having “acceptable” discrimination ability;  $0.8 \leq \text{AUC} < 0.9$  having “excellent” discrimination ability; and  $\text{AUC} \geq 0.9$  having “outstanding” discrimination ability.

## Findings

The logistic regression model for STEM vs. All Else found numerous variables to be statistically significant predictive factors. The AUC value for the associated ROC curve was 0.848 demonstrating the fitted model possessed excellent ability to discriminate between these two educational outcomes. The significant variables were overall BY math proficiency; BY science quartile; family composition; language minority status; frequency of parental discussions with the student regarding post high school plans; parental expectations of the student’s advancement; student expectation of personal educational attainment; parental marital status; type of high school the student planned to attend; the father’s highest level of education; the number of hours per week the student worked for pay; the student’s ability groups for math and science; the student’s math and science grades from grades 6 to 8; ACT math score; SAT math and verbal scores; Asian race, African-American race; and gender. Figure 1 below shows that a high proportion of correct STEM predictions were achieved with a comparatively low proportion of incorrect STEM predictions for the STEM vs. All Else model.

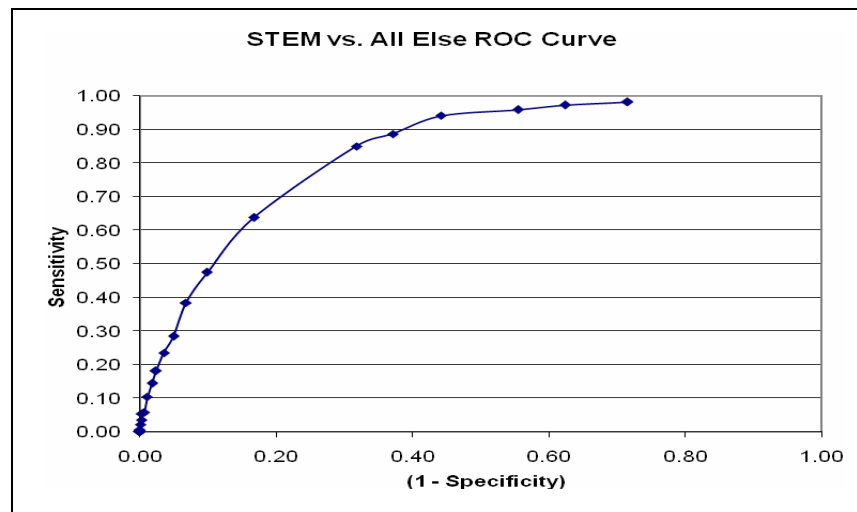


Figure 1 Sensitivity vs. (1-Specificity) for STEM vs. All Else model

The logistic regression model for STEM vs. STEM-Related also found numerous variables to be statistically significant predictive factors. The AUC value for the associated ROC curve was 0.720 showing the fitted model possessed acceptable ability to discriminate between these two educational outcomes. The significant variables were overall BY math proficiency; BY science and reading quartiles; family rules regarding the student’s television-watching habits; the number



of cigarettes the student smoked per day; the student’s ability groups for math and science; the student’s math and science grades from grades 6 to 8; ACT math score; SAT math score; the school’s percentage of white non-Hispanic 8<sup>th</sup> graders; the school’s base salary for beginning teachers with a B.A.; Asian race, and gender. Figure 2 below shows that a good proportion of correct STEM predictions were achieved with a correspondingly modest proportion of incorrect STEM predictions for the STEM vs. STEM-Related model.

