120th ASEE Annual Conference & Exposition
FRANKLY, WE DO GIVE A D*MN
June 23-26, 2013

Paper ID #7742

Formulating Predictive Models of Engineering Student Throughput

Dr. Gillian M. Nicholls, University of Alabama in Huntsville

Dr. Gillian M. Nicholls is an Assistant Professor of Industrial and Systems Engineering and Engineering Management, and a 2009-2010 Gray Faculty Fellow at the University of Alabama in Huntsville. Her research interests are in applying statistical analysis and optimization to supply chain management, transportation management, and engineering education. She holds the B.S. in Industrial Engineering (Lehigh University), Masters in Business Administration (Penn State University), M.S. in Industrial Engineering (University of Pittsburgh.), and Ph.D. in Industrial Engineering (University of Pittsburgh). Address: N149 Technology Hall, The University of Alabama in Huntsville, Huntsville, AL 35899; telephone (256) 824-6637; fax: (256) 824-6733; e-mail: gillian.nicholls@uah.edu.

Formulating Predictive Models of Engineering Student Throughput

Abstract

Engineering degree acquisition is a complex system that lacks tools for efficient management and goal optimization. A reliable model of engineering degree acquisition will help administrators to increase throughput and resource utilization. It will also aid engineering students in better managing their educational investment. A method is needed to quantitatively assess the factors that predict time to graduation for engineering students; explore the potential positive effects of intervention to affect critical factors; and examine the costs vs. benefits of increasing engineering student throughput rates. Changes in student course-taking patterns and degree requirements have lengthened the time to graduation for typical engineering students. This reduces the number of students educated in a four year time period and consumes additional resources in course enrollments, faculty time, and support staff labor. With tuition costs rising faster than inflation, the trends have undesirable results for both universities and students.

This paper reviews the relevant literature in order to begin developing a study design to model student progression through engineering degree acquisition as a complex system. Elements are expected to include transition probabilities, identifying critical factors, predicting time to graduation, estimating costs and benefits of potential interventions, and projected throughput of engineers earning bachelors' degrees. The main goal of the research is to achieve an actionable, applicable, and accurate decision modeling method for a student's progress to an engineering degree and a university's resulting throughput rate to provide decision-making tools for both students and administrators. The longer term goal is increasing STEM persistence and throughput.

Introduction

Garrett and Poock[1] and Hansen et al.[2] have observed that universities are under increasing pressure to educate students more effectively in terms of cost and time. This is particularly true for Colleges of Engineering which have seen the time to graduate with a bachelor's degree increase[3] from four years to five or more as the volume of content deemed required has increased and more students matriculate with insufficient academic preparation to progress quickly[4]. Wang[5, 6, 7] and Baum and Ma[8] have documented the increasing costs of college tuition. Weisbrod and Asch[9] have shown how economic difficulties have led state governments to press for a lowered rate of increase in funding for state-supported universities. Private donors who have often generously supported colleges and universities are also feeling financial strain as they contemplate their plans for charitable donations[9]. On the other side, Dunbar et al.[10] and Logue[11] found students are under increasing pressure to complete their degrees, find gainful employment in a very tough economic climate, and begin repaying student loans that are approaching a crushing level of burden. Both sides have a vested interest in increasing the efficiency and effectiveness of undergraduate engineering education.

Bell[12] discussed the similarities that universities have to a manufacturing production system that can be modeled to predict and improve throughput Students enter as part of the "raw materials";

are routed through various processes (courses); are inspected for quality and functional ability (homework, exams, course grades); and ideally emerge as finished products (graduates) possessing knowledge and skills that make them marketable to buyers (employers). Engineering undergraduates take a number of courses that are key gateways to continued progress through the system. These include Calculus (A-C), Differential Equations, Physics (I & II), Linear Algebra, and Electrical Circuits. Some engineering programs also require Statics, Engineering Economy, and Thermodynamics I. All of these are courses that students may have difficult completing successfully on the first attempt thus presenting a probability of repeating the courses (rework). A number of the early math and science classes may be taken at less expensive community colleges before the student transfers to a four year school[13] such as The University of Alabama in Huntsville. Other students may transfer in from different schools after a physical relocation. Some students join the engineering path after pursuing other endeavors including work, family life, or public service.

