

# Transfer Student Success in Engineering

Meryl Motika, Xinyi Hou

Center for Educational Effectiveness

May 2018

## Executive Summary

Transfer students make up a large part of the undergraduate student population at UC Davis yet our university lacks a systematic study of factors impacting their success. As a part of the Avenue E project, the Center for Educational Effectiveness (CEE) analyzed predictors of transfer student success in the College of Engineering. We used three measures of success: the probability of graduating in two years or less, probability of thriving defined as having a 3.0 or higher grade point average (GPA) with a 13 units in their fourth quarter, and GPA at graduation.

This study has four parts. First, we estimated the influence of demographic background and community college of origin on the three success measures. Second, we examined the pre-UC Davis courses of transfer students from some high, middle, and low-performing colleges to see how they prepared for transfer. Third, we used machine learning models to identify which courses from UC Davis were the best predictors of transfer students' thriving and graduation rates. Finally, we estimated the correlations of those influential course grades with two-year graduation and thriving in regression models controlling for background characteristics.

### Key findings:

- 1) Under-represented minorities and, to a lesser extent, first-generation college students underperformed compared to other students after transfer. Students completing a Transfer Admissions Guarantee did better, as did those who studied more engineering and social science prior to transfer.
- 2) There was differential success by community college of origin, although it was not extreme for Avenue E colleges.
- 3) Differences in instruction seem unlikely to explain the differential performance by college. Nearly all students took all of the core preparatory courses in math and engineering before transfer. Many students took these classes at multiple colleges. This cross-registration seemed to be based on geography rather than college quality. It is plausible that the college success rates stem from either regional factors not included in the analysis or advising and other support services rather than instruction.
- 4) Several courses were influential in the machine learning models but were unimportant in the regression in which we controlled for demographic factors. These courses tended to emphasize philosophy, writing, and other non-mathematical skills and to have high average grades. These courses' correlations with success may be related to demographic factors or may depend on interactions with other courses.

## Transfer Student Success in Engineering

- 5) After validation, the course grades most closely correlated with earning a high fourth-quarter GPA at UC Davis included Fluid Mechanics, Electromagnetics I, Engineering Economics, Heat Transfer, and Electrical Circuits and Systems. The course grades most correlated with graduating in two years were for Electronic Circuits I, Thermodynamics, Digital Systems I, Mechatronic Systems, and Chemical Kinetics.

### Introduction and Background

As UC President Janet Napolitano emphasized in the creation of the UC Transfer Students Initiative, opening our doors to transfer students is a core part of how UC creates educational opportunities for all Californians (Transfer Students Initiative, 2017). Transfer students are a large and diverse population at UC Davis and may face a different set of challenges here compared to freshman admits. In the period from 2011-2016 they made up around 23% of enrolled students and 37% of new admits, although the proportion in engineering was lower at 24% of new admits. Transfer students are more diverse than the freshman admits; 46% of transfers in the College of Engineering were the first in their families to attend college and 19% were under-represented minorities, as compared to 33% and 17% of junior and senior level freshman admits. Furthermore, a UC degree seems to enable many transfer students to obtain higher economic status than their parents (Freeling, 2015). The transfer program is a huge potential source of social mobility.

In this report, we explored four aspects of the transfer student experience:

- 1). How does the background of transfer students influence their performance at UC Davis?
- 2). Do transfer students complete all required courses before transferring, and are there identifiable patterns in how they complete those courses that predict success?
- 3). Which courses at UC Davis are most influential on transfer student success?
- 4). Considering the background and the influential courses in a model, how do these courses relate to transfer students' performance at UC Davis?

The overall goal was to find points of intervention to improve transfer students' outcomes in the UC Davis engineering program.

In the first model, we explored the relationships between transfer student background, preparation, and college of origin and three outcomes: the probability of thriving after the first year, the probability of graduating in two years, and GPA at graduation. In this context, we defined "thriving" as achieving a 3.0 or higher GPA with at least 13 credits in the first quarter of the second year. We also examined the distributions of demographic variables and performance of the transfers from specific community colleges.

The variability in performance among students from different community colleges raised questions about whether there were significant differences in transfer student preparation.

## Transfer Student Success in Engineering

Specifically, we were interested in the consistency across community colleges in completion of the courses required to transfer into engineering at UC Davis. Our engineering program has selective major review meaning that potential transfer students are required to complete a specific set of pre-requisite courses with a minimum GPA before they can be admitted, but the issue is still worth studying since requirements and actual behavior do not always coincide. Descriptive analyses were used to answer these questions.

Our next question was whether we could identify specific courses at UC Davis that had a significant influence on the success of our transfer students in engineering, which we focused on in the third part of the work. Specifically, we aimed to find the top ten most influential courses for two-year graduation and to thrive after the first year. These specific measures of student performance were selected because they are binary and many students do not accomplish them. These characteristics allowed us to use machine learning predictive techniques that are well suited to identifying pivotal courses.

In the fourth part, we validated the results of the machine learning models by incorporating the course-based factors into a more general regression model of transfer student success. The explanatory variables for this model included the student's background, preparation, and grades in the pivotal courses at UC Davis from the machine learning model. The outcomes of interest were the probability of graduating in two years and the probability of thriving after the first year.

In the remainder of the report, we will describe data preparation, methods, and the results of these four parts of the analysis.

### Data Preparation

The data for this study come from Admissions and Registrar records at UC Davis and were provided to CEE for our ongoing work in support of undergraduate education. The data included course enrollment and grade records as well as demographic information and other admissions data including the college of origin for each student enrolled since 2001. For this project the sample was limited to undergraduate students who transferred into the College of Engineering at UC Davis and took courses between Fall 2008 and Fall 2015.