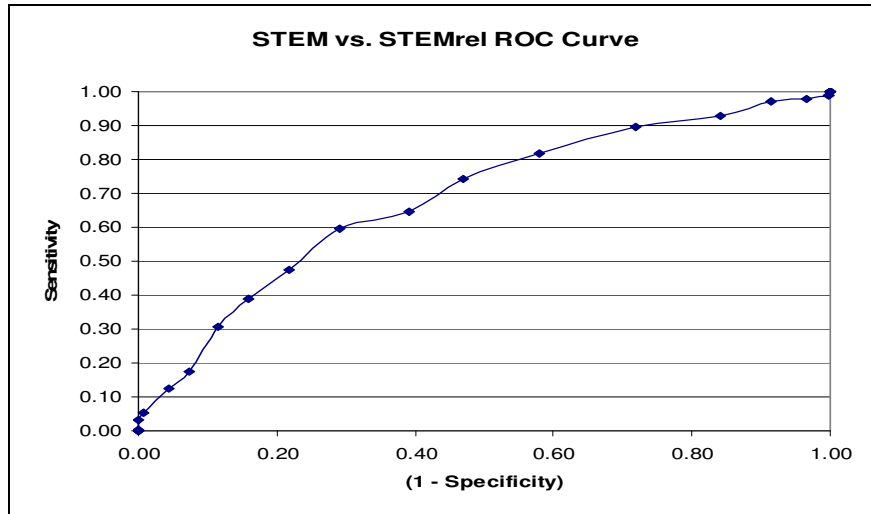


Figure 2 Sensitivity vs. (1-Specificity) for STEM vs. STEM-Related model

The logistic regression model for STEM vs. Non-STEM found several variables to be statistically significant predictive factors. The AUC value for the associated ROC curve was 0.743 showing the fitted model possessed acceptable ability to discriminate between these two educational outcomes. The significant variables were BY science quartile; family composition; the number of hours per week the student spent on homework; frequency of parental discussions with the student regarding post high school plans; parental expectations of the student’s advancement; parental marital status; type of high school the student planned to attend; the student’s math and science grades from grades 6 to 8; ACT math and English scores; SAT math and verbal scores; the school’s base salary for beginning teachers with a B.A.; Asian race; African-American race; and gender. Figure 3 below shows that a good proportion of correct STEM predictions were achieved with a correspondingly modest proportion of incorrect STEM predictions for the STEM vs. STEM-Related model.

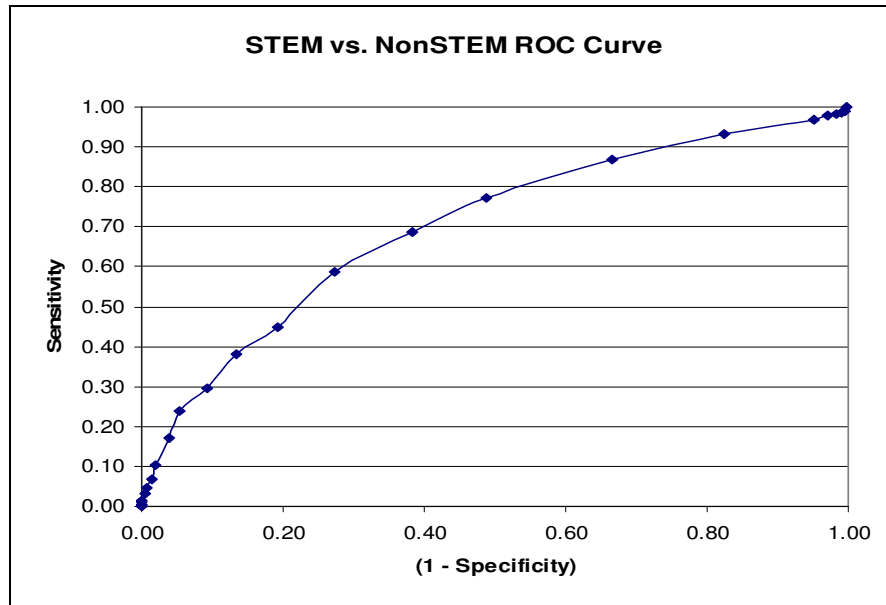


Figure 3 Sensitivity vs. (1-Specificity) for STEM vs. Non-STEM model

The logistic regression model for STEM vs. Sub-4Yr Degree found numerous variables to be statistically significant predictive factors. The AUC value for the associated ROC curve was 0.924 showing the fitted model possessed outstanding ability to discriminate between these two educational outcomes. The significant variables were BY math quartile; the number of hours per week the student spent on homework; limited English proficiency status; having a family rule about the student maintaining grade point average; frequency of parental assistance with homework; parental expectations of the student’s advancement; student expectation of personal educational attainment; type of high school the student planned to attend; the father’s highest level of education; the number of cigarettes the student smoked per day; the student’s confidence of advancing beyond high school; the number of hours per week the student worked for pay; the student’s ability group for math; the student’s math and science grades from grades 6 to 8; ACT math score; SAT math score; Asian race, African-American race; and gender. Figure 4 below shows that a good proportion of correct STEM predictions were achieved with a correspondingly low proportion of incorrect STEM predictions for the STEM vs. Sub-4Yr Degree model.

