Institutional Effectiveness: Predicting Stop-outs based on High School GPA Deviations and Unmet Need

By Jake Morrison, Almat Yeraly, and Roy Caligan

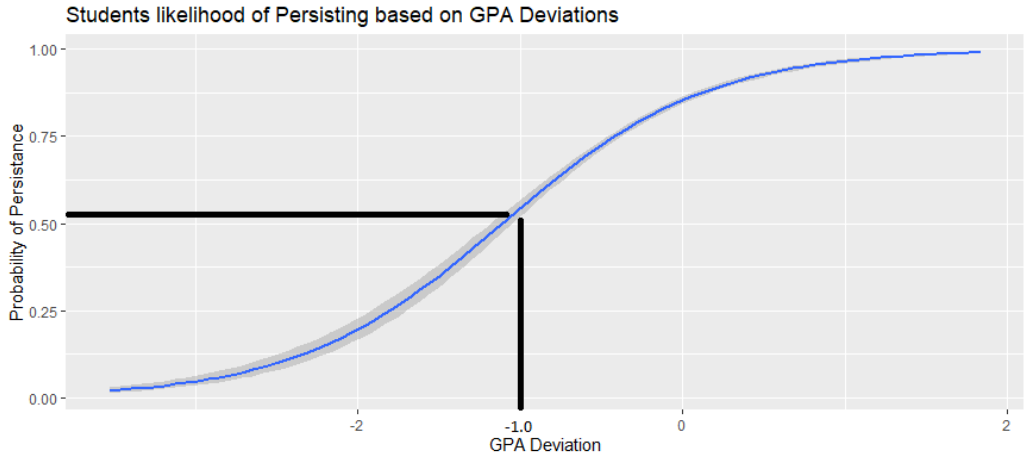
In this study, we attempt to key in on first time in College (FTIC) student's perception of self-efficacy and the effects their grades have on their ability to remain in College. Taking a student's high school GPA, we create a measure that takes the difference between final high school GPA and cumulative college GPA at the end of every term.

GPA Deviation at end of Spring Term

Minimum	Maximum	Mean	Median
-3.528	1.833	-0.177	-0.079

We then use this derivation to determine at what point a student is less likely to return to the University. Along with GPA deviation, we incorporate unmet financial need, number of credits a student brings into the university, an indicator for Running Start participation, an indicator for First Year Experience course taken, and a number of individual level controls such as race/ethnicity and gender.

The model[1] demonstrates a statistically significant effect of GPA deviations and unmet need on student's likelihood of persisting into the second year of their education at EWU. We model persistence from fall term to winter and spring, but present results of GPA deviation after spring quarter and likelihood of persisting into the following fall. The model predicts for every 1.00-point decrease in GPA a student is 26.40 percentage points less likely to persist[2]. Similarly, for every $1,000 of additional unmet need a student is 1.26 percentage points less likely to persist[3].



The estimates provided by our model may be used in harmony with a Naïve Bayes[4] machine learning algorithm that classifies students as being at risk of stopping out from the university after their first year. The results from these two models can help EWU administrators more effectively deploy resources and direct student level interventions. Along with this, we can use the estimates to forecast retention numbers and enrollment into future terms. The models used in the analysis may be expanded to differing student populations, such as continuing students, transfer students and returning students.

---

[1] We use a multivariate regression model with cohort year fixed effect and standard errors clustered at the major level.

[2] Confidence interval: [0.02495,0.2786] (*CI* = 95%, $p < 0.01$)

[3] Confidence interval: [-0.015,-0.010] (*CI* = 95%, $p < 0.01$)

[4] Naïve Bayes model demonstrates 80.36% prediction accuracy