For models of 2-year graduation, the sample was restricted to students whose most recent major after two years was in engineering. Only students who had articulation records of courses taken prior to transferring at either a community college or UC were considered for the analysis. There were 1,223 students who met the aforementioned requirements.

For models of thriving after the first year, the sample was restricted to students whose second-year first quarter major was in engineering. For the articulation records, the requirement was the same as above. 1,240 of these students were included in the analysis.

## Transfer Student Success in Engineering

The samples described above were used for the first and fourth parts of the analysis. In the second part, we focused on specific community colleges chosen because they were part of Avenue E. In the third part, we expanded the samples to include students whose demographic information was missing but added a requirement that the student had records for the courses of interest at UC Davis, resulting in samples of 1,219 students for the graduation model and 1245 for the thriving model.

Table 1 shows the main sample's demographic and performance statistics and compares these transfer students to freshman admits in engineering. Transfer students in the College of Engineering were more likely to be male, first generation, and from an under-represented minority group than freshman admits who reached their junior or senior year. Transfer students in this college performed about the same as freshman admits regarding GPA.

	<i>Transfer students</i>		<i>Engineering college Junior or Senior level non-transfers</i>	
	Main sample	Graduates	All	Graduates
<i>Female</i>	12%	12%	29%	27%
<i>Under-represented minority</i>	18%	17%	17%	15%
<i>Minority- Black</i>	3%	2%	2%	1%
<i>Minority-Latino/a</i>	15%	15%	14%	13%
<i>First generation</i>	46%	46%	32%	31%
<i>Low-income</i>	23%	23%	29%	28%
<i>Graduated in 2 years</i>	51% of eligible	59%	In 4 years: 59% of eligible	In 4 years: 59% of graduates
<i>Most recent average GPA</i>	2.98	3.09	3.02	3.10
<i>N</i>	1,240	1,069	4,534	2,748

Table 1: Sample Characteristics

### Part one: Historical Model of Transfer Student Success in Engineering

We studied the influences of general background and college of origin on student success. This part was a data mining rather than a hypothesis-driven analysis. The results were estimated using mixed-effects models with the college of origin represented as random effects. As such, the estimated effects of the background were independent of the college of origin and vice versa.

### Background and Preparation Analysis

We investigated effects of a transfer student’s background on three different outcomes: thriving after the first year, the probability of graduating in two years, and GPA at graduation. Table 2 shows the results of this analysis. Coefficients for thriving and two-year graduation should be interpreted as an increase or decrease in the student’s probability of graduating on time or thriving compared to the baseline student. In the third model, the coefficients can be interpreted as the effect of membership in that group (for demographic variables) or 1 additional course in that subject (for courses taken prior to transfer) on the student’s final GPA.

	<i>Thriving</i>		<i>Two-year graduation rate</i>		<i>GPA at graduation</i>	
<i>Minority</i>	-0.083*	0.037	-0.125***	0.036	-0.095**	0.036
<i>Low-income</i>	0.025	0.035	-0.007	0.034	-0.049	0.033
<i>First generation</i>	-0.072*	0.030	0.033	0.029	-0.074*	0.029
<i>Female</i>	-0.051	0.044	-0.076	0.043	-0.053	0.042
<i>Transfer Admission Guarantee</i>	0.060*	0.029	0.144***	0.028	0.080**	0.027
<i>Engineering courses</i>	0.009	0.007	0.037***	0.006	0.019**	0.006
<i>Biology courses</i>	-0.002	0.010	0.000	0.010	0.011	0.009
<i>Physical science courses</i>	0.007	0.005	0.023***	0.005	0.001	0.005
<i>Math courses</i>	-0.006	0.015	0.003	0.014	-0.008	0.014
<i>Social science courses</i>	0.008	0.005	0.023***	0.005	0.014**	0.005
<i>English courses</i>	0.030	0.017	0.027	0.016	0.018	0.015
<i>Language, culture, and arts courses</i>	-0.003	0.005	-0.003	0.005	0.010*	0.005
<i>Other courses</i>	-0.001	0.009	0.012	0.008	-0.010	0.008
<i>N</i>	1,240		1,223		1,055	

\* p<0.05, \*\*p<0.01, \*\*\*p<0.001. Coefficients and standard errors from linear mixture model regressions including random effects for the college of origin. Baseline: non-minority average income non-first generation male with no TAG and average numbers of courses.

*Table 2: Regressions of student success measures on pre-transfer characteristics*

Underrepresented minorities had more difficulty than others with all three outcomes. First-generation students had lower thriving probability and GPA at graduation compared to students whose parents attended college. Students who completed TAG agreements were more likely to graduate in two years and earned a little bit higher GPAs than the other students.

Extra preparation in engineering helped students with the last two outcomes, and pre-transfer work in social sciences and humanities also seemed to increase final GPAs. Additional courses in physical sciences and social sciences helped students graduate on time.

## Transfer Student Success in Engineering

While these results are not shown, we investigated differential effects of background characteristics for minority students. We found no significant differences for minority students; the same types of courses and background characteristics that helped transfer students in general also seemed to help minority transfer students in particular.

### College of Origin Analysis

In addition to examining effects of transfer students' characteristics, we studied average success of students from different community colleges. We estimated these effects using a random effects model controlling for the factors in the previous analysis. Most had no clear correlation with success, but a few colleges tended to have students do unusually well or poorly here. Table 3 shows average outcomes and characteristics of students from the colleges identified as outliers through this analysis compared to transfer students overall. Colleges with very small numbers of students in the UC Davis engineering program are excluded from the table even if their average effects were large to avoid small sample bias.