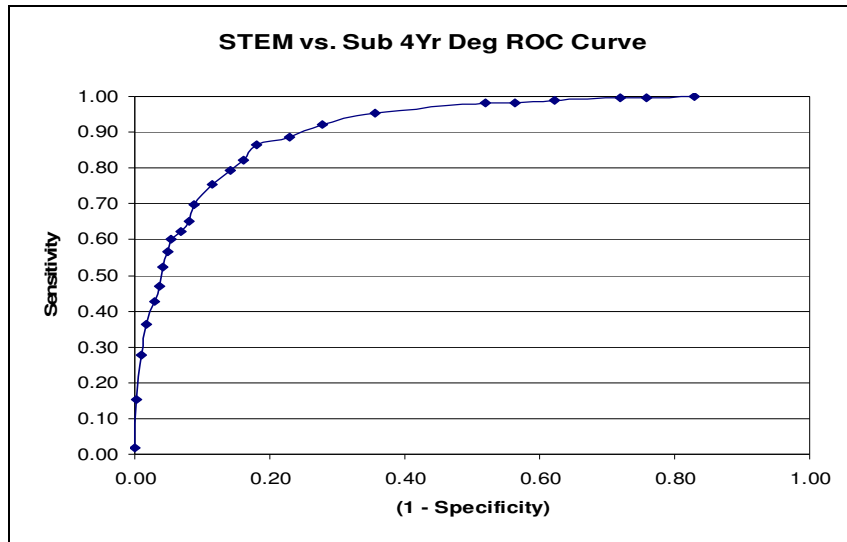


Figure 4 Sensitivity vs. (1-Specificity) for STEM vs. Sub-4Yr Deg model

The logistic regression model for STEM vs. No-Degree found numerous variables to be statistically significant predictive factors. The AUC value for the associated ROC curve was 0.919 showing the fitted model possessed outstanding ability to discriminate between these two educational outcomes. The significant variables were BY math and science quartiles; family composition; language minority status; having family rules about the student’s television watching habits; having a family rule about the student maintaining grade point average; the highest levels of education earned by the student’s father and mother; student expectation of personal educational attainment; the student’s confidence of advancing beyond high school; the number of hours per week the student worked for pay; the student’s math, science, and English grades from grades 6 to 8; ACT math and English scores; SAT math and verbal scores; the school’s number of students in remedial reading; and Asian race. Figure 5 below shows that a good proportion of correct STEM predictions were achieved with a correspondingly low proportion of incorrect STEM predictions for the STEM vs. No-Degree model.

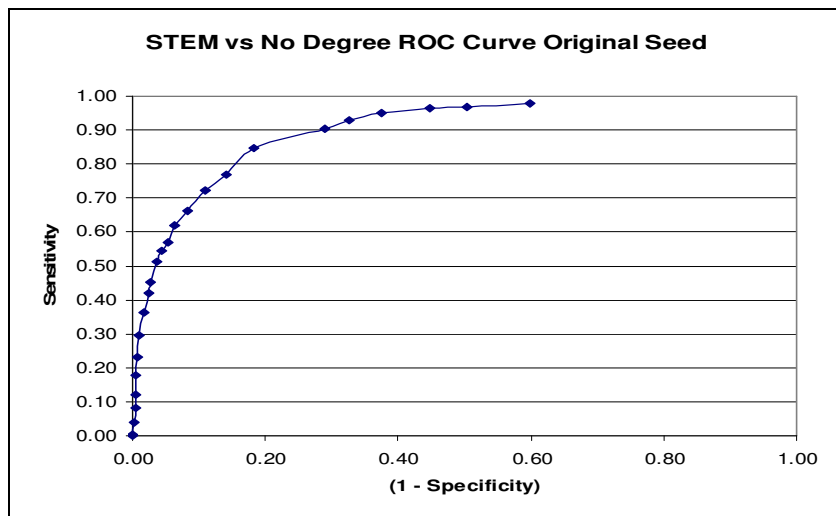


Figure 5 Sensitivity vs. (1-Specificity) for STEM vs. No-Degree model

## Conclusions

The findings corroborated the prior research in that academic ability, previous academic performance, personal attitudes, race/ethnicity, and gender are predictors of STEM degree achievement. The results also demonstrated that logistic regression models can achieve excellent predictive accuracy between two educational outcomes. The predictive accuracy improved when the two educational outcomes were less similar. For example, the models for STEM vs. All Else, Sub-4Yr Degree, or No-Degree had much greater predictive ability than those for STEM vs. STEM-Related or Non-STEM. The ROC curves for the different models indicated the tradeoffs between correct and incorrect STEM predictions. They illustrate the percentage of incorrect STEM predictions that would be incurred to achieve a corresponding level of 80% of the correct STEM predictions.

If the goal of an intervention program is to balance the sensitivity and specificity then both are plotted against a range of prediction cutpoint values and the intersection point is the ideal cutpoint value. If the goal is to optimize correct predictions of STEM, then the cutpoint can be selected accordingly with little concern for the specificity. More likely, there will be cost constraints that require education policy makers to balance the potential increase in STEM students resulting from a potential pro-STEM intervention with the cost per pupil. If this is the case, the cutpoint can be chosen to offer an intervention to all students that are suggested to be moderate STEM degree candidates without exceeding the budgetary limit. This would mean potentially leaving out the students who are already highly capable and likely to consider a STEM degree in favor of offering the intervention to more students that might not otherwise achieve a STEM degree.

