Supplemental Predictive Model of Spring to Fall for First-Time Full-time Undergraduate Students



2019

Institutional Effectiveness and Accreditation

University of Idaho

3/14/2019

Contents

[Reexamination of Spring to Fall Progression 2](#_Toc3466362)

[Executive Summary 2](#_Toc3466363)

[Development of Spring to Fall Prediction Model First-Time Full-time Undergraduates 3](#_Toc3466364)

[Treatment of Racial Self-Identification: 4](#_Toc3466365)

[Summary Findings (take-away) 4](#_Toc3466366)

[References 5](#_Toc3466367)

[Appendix A: Model Maximum Likelihood Analysis Summary 7](#_Toc3466368)

***Contact: Dale Pietrzak, Ed.D., Director, Institutional Effectiveness and Accreditation***

*dalepietrzak@uidaho.edu*

# Reexamination of Spring to Fall Progression

## Executive Summary

There are several implications of this examination of the final model.

1. Academic performance continues to be the most significant factor impacting new student retention.
2. The new variable “registration in 15 or more credit hours in the fall” appears as a relatively strong predictor of spring to fall progression/retention for FTFT students. The two variables of institutional credit hours and a student’s first math course at the remedial level were borderline significant predictors.
3. This model can also serve as the basis for a decision tree-based decision-making process when conceptualizing how to think about students’ risk of exodus. The relative importance and grouping of predictors can be used to prioritize what areas of a student’s situation to consider first, and as most important. However, as “importance” is a personal element, the relative importance in the general population compared to the individual student can be quite varied.
4. As discussed previously, the elements of student behavior and attitude are largely missing from this model. Steps are underway to remedy this, but they are not yet operational. Data collected by VandalStar may enable such development here. Further, after conversations with Advising, and other potential end users, the Freshman Survey (CIRP) administered by IEA is scheduled to be replaced (pilot) with ACT Engage. This would be administered soon after a student’s arrival on campus and should allow for the development of key data in this area. Lastly, a Student Experience Survey is being piloted (second semester) and it may provide information to assist in the refinement of the model over time.
5. A note of caution. This model, like similar models elsewhere, still has a higher than desired false classification rate. This is to some extent an artifact of such models but has sufficient utility to be used to enter an advising discussion. This also means that this approach is currently not likely to lead to greater precision in predicting numbers for next-term enrollments than other approaches.

## Development of Spring to Fall Prediction Model First-Time Full-time Undergraduates:

This model was developed using the same methodology as described in the report covering the development of a model of spring to fall progression/retention for all undergraduates. This model is based on those who were enrolled full-time in the spring at census date. These further had to have been new first-time students in the previous fall.

This model included several new variables that had not been examined previously. These included registration in 15 or more semester hours the previous fall term, first course taken in English or math was remedial, and tuition minus aid differences.

The trimmed model completed appropriately (found solution). The overall model has an area under the curve (ROC) of 0.90, and rescaled R-Square of 0.37. Pearson Goodness of fit indicated a reasonable fit of the model to the data. This is a relatively strong model that is consistent with information available from the literature and conferences.

The relative importance of the various elements in the model accounting for student exodus is shown in the chart below. We see the largest factor by far associated with leaving U of I from spring to fall for FTFT students is academic performance (cumulative GPA). The next items appear to be Greek membership followed by registration in 15 or more hours the previous fall term. The next pair of items are ethnicity-related. The next grouping appears to be made up of College of Art and Architecture, tuition-aid differences, international student standing, and academic standing. The last grouping is made up of College of Natural Resources, first-generation American Indian/Native American, and Asian.

A summary list of the statistically significant predictors and their operation is provided below.

* Cumulative GPA (lower is at risk)
* Greek Member (not being a Greek is higher risk)
* Registered for 15 or more fall hours (not taking 15 or more is at risk)
* Two or More Races (identification as other than 2 or more is at risk)
* Black or African American (identification as other than Black or Af. Am. is at risk)
* College of Art and Architecture (CCA is at risk)
* Tuition Aid Difference (the greater the amount due is at risk)
* International (being other than International is at risk)
* Academic Standing (being other than in Good standing at risk)
* College of Natural Resources (CNR is at risk)
* First Generation (being 1st generation is at risk)
* American Indian or Native American (identification as Am. Ind. at risk)
* Asian (identification as Asian at risk)

### Treatment of Racial Self-Identification:

The approach used here coded racial self-identification relative to the combined general population and within the context of all variables in the model (for example, white compared to non-white, or Native American compared to all non-natives, etc.). Usually ethnic predictors are related to campus environment and engagement factors.

## Summary Findings (take-away):

There are several implications of this examination of the final model.