	<i># in sample</i>	<i>Thriving Rate</i>	<i>2-year grad. rate</i>	<i>GPA at grad.</i>	<i>Percent minority</i>	<i>Percent low-inc.</i>	<i>Percent women</i>
<i>All</i>	1,240	48%	51%	2.98	18%	23%	12%
<i>College A</i>		29%	38%	2.71	45%	36%	24%
<i>College J</i>		20%	22%	2.80	35%	26%	16%
<i>College C</i>		57%	55%	3.18	20%	22%	15%
<i>College K</i>		58%	59%	3.19	11%	10%	15%
<i>College L</i>		54%	63%	3.36	15%	33%	21%
<i>College M</i>		55%	54%	3.23	21%	29%	5%
<i>College N</i>		56%	74%	3.26	13%	26%	16%

Table 3: Average outcomes and characteristics for UC Davis Engineering students from selected community colleges

Students from College J and College A performed poorly on average compared to other students. On the other hand, students from College N, College L, College K, College M, and College C did well on average. It is noteworthy that the low-performing colleges contributed much higher proportions of minorities and low-income students to our program. This makes uncovering the reasons for their poor performance relevant to the goals of Avenue E.

Among the target colleges for the Avenue E program, very few correlations with success were significant at the 5% level. Students from College G and College C performed somewhat better than average, and students from College A performed poorly as noted above. The others did not stand out as particularly strong or particularly weak.

### Part Two: Articulation Courses Analysis

The effects of pre-transferring courses and colleges of origin suggested a hypothesis that transferred courses, known commonly as articulations, might be inconsistent between students and cause differences in success. To test this hypothesis we examined articulation records for students in engineering from four specific colleges in detail. We selected College A to represent lower-achieving colleges, College C to represent higher-achieving colleges, and College L and College I to represent more typical-achievement colleges.

What we found indicated that most transfer students accomplished all of the required courses, showing the effectiveness of the selective major review. Some students met these requirements by taking the courses at different community colleges. For example, a student from College A might take several of the required courses at College C, especially computer science related courses.

It is notable that students from all colleges cross-register, so a student from College A is not getting all of their instruction at that institution. The numbers of cross-registering students and the colleges they attend seem to stem mainly from geography; there is no significant correlation between college success from Part One and the proportion of students taking required courses at other colleges. Overall, the findings on articulations contradict the hypothesis that students from low-achieving colleges are unable to access instruction or receive lower quality instruction in key courses.

### Part Three: Influential Courses for Engineering Transfers

Our third goal was to identify the courses at UC Davis that best predicted two-year graduation and thriving after the first year. To find the key courses, we considered two decision tree-based ensemble machine learning algorithms: random forest and XGBoost. The two models are known to be “strong learners,” meaning they develop close correlations with outcomes, and to have good performance in prediction. After applying the algorithms we combined the results and selected the 10 strongest predictors for each outcome.

#### Models

##### Decision Tree

Both random forest and XGBoost are decision-tree based algorithms. In a decision tree, an input is entered, and as it traverses down the tree, the data gets divided into smaller and smaller sets. At each node, a binary variable is used to split the data into two branches. A measure of the homogeneity of the target variable within the subset, like Gini impurity or entropy, is used to determine the best variable to use for the split at that node.

## Transfer Student Success in Engineering

Figure 1 shows how this process works in our data. This diagram shows the top three levels of one decision tree in our fitted model. At the top, we have 848 students, and a binary variable describes if a student gets A- or higher in EEC130A. There are 38 students with the answer yes and 810 with no. Of those students who do get the A-, most graduate in two years, while the other group is more variable. The next node in the tree for students who did not get the A- in EEC130A is whether they got an A- in ECH051. The tree continues in the same manner, dividing students according to their grades in different courses and determining which courses are most predictive of graduating in 2 years.