Prior research by researchers including Astin[14]; Zhang, Anderson, Ohland, and Thorndyke[15], Zwick and Sklar[16], and Nicholls, Wolfe, Bestfield-Sacre, and Shuman[17] has shown that factors such as demographic characteristics, high school GPA, SAT/ACT scores, and grades earned in high school math and science classes have been found to be predictive of subsequent academic performance in college classes. Success in early classes in college may be expected to influence success in subsequent classes. Knight, Nicholls, and Componation[18] found other factors including credit load, part or full time employment, family commitments, extracurricular activities, transfer status, non-traditional student status, commuter status, etc. have also been shown to predict student performance. Orr, Ramirez, Ohland, and Lundy-Wagner[19] documented that some of these variables such as gender, race/ethnicity, and socioeconomic status may be statistically significant but are informative rather than actionable. Other variables are more actionable and capable of being affected. Min, Zhang, Long, Anderson, and Ohland[20] identified these actionable variables to include math skills, verbal skills, credit load, and course enrollment patterns.

Henschen[21] discussed efforts by The University of Kentucky to build a predictive model for student retention using variables including high school GPA, college entrance exams, and Learning Management System (LMS) utilization patterns. Henschen also reported on Taylor University's use of logistic regression to predict student retention from high school GPA, college entrance exam scores, freshman students' fall term GPA, fall term credit hour completion percentage, number of courses with D or F grades as of Fall midterm, and credit hours attempted in the spring term. The predictive results showing at-risk students are used to make intervention attempts. Raimondo[22] described analysis at the University of Michigan to assess within class performance by students and offer guidance via a digital resource called "E^2Coach"s to assist them in improving their performance trajectory. McKay[23] has used E^2Coach to interact with physics students predicted to be at risk of not succeeding and provide tailored feedback to all enrolled students that they can use to adjust their strategy in the course.

Universities have constrained resources including enrollment capacity, faculty, staff, lab space, etc. When students enroll in a class and subsequently withdraw or need to repeat it to pass successfully, both the university and the student have lost resources. The student's progression towards a degree is delayed; the risk of dropping out increases, and the cost of the course is lost.

The university earned the student's tuition money but the overall student throughput is hampered. When students take longer to graduate, spaces that could have been opened up for new students are not available.

Sandmeyer, Dooris, and Barlock[24] stated that universities need tools to plan their resource allocation and educational policies. These are issues that affect engineering student throughput, cost efficiency, and allow for investment in promising growth areas. Having accurate projections of the cost and time resources that will be required to graduate an incoming cohort of engineering students and what the costs and benefits will be of making different allocations is key for university decision makers. Students will benefit from tools that realistically estimate the time it will take them to graduate and the resulting cost to complete their education. Knowledge is empowering. Armed with a more realistic understanding of the financial investment being made, students can make better decisions about how to allocate their time, effort, and financial resources to completing their education. Rothstein and Rouse[25] examined the financial constraints students face after taking graduating with large debts have serious, long term effects on career choices and future financial decisions. The rising tide of student debt could have very adverse consequences for the U.S. economy[26] . Hard decisions and choices need to be made by all stakeholders in the educational process in order to improve the process overall.

Research Goals

The purpose of this research is to develop a set of predictive models that utilize student characteristics of demography, academic performance, and course-taking patterns to determine their probability of completing an engineering degree within four to six years and project this to a university's throughput rate of students graduating with a bachelor's degree in engineering. A subset of the models will focus upon the progression of students that are members of under-represented groups including females, African-Americans, Asian, Hispanic/Latino, and non-traditional students such as home-schooled, transfer, part-time, and those returning to school after another career. Based on data describing the incoming cohort of students and the university's engineering programs, administrators will have a quantitative tool to estimate the numbers of engineering students progressing through the program and consequently what allocation of resources will be required to serve them. The data used in the predictive modeling will be drawn from university admittance, enrollment, and academic performance records to increase the likelihood of successful dissemination and adaptation at other educational institutions. Analysis of the costs and benefits of interventions designed to aid engineering students struggling with coursework or considering switching majors will aid policy-makers in assessing the fiscal impact of attempting to improve throughput with pro-engineering interventions.