The type of intervention program(s) envisioned would focus on students that could succeed in obtaining a STEM degree but might not already possess an interest in the topic as well those students who require extra assistance in order to succeed in studying STEM. A program to increase the students' awareness of STEM and their personal motivation to pursue it could assist in increasing the number of students that choose a STEM major vs. a STEM-Related or Non-STEM major. A program to assist the students at risk of not succeeding in STEM to strengthen their academic skills could increase the number of students that are capable of attempting and completing a STEM degree.

Adapting these findings for future research would involve gathering comparable measures of academic ability, academic performance to date, attitudes towards school and career, English proficiency, school/social/professional time commitments, family structure, parental involvement in the student's education, parental education levels, demographic data, and school characteristics. Some of these variables such as academic performance and ability are capable of being influenced with educational intervention programs. Other variables such as highest level of parental educational attainment, family composition, and parental marital status are beyond the scope of educators to affect. These variables as well as race/ethnicity and gender can be used by educators in identifying students "at risk" of not pursuing a STEM degree for special encouragement efforts. Some variables cannot be directly affected by educators but could be indirectly affected by encouraging parents to consider adopting new habits. For example, school administrators could attempt to impress upon parents the benefits of supporting students'

educational attainment; having family rules about academic performance; and encouraging students to focus on school.

Future work is anticipated in understanding the effect of the individual variables in predicting a STEM degree. Quantifying the benefit of shifting a student's academic picture through changing the values of individual variables by one unit will allow educators to determine the resulting value in intervention efforts. The most valuable variables for developing intervention programs will be those that are directly controllable and have the greatest impact on increasing the estimated probability of a STEM outcome.

## Bibliography

[1] National Science Foundation, Division of Science Resources Statistics, *Graduate Students and Postdoctorates in Science and Engineering: Fall 2002*, NSF 05-310, Project Officers: Julia D. Oliver and Emilda B. Rivers (Arlington, VA 2004). (available from NSF website <http://www.nsf.gov/statistics/nsf04318/>)

[2] Commission on Professionals in Science and Technology (CPST), data derived from the American Association of Engineering Societies' (AAES) Engineering Workforce Commission report, *Engineering and Technology Enrollments, Fall 1990 through 2004*.

[3] National Science Foundation, Division of Science Resources Statistics, *Science and Engineering Degrees, by Race/Ethnicity of Recipients: 1992-2001*, NSF 04-318, Project Officers: Susan T. Hill and Jean M. Johnson (Arlington, VA 2004). (available from NSF website <http://www.nsf.gov/statistics/nsf05310/>)

[4] Astin, Alexander W., *What Matters in College?: Four Critical Years Revisited*, Jossey-Bass Publishers, San Francisco, 1993.

[5] Astin, Alexander W., "Studying How College Affects Students: A Personal History of the CIRP," *About Campus*, July-August 2003, pp. 21-28.

[6] Zhang, Guili, Timothy J. Anderson, Matthew W. Ohland, and Brian R. Thorndyke, "Identifying Factors Influencing Engineering Student Graduation: A Longitudinal and Cross-Institutional Study," *Journal of Engineering Education*, Vol. 93, No. 4, October 2004, pp. 313-320.

[7] Smyth, Frederick L. and John J. McArdle, "Ethnic and Gender Differences in Science Graduation at Selective Colleges with Implications for Admission Policy and College Choice," *Research in Higher Education*, Vol. 45, No. 4, June 2004, pp. 353-381.

[8] Besterfield-Sacre, Mary, Cynthia J. Atman, and Larry J. Shuman, "Characteristics of Freshman Engineering Students: Models for Determining Student Attrition in Engineering," *Journal of Engineering Education*, Vol. 86, No. 2, April 1997, pp. 139-149.

[9] U.S. Department of Education. National Center for Educational Statistics. *Entry and Persistence of Women and Minorities in College Science and Engineering Education*. NCES 2000-601, by Gary Huang, Nebiyu Taddese, and Elizabeth Walter. Project Officer, Samuel S. Peng. Washington, DC: 2000.

[10] Nicholls, Gillian M., Harvey Wolfe, Mary Besterfield-Sacre, Larry J. Shuman, and Siripen Larpiattaworn, (2005), "A method for identifying variables for STEM intervention," *Journal of Engineering Education*, Vol. 96, No. 1, pp. 33-44, January 2007.

[11] National Center for Education Statistics, National Education Longitudinal Study of 1988, Project Officers: Peggy Quinn and Jeffrey T. Owings (Washington, DC), <http://nces.ed.gov/surveys/nels88/index.asp>.

[12] Hosmer, David W. and Stanley Lemeshow, "Applied Logistic Regression 2<sup>nd</sup> edition," John Wiley & Sons, Inc., New York, NY, 2000

[13] See reference 10.

[14] Fawcett, Tom, "An Introduction to ROC Analysis," Elsevier B.V, 2005, <http://www.csee.usf.edu/~candamo/site/papers/ROCintro.pdf>

[15] See reference 12.