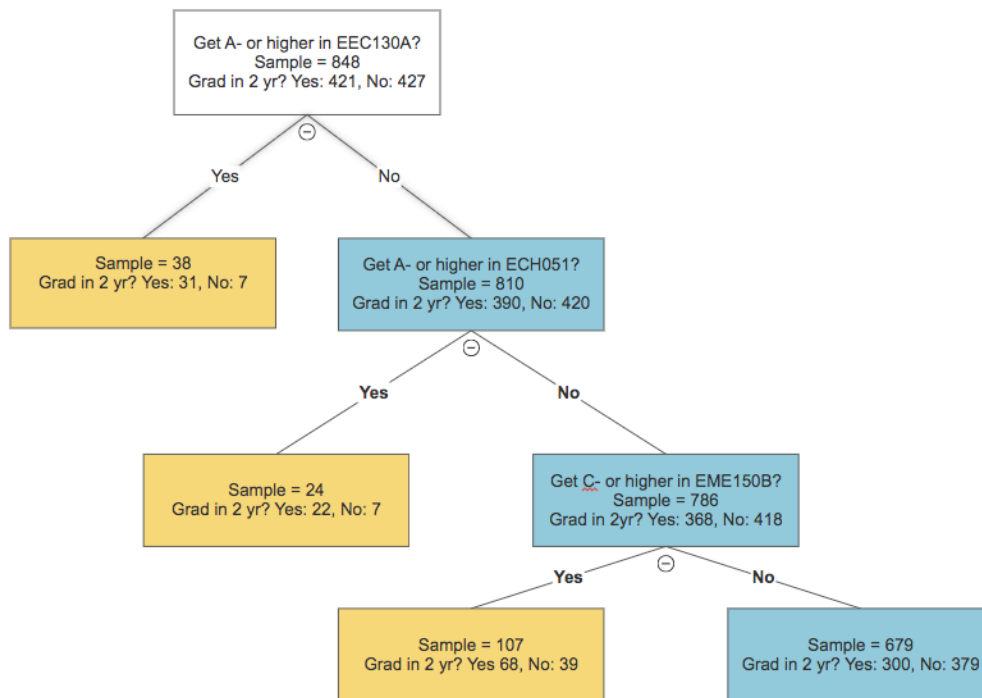


Figure 1: Decision tree based on our data

### Random Forest

Random forest operates by constructing a multitude of decision trees and outputting the predicted result that is the mode of the results from all decision trees (Breiman, 2001). Each decision tree is built from a small subset of variables at random and a random sample with replacement of the data.

### XGBoost

XGBoost is short for "Extreme Gradient Boosting" (Chen & Guestrin, 2016). XGBoost also operates by constructing multiple decision trees. What is different from the random forest is that the trees are dependent. The boosting algorithm puts more weight on challenging-to-classify instances and less on those already handled well. New decision trees are added



## Transfer Student Success in Engineering

sequentially that focus their training on the more difficult patterns (Brownlee, 2016). Generally, it performs better than random forest, and it is the top choice for the Kaggle, an online data science competition platform. However, it has some weaknesses compared with random forest, primarily that XGBoost needs more time for setting up and running because it is hard to compute in parallel and has more hyperparameters. It is also more prone to overfitting, especially in a high dimensional problem (Caruana Karampatziakis, & Yessenalina, 2008).

### Input Variables

In order to focus on courses, we included only course and major data in these models; we excluded demographics and other personal information that might predict a student's success. This reduced our ability to predict any individual student's outcome but allowed us to identify highly predictive courses without interference from interactions between course selections and demographics. These results can then be combined with demographic and other background information to learn how the courses fit into a larger model of student success, as we do in the final part of the study.

For the thriving model, the 70 courses most commonly taken by transfer students in their first year were included, while for the graduation model we used the most common 100 courses taken at any time before graduation. For each of these courses, grades and the majors of students were inputted to the model.

In a decision tree, each variable must be True/False. In order to fit course grades into this model we used cut points in the grades. We created three kinds of cut points associated with each course: *C-*, *A-* or *not take*. For the cut point *C-*, if a student gets *C-* or higher in that course, the corresponding variable would be *True*, otherwise, *False*. The cut point *A-* followed the same logic. When a student did not take the course, the variable with the cut point *not take* would be *True*. Thus, for every course, there were three corresponding variables. For example, if a student earned a *C* in ENG103, the variable associated with cut point *A-* would be *False*, the variable associated with cut point *C-* would be *True*, and the variable associated with *not take* would be *False*.

## Results

### Graduation Model

The top 10 most predictive variables from our two-year graduation model are listed below in Table 4.

<i>Course</i>	<i>Title</i>	<i>Cut Point</i>
<i>EEEC110A</i>	Electronic Circuits I	<i>C-</i>
<i>ENG190</i>	Prof Resp Of Engr	<i>A-</i>
<i>ECH148A</i>	Chemical Kinetics	<i>C-</i>

## Transfer Student Success in Engineering

ENG105	Thermodynamics	C-
ECS188	Ethics in an Age of Tech	A-
EME171	Mechatronic Systems	A-
ENG103	Fluid Mechanics	A-
ENG190	Prof Resp Of Engr	C-
UWP102E	Writing in Engineering	Not Take
EEC180A	Digital Systems I	A-

Table 4: Top 10 Influential Courses for 2 Year Graduation

Different cut points were informative for different courses. These cut points can help us understand how the course predicted student success. For example, if the cut point C- is most informative then a student who cannot pass the course is likely to have trouble graduating on time. For an A- cut point, students who do very well in that course might be particularly well prepared to graduate on time while passing that course is not enough to signal anything in particular. The relationship between the course grades and the outcomes is not necessarily causal but getting a good grade in the course is a signal that a student has whatever latent traits it takes to graduate on time.

There are some similarities among the courses we picked, so we put them into four categories. We describe the DFW (D, F, and withdrawal) rate, grade gaps between transfers and freshman admits, grade distributions and some other details below and in Table 5. The grade distribution plots and more information about these courses are shown in Figure 2 and Table 9 in the appendix.

- EEC110A (Electronic Circuits I), ENG105 (Thermodynamics), ENG103 (Fluid Mechanics), and EEC180A (Digital Systems I):** These four were all large classes with high DFW rates. They were challenging courses for all students, not only for transfer students. If a student got an A in the class, it might be a strong signal for graduating on time. If a student could not get C- or higher on these required courses, it would make graduating on time difficult. Transfer students on average performed less well than freshman admits in EEC180A, but better in ENG103 and ENG105. There was no significant gap in EEC110A.
- ENG190 (Professional Responsibility of Engineers), ECS188 (Ethics in an Age of Tech), and UWP102E (Writing in Engineering):** These three courses were not typical engineering and technical courses. Based on the grade distributions, they were all GPA boosting courses, which meant if students could not get an A they might have serious difficulty with other courses. Furthermore, ENG190 had a significant grade gap and DFW rate gap between transfer and freshman-admit students, suggesting it is more difficult for transfer students.
- ECH148A (Chemical Kinetics):** This course had a very low DFW rate, thus if a student got C- or lower it was a strong signal for not graduating on time.