There are two overall goals of the proposed model in this research. The first goal is to examine the process of engineering students' course-taking and determine how significant factors predictive of successfully earning a bachelor's degree within four to six years can be used to model a university's throughput of engineering students. This goal will be met by collecting data about actual students and model their educational outcomes within the larger system of the engineering program.

The second goal is to share the results and methodology of creating these predictive models with engineering educators and university administrators for adaptation and adoption at other institutions. The methodology will thus need to include reflections of which aspects are most sensitive to differences in institutions or their academic policies. This goal will be met by sharing the results through scholarly publications and demonstrations at educational conferences. Ultimately, a tool adopted for university planning would produce information that pre-college students and high school guidance counselors could use in projecting the cost of completing a degree at a given university based on the typical length of time to graduate with a bachelor's in engineering.

Background

The source of data for this study will be the records of undergraduate engineering students at The University of Alabama in Huntsville (UAHuntsville). The diversity of the undergraduate engineering student population between 2005 and 2012 was examined to ascertain whether this source of data would provide a sufficient numbers and heterogeneity of population. The percentage of UAHuntsville engineering students that are from underrepresented racial/ethnic populations is approximately 17%. These students are primarily Black/African-American, Asian/Pacific Islander, and Hispanic/Latino in descending order. The percentage of female students in engineering is approximately 20%. These percentages for engineering students earning bachelors' degrees are approximately 14% and 20%, respectively. Since 2007-2008 a total of 166 undergraduate engineering students from underrepresented racial/ethnic populations earned a bachelors' degree. Overall, the entire undergraduate student population in the Fall of 2011 of nearly 5,600 students was approximately 14.5% Black/African-American, 3.4% Asian, 2.8% Hispanic/Latino, 1.4% American Indian/Alaskan Native, and 1.6% biracial/multiracial. Preliminary analysis indicates that approximately 42% of UAHuntsville engineering students persist to graduate within six years, and that the percentage of engineering students that transferred in from other schools is nearly 50%. Over 26% of the undergraduate engineering students enrolled in Spring 2012 were taking class on a part-time basis.

Proposed Research Design

Research Questions

- Can models of engineering student progression through core gateway classes accurately predict time to graduation and probability of graduating in 4 – 6 years?
- Can such models focusing upon students that are members of under-represented groups accurately predict time to graduation and probability of graduating in 4 – 6 years?
- What are the costs and benefits of potential pro-engineering educational interventions that affect the values of actionable predictors to improve student's probability of graduating in 4 – 6 years?
- Can the progression of engineering students through a university program be accurately modeled to predict student throughput?
- Can factors be identified that allow universities to adjust throughput and evaluate the required resource allocations?

Data Collection

The University of Alabama in Huntsville meticulously collects student data including prior academic records, demographic information, declared major(s), academic performance metrics, and course enrollment records. This data is maintained and available via its student information database. These records will be obtained for the engineering students that have matriculated at UAHuntsville since 2005 with ongoing data collection through 2015. This will provide a large dataset of engineering students as they began their studies, continued through later coursework, and ultimately either graduated, switched majors, transferred, dropped out, or remained in school. Since 2005, a total of 1,825 students have graduated UAHuntsville with a bachelor's of science in engineering for an average of 228 students per year. From 2005 through June, 2012 the total undergraduate engineering enrollment has risen from 1,400 to nearly 1,800 students. The number of engineering students earning bachelors' degrees from 2005 through 2015 is estimated at nearly 3,000 students.

The data collection will start by studying the course enrollment and success rates for a subset of gateway courses at UAHuntsville to measure the probabilities of successful completion (earning an A, B, or C), unsuccessful completion (earning a D or F), and withdrawal for students given a their individual sets of characteristics and factors. The gateway class sizes at UAHuntsville are sizeable enough to provide an extensive set of records over the anticipated 10 year period. For example, during the 2011-2012 academic year student enrollment figures for the Calculus A – C sequence were 608, 486, and 483 students, respectively. Similarly, the total 2011-2012 enrollment for the Physics 1 and 2 courses were 422 and 371, respectively. The percentage of students within a category of background variables such as gender, minority/majority, transfer status, or math SAT/ACT scores that completed the course successfully vs. unsuccessfully on the first attempt will be used to gauge the transition state probabilities. A general set of conditional probabilities will be estimated for the probability of successful completion on repeated attempts for the gateway courses.