## Transfer Student Success in Engineering

- **EME171 (Mechatronic Systems):** Not many students achieved A- or higher in this class. If a student did manage an A, they had a high probability of earning their degree in two years.

Course	Title	When <sup>(A)</sup>	Num Of Stud/Yr	Grade Gap <sup>(B)</sup>	DFW% <sup>(C)</sup>
EEC110A	Electronic Circuits I	Yr 1 W	129.22	-0.04	18%
ENG190	Prof Resp Of Engr	Yr 1 W/S	419.67	0.15	2%
ECH148A	Chemical Kinetics	Yr 2 F	104.80	0.02	2%
ENG105	Thermodynamics	Yr 1 W	340.90	-0.19	13%
ECS188	Ethics in an Age of Tech	Yr 2 Any	132.10	0.05	2%
EME171	Mechatronic Systems	Any F/W	89.70	0.05	7%
ENG103	Fluid Mechanics	Yr 1 F	323.20	-0.12	16%
UWP102E	Writing in Engineering	Any	152.60	0.09	2%
EEC180A	Digital Systems I	Yr 1 F	129.30	0.15	15%

(A) When: The most common time to take the course is derived from the program suggested plan and the CEE tool: Know Your Students.

(B) Grade Gap: Freshman admit average grade – Transfer average grade

(C) DFW: Percent of students earning D, F, or withdrawing from the course

Table 5: Simple summary of the top 10 courses for two-year graduation

### Thriving Model

As we only considered the courses which students took in the first year for the thriving model, we got a different set of highly predictive courses from the thriving model. Unlike in the graduation model, we did not find variables associated with the C- cut point to be important in this model. Instead, most of the highly predictive variables used the A- cut point.

Course	Title	CUT POINT
ENG102	Dynamics	A-
ENG103	Fluid Mechanics	A-
ENG106	Engin Economics	A-
ENG104	Mech of Materials	A-
ENG190	Prof Resp Of Engr	Not Take
ECH142	Heat Transfer	A-
EME150A	Mechanical Design	A-
ENG100	Elec Circuits & Systems	A-
EEC130A	Electromagnetics I	A-
EME050	Manufacturing Processes	A-

Table 6: Top 10 Courses for Thriving

Table 6 summarizes the ten most predictive course variables for thriving. The detailed information and grade distributions of these courses are shown in the appendix (Figure 3 and Table 10) and summarized in Table 7. As with the graduation model, we can put these ten courses into four categories.

- **ENG103 (Fluid Mechanics), ENG104 (Mech of Materials), and EEC130A (Electromagnetics I):** These three courses were difficult for all students with low average grade and high DFW rate. In ENG103 and ENG104, transfer students did better than the freshman admits.
- **ENG102 (Dynamics) and ENG106 (Engin Economics):** These two courses were of medium difficulty. The percentage of students getting A and B were similar. The DFW rates were not high. Transfer students got higher grades than freshman admits in ENG102.
- **ECH142 (Heat Transfer):** This course was also not very difficult to pass, but getting A was not easy in this class. A student with an A was likely to also excel in other courses.
- **ENG100 (Elec Circuits & Systems), EME050 (Manufacturing Processes), EME150A (Mechanical Design), and ENG190 (Prof Resp Of Engr):** The average grades of these four courses were 3.0 or above. Thus, if a student could not achieve a high grade, there was a lower probability of thriving in the 2<sup>nd</sup> year. For ENG190, students who not take the course during their first year may have had a grade boost in their 2<sup>nd</sup> year 1<sup>st</sup> quarter by taking this class at that time – they are more likely to meet our definition of thriving, but this does not mean they are doing better overall.

Course	Title	When <sup>(A)</sup>	Num Of Stud/Yr	Grade Gap <sup>(B)</sup>	DFW% <sup>(C)</sup>
ENG102	Dynamics	Yr 1 F	303.60	-0.17	7%
ENG103	Fluid Mechanics	Yr 1 F	323.20	-0.12	16%
ENG106	Engin Economics	Yr 1 W	181.56	0.00	5%
ENG104	Mech of Materials	Yr 1 W	313.20	-0.13	11%
ENG190	Prof Resp Of Engr	Yr 1 W/S	419.67	0.15	2%
ECH142	Heat Transfer	Yr 1 S	102.22	-0.01	6%
EME150A	Mechanical Design	Yr 1 F/S	141.70	-0.09	5%
ENG100	Elec Circuits & Systems	Yr 1 W	246.89	-0.02	7%
EEC130A	Electromagnetics I	Yr 1 W/F	97.20	-0.06	13%
EME050	Manufacturing Processes	Yr 1 W/F	144.20	-0.07	1%

Table 7: Simple Summary of the Top 10 Influential Courses for Thriving

## Transfer Student Success in Engineering

- (A) When: The most common time to take the course is derived from the program suggested plan and the CEE tool: Know Your Students.  
 (B) Grade Gap: Freshman admit average grade – Transfer average grade  
 (C) DFW: Percent of students earning D, F, or withdrawing from the course

### Part Four: Background, Preparation, and Key Courses Model

While the machine learning models are excellent for providing lists of key courses, they do not help with interpretation. In order to validate and interpret the relevance of these courses, we used the same background and preparation variables as in Part One along with the key courses identified by the machine learning models in linear probability models of thriving and 2-year graduation. For each course, we used the same binary grade variable that received a high importance rating in the machine learning model. The results are shown in Table 8. Coefficients can be interpreted as in Part One.