Additional variables to be gleaned from the enrollment records include but are not limited to the groups of classes taken together, credit load per semester, semester vs. overall grade point average, and declared major. The major per semester is expected to provide time interval data to be used for determining whether and when a student that switched out of engineering left the engineering pathway.

Data Analysis

This data will be used to model engineering student progression through the different majors and predict the probability of successful graduation within four, five, or six years for a student with a given vector of variables including prior high school performance, prior college performance, patterns of enrollment, and demographic characteristics. Actionable factors found to be significantly predictive will be useful in guiding any examination of intervention efforts designed to shift the characteristics of a given student to a band predictive of higher success.

University administrative decision makers need a "bigger picture" than the probability of an individual student graduating in four vs. six years. The progression of students through an undergraduate degree program is a complicated and lengthy process requiring considerable resource investments by the university. The process is not purely linear since there is not a single course routing plan across all of the engineering disciplines, and even within disciplines, students do not follow identical paths through coursework. Students take courses in different assortments and may choose to take less than the anticipated number of courses per semester for a full time student. When a student withdraws from a course or completes it unsuccessfully, this creates "loops" in the process of earning a degree which can notably disrupt the normal course-taking pattern for that student's engineering discipline. Ohland, et al.[27] found that the number of students switching into engineering from a different major is not generally high, but there is some switching between engineering disciplines that can further disrupt the course-taking pattern.

Modeling the process of educating engineering students requires a means of capturing the feedback inherent in the education process where students repeat courses, are affected by choices in which classes to simultaneously enroll in, and students enter the program at different time points from different sources. The method selected for this modeling is a combination of simulation of student outcomes and system dynamics modeling. Sterman[28]; Schaffer[29]; Thompson and Reimann[30]; and Williamson, Robertson, and Casey[31] used these tools to examine complex scenarios. Simulation will be applied to model the probable patterns of students through the different entrances and exits from the engineering pathway and to test the results of the predictive models for student outcome and time to graduation. System dynamics modeling will be applied to examine the behavior of the complex system of undergraduate engineering education over time as engineering students are "processed" towards becoming a cohort of graduating engineers.

Validation of Predictive Models

Collecting data over a 10 year period between 2005 – 2015 will permit the predictive models to be validated by randomly dividing the student records up into two sets. One set that will be used to fit the models, and a second set that will be held in reserve to test the accuracy of the models when applied to fresh data. Part of the planning phase of the research will involve determining how best to divide the data to avoid the effects of other factors. For example, the data could be divided by years with one set containing all students that matriculated between 2005 and 2008 and the second set containing all students that matriculated between a sequence of years later in the time scale. Another option is to randomly select a portion of all the students that matriculated between 2005 and 2012 to use for fitting and use the remaining data for testing.

Cautionary Observations

It should be noted that this research has the potential to be controversial, and there are some risks associated with it. The concept of modeling university throughput of engineering students based upon principles more commonly associated with a manufacturing setting may seem jarring and not respectful of students' individuality. Each student is a unique human with a myriad of personality traits, characteristics, and life experiences that affect how they experience the educational process. Most individuals involved in engineering education over some time have observed students appearing to have great potential to be successful in an engineering degree program that nonetheless either struggled or became disengaged with the end result of departing the STEM track. Some leave due to difficulties in progressing or simply because they chose another path that held greater appeal. Conversely, most educators have also observed students appearing to be more challenged in completing an engineering degree that ultimately overcame the perceived obstacles and doggedly persevered to achieve their desired degree(s). These students represent “outliers” that are difficult for any model to account for in generating predictive results. There are individual human characteristics that affect students’ attitudes, decisions, and behaviors. Campus climate may also be a factor for some students. These characteristics and factors are difficult to measure and control for. Given these facts, the proposed tool for university policymakers has the potential to be used in ways that have negative consequences for engineering education.