Most of the key courses we added to the model have significant effects on the outcomes, which validates the results of the machine learning models. However, the non-typical engineering courses, like ENG190 and UWP102\_E, had small coefficients and were not statistically significant in the two-year graduation regression model. The EME150A and EME050A were also insignificant in the new model. The predictive value of these courses may be better explained by the background variables from the original regression models. The influences of the demographic variables and the pre-transferring courses were similar to before, suggesting that the courses with large and significant coefficients in this model predict success for different reasons than the background variables.

	<i>Course</i>	<i>Cut Point</i>	<i>Two-year graduation rate</i>	<i>Thriving after 1<sup>st</sup> year</i>
EEC110A	Electronic Circuits I	C-	0.271***	0.045
ENG190	Prof Resp Of Engr	A-	0.035	0.038
ECH148A	Chemical Kinetics	C-	0.472***	0.053
ENG105	Thermodynamics	C-	0.191***	0.036
ECS188	Ethics in an Age of Tech	A-	0.138*	0.067
EME171	Mechatronic Systems	A-	0.157**	0.054
ENG103	Fluid Mechanics	A-	0.100*	0.044

## Transfer Student Success in Engineering

ENG190	Prof Resp Of Engr	C-	0	0.033	
EEC180A	Digital Systems I	A-	0.258***	0.057	
UWP102E	Writing in Engineering	Not Taking	-0.021	0.029	
ENG102	Dynamics	A-		0.107*	0.044
ENG103A	Fluid Mechanics	A-		0.232***	0.046
ENG106	Engin Economics	A-		0.282***	0.053
ENG104	Mech of Materials	A-		0.116*	0.045
ENG190	Prof Resp Of Engr	Not Taking		-0.070*	0.03
ECH142	Heat Transfer	A-		0.552***	0.046
EME150A	Mechanical Design	A-		0.107	0.06
ENG100	Elec Circuits & Systems	A-		0.199***	0.046
EEC130A	Electromagnetics I	A-		0.325***	0.067
EME050	Manufacturing Processes	A-		0.002	0.047

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Results from linear mixture model regressions including random effects for college of origin.

*Table 8: Regressions of student success measures on pre-transfer characteristics and key courses at UC Davis*

## Conclusions

In this study we examined background and demographic factors, preparation at community college, and courses at UC Davis to identify the strongest predictors of transfer students success in our engineering program. We found that demographic factors, in particular minority status and TAG, continued to be important even after considering coursework before and after transfer. While students from different community colleges had varying success rates, these differences seem to stem from factors other than the instruction they received; nearly all transfer students completed all of the core preparatory courses, often at colleges with much higher or lower average GPAs or rates of thriving or two-year graduation compared to their official college of origin.

The UC Davis courses that best predicted success included a mix of universally challenging courses, more average large courses, and less-mathematical courses in which many students typically earned high grades. After testing these courses in regression models including demographic factors, those less-mathematical courses did not appear to be as important while several of the courses that both transfers and freshman admits find difficult were strongly

## Transfer Student Success in Engineering

predictive of transfer students' success. Specifically, the course grades most closely correlated with thriving at UC Davis included Fluid Mechanics, Electromagnetics I, Engineering Economics, Heat Transfer, and Electrical Circuits and Systems. High grades in these courses indicated a high probability of earning a GPA over 3.0 with a full courseload in the student's fourth quarter. The course grades most correlated with graduating in two years were for Electronic Circuits I, Thermodynamics, Digital Systems I, Mechatronic Systems, and Chemical Kinetics. Very low grades or failure in some of these courses may push students off the course sequence needed to graduate on schedule, requiring them to delay more advanced courses until they can pass the prerequisite. We hope that with the continued support of advising staff at the college of engineering we can learn more in the coming years about using this information to help students.

References:

Breiman, L. (2001). Random Forests. *Machine Learning* ,45(1),5-32.

doi:<https://doi.org/10.1023/A:1010933404324>

Brownlee, J. (2016). A Gentle Introduction to the Gradient Boosting Algorithm for Machine Learning. Retrieved from

<https://machinelearningmastery.com/gentle-introduction-gradient-boosting-algorithm-machine-learning/>

Caruana, R., Karampatziakis, N., & Yessenalina, A. (2008, July). An empirical evaluation of supervised learning in high dimensions. In *Proceedings of the 25th international conference on Machine learning* (pp. 96-103). ACM.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system.

In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). ACM.

Freeling, Nicole. (October 2015) Napolitano: Transfer pathway is key to UC's success. Retrieved from: <https://www.universityofcalifornia.edu/news/napolitano-addresses-nations-community-college-leaders>

University of California (Dec 2017). Student Opportunity, Transfer Students Initiative. Retrieved from: <https://www.universityofcalifornia.edu/initiative/student-opportunity/transfer/transfer-students-initiative>



Appendix:

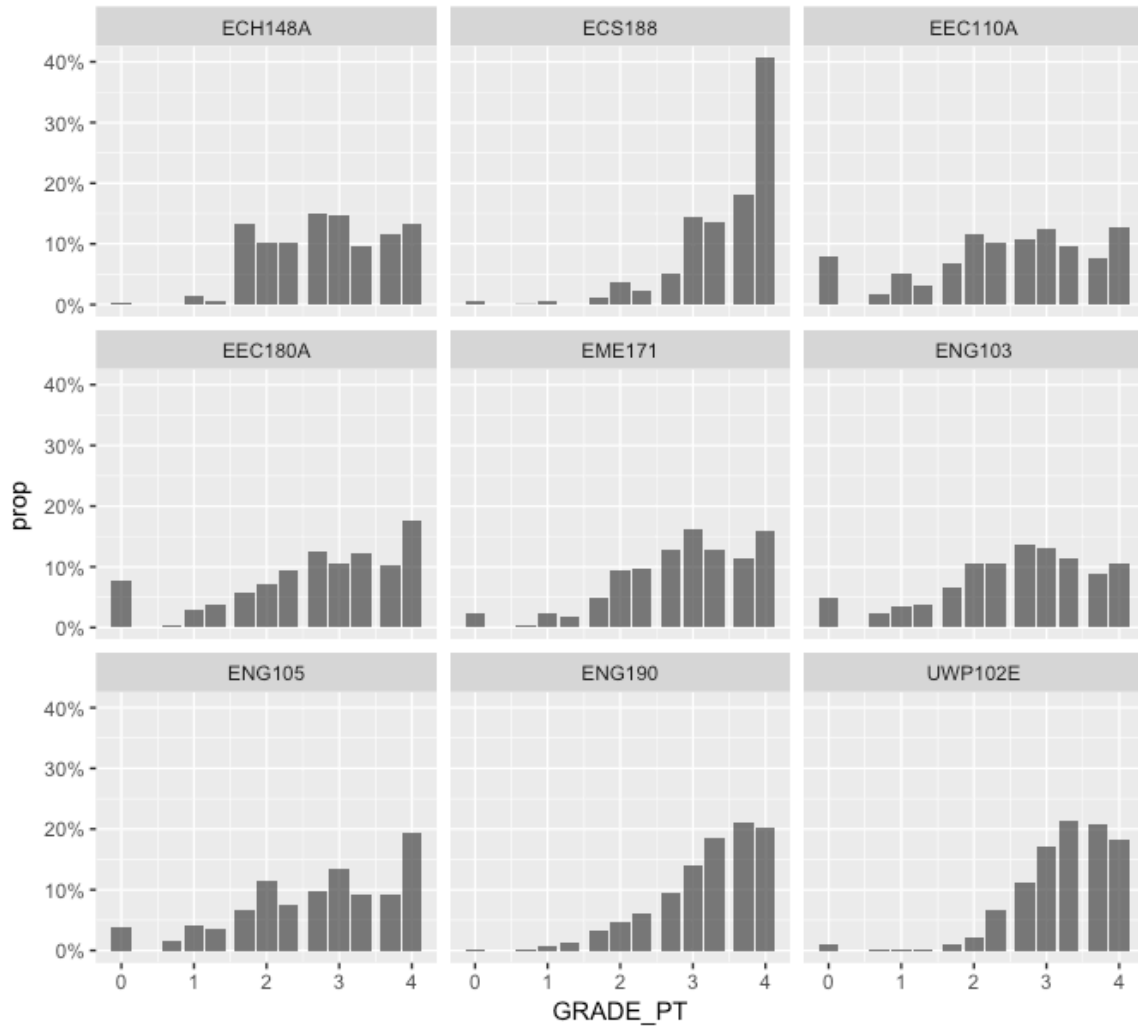


Figure 2: Top Courses Grade Distributions for the Graduation Model

### Transfer Student Success in Engineering

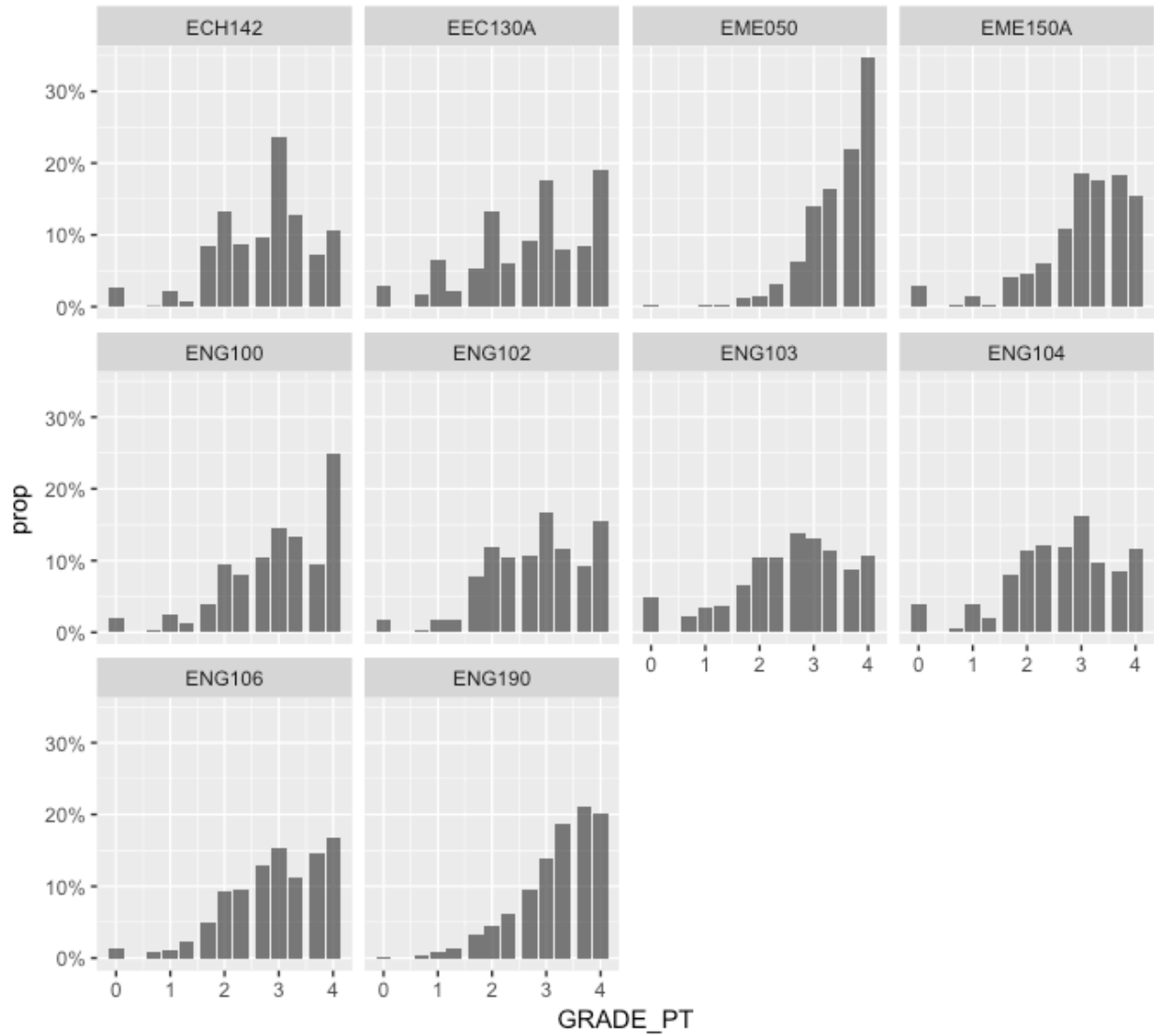


Figure 3: Top Courses Grade Distributions for the Thriving Model

Course	Title	Cut Point	RF Rank	XGB Rank <sup>1</sup>	#Stud/ Yr	#Trans fer/Yr	Transfe r Grade	Grade Gap	Grade of Triple	Triple gap <sup>2</sup>	DFW%	DFW% of Transfer	DFW % Gap <sup>3</sup>
EEC110A	Electronic Circuits I	C-	2	1	129.22	37.44	2.49	-0.04	2.27	0.22	18%	17%	3%
ENG190	Prof Resp Of Engr	A-	1	6	419.67	108.22	3.09	0.15	3.03	0.18	2%	4%	-2%
ECH148A	Chemical Kinetics	C-	4	Not in top 10	104.80	25.80	2.79	0.02	2.62	0.20	2%	3%	-1%
ENG105	Thermodynamics	C-	Not in top 10	4	340.90	89.30	2.85	-0.19	2.55	0.19	13%	11%	3%
ECS188	Ethics in an Age of Tech	A-	Not in top 10	5	132.10	47.10	3.42	0.05	3.39	0.07	2%	2%	-1%
EME171	Mechatronic Systems	A-	3	7	89.70	26.10	2.82	0.05	2.58	0.33	7%	9%	-3%
ENG103	Fluid Mechanics	A-	5	Not in top 10	323.20	90.70	2.64	-0.12	2.32	0.30	16%	13%	2%
ENG190	Prof Resp Of Engr	C-	7	3	419.67	108.22	3.09	0.15	3.03	0.18	2%	4%	-2%
UWP102 E	Writing in Engineering	Not Take	8	2	152.60	45.80	3.18	0.09	3.19	0.07	2%	3%	-1%
EEC180A	Digital Systems I	A-	6	9	129.30	37.50	2.57	0.15	2.39	0.33	15%	16%	-2%

Table 9: Information for Top Courses from the Graduation Model