In a scenario where a university had more applicants to its engineering program than spaces available, there may be temptation to use the findings of this research to select a cohort that was predicted to have higher probabilities of graduating in four years and a correspondingly higher throughput. Such a policy could reduce the likelihood of acceptance for students that don’t have factors sufficiently predictive of success but who possess other unmeasured characteristics driving them to be successful in engineering. The goal of universities to offer a diverse student population would mitigate the likelihood of using the research in a way that selected a very homogenous, non-diverse cohort. However, there are students that enter engineering from non-traditional pathways and these students with the potential to be outlier successes may be overlooked when selecting students for acceptance. The best way to counteract this tendency is to remind users of the research that it is a tool for improving student outcomes and university throughput by aiding decisions about program design and resource allocation. Reminding users to consider how they utilize the tool will aid in lessening unconscious bias. Yet, there is still a risk that users may intentionally seek to use it for short-term advantages. This is less of a concern for schools that have excess enrollment capacity and are eager to recruit additional students from all walks of life to the engineering profession.

While this risk is real, the value of having a methodology for creating predictive models that are accurate for most students and that enable better evaluation of resource allocation is significant. The environmental factors creating financial pressures on universities and students are unlikely to recede soon. The potential positive applications of the proposed tool are also significant. Namely, that identifying students at risk of switching out of engineering or not succeeding presents an opportunity to provide cost effective pro-engineering support to improve retention. Identifying students that are good candidates to enter engineering will aid in recruitment.

Next Steps

The next phase of this research involves refining the research design, collecting the historical data at the University of Alabama in Huntsville, and estimating the transition probabilities given broad background variables. It is anticipated that valuable feedback about this proposed research will be obtained through the conference process.

Bibliography

1. Garrett, M.D. and M.C. Poock, “Resource Allocation and Management Strategies in Times of Fiscal Constraint and Their Impact on Students,” *College Student Journal*, Dec 2011, Vol. 45 Issue 4, p882-890.
2. Hansen, R.J., W.S. McClure, & S. Parkes, “Dangerous Opportunities: Strategic Implications of the Global Financial Crisis,” *Continuing Higher Education Review*, 73, 184-193.
3. National Center for Education Statistics (NCES), U.S. Department of Education, “Six-Year Attainment, Persistence, Transfer, Retention, and Withdrawal Rates of Students Who Began Postsecondary Education in 2003–04,” by Paul Skomsvold, Alexandria Walton Radford, and Lutz Berkner July 2011, NCES 2011-152, http://nces.ed.gov/pubs2011/2011152.pdf
4. Offenstein, J. C. Moore, & N. Shulock, “*Advancing by Degrees: A Framework for Increasing College Completion*,” Institute for Higher Education Leadership & Policy (IHELP), 2010, http://www.eric.ed.gov.elib.uah.edu/contentdelivery/servlet/ERICServlet?accno=ED511863
5. Wang, Penelope, 2008, “Is college still worth the price?,” CNNMoney.com Money Magazine, August 20, 2008, updated April 13, 2009. http://money.cnn.com/2008/08/20/pf/college/college_price.moneymag/index.htm
6. Wang, Penelope, 2008, “The real return on your college investment,” CNNMoney.com Money Magazine, August 20, 2008, updated April 13, 2009. http://money.cnn.com/2008/08/20/pf/college/college_price.moneymag/index3.htm
7. Wang, Penelope, 2008, “What’s really driving skyrocketing prices,” CNNMoney.com Money Magazine, August 20, 2008, updated April 13, 2009. http://money.cnn.com/2008/08/20/pf/college/college_price.moneymag/index2.htm
8. Baum, S. and J. Ma, “Trends in College Pricing, 2011 – Trends in Higher Education Series,” College Board Advocacy & Policy Center, 2011, http://trends.collegeboard.org/downloads/College_Pricing_2011.pdf
9. Weisbrod, B.A. & E.D. Asch, “The truth about the “Crisis” in Higher Education Finance,” Change, 42(1), 23-29.
10. Dunbar, A., D. Hossler, D. Shapiro, J. Chen, S. Martin, V. Torres, D. Zerquera, & M. Ziskin, “National Postsecondary Enrollment Trends Before, During, and After the Great Recession,” National Student Clearinghouse Research Center Signature Report, July 2011, http://research.studentclearinghouse.org/
11. Logue, A.C. “Tough Conversations – Talking about student debt and default across the campus,” *University Business*, October 2011, pg. 44-48.
12. Bell, E.C., “*A College of Business Administration as a Production System*,” The Academy of Management Journal, Vol. 17 No. 2, June, 1975, pgs. 306-317.
13. Sallie Mae, Inc., “How America Pays for College 2012,” https://www1.salliemae.com/NR/rdonlyres/75C6F178-9B25-48F5-8982-41F9B3F35BF6/0/HowAmericaPays2012.pdf
14. Astin, A. W., 1993, *What matters in college? Four critical years revisited*, San Francisco, CA; Jossey-Bass Publishers.
15. Zhang, G., T.J. Anderson, M. W. Ohland, & B.R. Thorndyke, 2004, “Identifying factors influencing engineering student graduation: a longitudinal and cross-institutional study,” *Journal of Engineering Education* 93(4), pgs. 313-320.
16. Zwick, R. & J. G. Sklar, 2005, “Predicting college grades and degree completion using high school grades and SAT scores: the role of student ethnicity and first language,” *American Educational Research Journal*, 42(3), pgs. 439-464.