<sup>1</sup> RF Rank and XGB Rank: Feature importance ranks of the random forest and XGBoost models

<sup>2</sup> Triple: which means a first-generation, under-represented minority and low-income student.

Triple gap: GPA of Non-Triple – GPA of Triple

<sup>3</sup> DFW% Gap: DWF Rate of Freshman - DWF Rate of Transfer Students

Course	Title	Cut Point	RF Rank	XGB Rank	#Stud/ Yr	#Transfer/ Yr	Transfer Grade	Grade Gap	Grade of Triple	Triple gap	DFW%	DFW % of Transfer	DFW % Gap
ENG102	Dynamics	A-	2	1	303.60	78.60	2.92	-0.17	2.56	0.29	7%	6%	0%
ENG103	Fluid Mechanics	A-	1	2	323.20	90.70	2.64	-0.12	2.32	0.30	16%	13%	2%
ENG106	Engin Economics	A-	4	3	181.56	39.56	2.92	0.00	2.70	0.28	5%	6%	0%
ENG104	Mech of Materials	A-	3	5	313.20	84.50	2.71	-0.13	2.31	0.38	11%	10%	1%
ENG190	Prof Resp Of Engr	Not Take	Not in top 10	4	419.67	108.2	3.09	0.15	3.03	0.18	2%	4%	-2%
ECH142	Heat Transfer	A-	5	Not in top 10	102.22	23.78	2.73	-0.01	2.59	0.15	6%	4%	2%
EME150A	Mechanical Design	A-	Not in top 10	6	141.70	39.70	3.10	-0.09	2.84	0.23	5%	4%	2%
ENG100	Elec Circuits & Systems	A-	6	Not in top 10	246.89	61.33	3.00	-0.02	2.90	0.10	7%	7%	-1%
EEC130A	Electromagnetics I	A-	7	Not in top 10	97.20	29.20	2.75	-0.06	2.51	0.21	13%	11%	4%
EME050	Manufacturing Processes	A-	Not in top 10	7	144.20	40.00	3.51	-0.07	3.31	0.18	1%	1%	0%

Table 10: Information for Top Courses from the Thriving Model