17. Nicholls, Gillian, Wolfe, Harvey, Besterfield-Sacre, Mary, & Shuman, Larry, "Predicting STEM Degree Outcomes Based on 8th Grade Data and Standard Test Scores." *Journal of Engineering Education*, July 2010, Vol. 99, No. 3, pgs 209-223.

18. Knight, A., Nicholls, G., & Componation, P. "Measuring the Effect of On-Line Homework Procedures on Student Exam Performance." Proceedings of the 2012 American Society of Engineering Education Conference, San Antonio, TX, June 2012.

19. Orr, M.K., N.M. Ramirez, M.W. Ohland, & V. Lundy-Wagner, "*Using High School and District Economic Variables to Predict Engineering Persistence*," Proceedings of the American Society for Engineering Education, San Antonio, TX, June 2012.

20. Min, Y., G. Zhang, R.A. Long, T.J. Anderson, & M.W. Ohland, "Nonparametric Survival Analysis of the Loss Rate of Undergraduate Engineering Students," *Journal of Engineering Education*, April 2011, Vol. 100, No. 2, pgs 349-373.

21. Henschen, D., "To Avoid Nasty Surprises, Higher Ed Turns To Prediction," *Information Week*, February 2013, pgs. 5-6, http://twimgs.com/infoweek/green/012813ed/InformationWeek_Education_2013_02.pdf .

22. Raimondo, J., "Analytics' Big Test: Can data analysis keep students on track and improve college retention rates?," *Information Week*, February 2013, pgs. 7-14, http://twimgs.com/infoweek/green/012813ed/InformationWeek_Education_2013_02.pdf .

23. McKay, T., "Leveraging analytics to study and improve learning," Next Generation Learning Challenges Blog, February, 14 2013, http://www.nextgenlearning.org/blog/leveraging-analytics-study-and-improve-learning .

24. Sandmeyer, L.E., M. Dooris, & R.W. Barlock, "*Integrated Planning for enrollment, facilities, budget, and staffing: Penn State University*," New Directions for Institutional Research, Fall 2004, Vol. 2004, Issue 123, Pg. 89-96.

25. Rothstein, J. & C.E. Rouse, "Constrained after college: Student loans and early-career occupational choices," *Journal of Public Economics*, 2011, Vol. 95, pgs. 149-163.

26. Consumer Reports, May 2012, "*Student debt: Your Threat – It can have an impact on the entire economy*."

27. Ohland, Matthew W., Sheri D. Sheppard, Gary Lichtenstein, Ozgur Eris, Debbie Chachra, & Richard A. Layton, "Persistence, Engagement, and Migration in Engineering Programs," *Journal of Engineering Education*, July, 2008, pgs. 259-278.

28. Sterman, J.D. (1994) Learning in and about complex systems. System Dynamics Review, 10 (2-3), 291-330. http://hdl.handle.net/1721.1/2504

29. Shaffer, S.C., "System Dynamics in Distance Education and a Call to Develop a Standard Model," *International Review of Research in Open and Distance Learning*, Vol 6, #3, Nov., 2005.

30. Thompson, K. & P. Reimann, "Patterns of use of an agent-based model and a systems dynamics model: The application of patterns of use and the impacts on learning outcomes," Computers & Education 54 (2010), pgs. 392-403.

31. Williamson, R.L, J.S. Robertson, & L.B. Casey, "Using a dynamic systems approach to investigating postsecondary education and employment outcomes for transitioning students with disabilities," *Journal of Vocational Rehabilitation*, 33 (2010), pgs. 101-111.