1. Academic performance continues to be the most significant factor.
2. The new predictor of registration in 15 or more credit hours in the fall appears as a relatively strong predictor of spring to fall progression/retention for FTFT students. . The two variables of intuitional credit hours and a student’s first math course at the remedial level were borderline significant predictors.
3. This model can also serve as the basis for a decision tree-based decision-making process when conceptualizing how to think about students’ risk of exodus. The relative importance and grouping of predictors can be used to prioritize what areas of a student’s situation to consider first, and as most important. However, as “importance” is a personal element, the relative importance in the general population compared to the individual student can be quite varied.
4. As discussed previously the elements of student behavior and attitude are largely missing from this model. Steps are underway to remedy this, but they are not yet operational. Data collected by VandalStar may enable such development here. Further, after conversations with Advising, and other potential end users, the Freshman Survey (CIRP) administered by IEA is scheduled to be replaced (pilot) with ACT Engage. This would be administered soon after a student’s arrival on campus and should allow for the development of key data in this area. Lastly, a Student Experience Survey is being piloted (second semester) and it may provide information to assist in the refinement of the model over time.
5. A note of caution. This model, like similar models elsewhere, still has a higher than desired false classification rate. This is to some extent an artifact of such models but has sufficient utility to be used to enter an advising discussion. This also means that this approach is currently not likely to lead to greater precision in predicting numbers for next-term enrollments than other approaches.

## References

Allison, P. (2014). Measure of fit for logistic regression. Paper 1485-2014. Statistical Horizons and University of Pennsylvania. (SAS)

Bean, J. & Eaton, S. (2001). The psychology Underlying Successful Retention Practices. *Journal of College Student Retention*, 3(1) 73-89.

Caison, A. L. (2006). Analysis of Institutionally Specific Retention Research: A Comparison Between Survey and Institutional Database Methods. *Research in Higher Education* 48(4): 435-451*.*

Collier, P. & Morgan, D. (2007). “Is that paper really due today?”: Differences in first-generation and traditional college students’ understanding of faculty expectations. *Higher Education*. DOI:10.1007/s10724-007-9065-5.

Demetriou, C. & Schmitz-Sciborski, A. (2012). Integration, motivation, strengths and optimism: retention theories past, present and future. In R. Hays (Ed.) Proceedings of the 7th National Symposium on Student Retention, 2011, Charleston (pp. 300-312). Norman, OK: The University of Oklahoma.

DesJardins, S. T. (2002). An Analytical Strategy to Assist Institutional Recruitment and Marketing Efforts. *Research in Higher Education* 43(5).

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters,* 27: 861-874.

Herzog, S. (2019), Student Success Perditions at your fingertips: Developing online dashboards with Microsoft Power BI. Percontor Workshop, 02-21-19.

Herzog, S. (2019). Exploring the power of predictive analytics: A step-by-step introduction to building a student-at-risk model. Percontor Workshop, 02-07-19.

Herzog, S. (2006). “Estimating student retention and degree-completion time: Decision trees and neural networks vis-à-vis regression.” In J. Luan & C. Zhao (eds.), *Data Mining in Action: Case Studies of Enrollment Management. New Directions for Institutional Research,* no. 131. San Francisco, CA: Jossey-Bass*.*

Herzog, S. (2005). “Measuring determinants of student return vs. dropout/stopout vs. transfer: a first-to-second year analysis of new freshmen.” *Research in Higher Education,* 46(8): 883-928.

Hosmer, D. & Lemeshow, S. (2000). *Applied Logistic Regression (Second Edition).* New York: John Wiley & Sons, Inc.

Luke, C. (2009), "An Examination of Psychological Factors That Predict College Student Success and Retention." Ph.D. dis., University of Tennessee. <https://trace.tennessee.edu/utk_graddiss/71>.

Pascarella, E. T., & Terenzini, P. T. (2005). *How College Affects Students: A Third Decade of Research, Volume 2.* San Francisco, CA: Jossey-Bass

Richardson, M., Abraham, C. & Bond, R. (2012). Psychological correlates of university students’ academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353-387. DOI: 10.1037/a0026838.

Rodgers, K. & Summers, J. (2008). African American Students at Predominantly White Institutions: A Motivational and Self-Systems Approach to Understanding Retention. *Educ Psychol Rev*, 20, 171-190. DOI 10.1007/s10648-008-9072-9.

Thompson, D. (2009). Ranking Predictors in Logistic Regression. Paper D10‐2009. Assurant Health. (SAS)

## Appendix A: Model Maximum Likelihood Analysis Summary

**Note many variables are coded in the software by defaults as:**

|  |  |
| --- | --- |
| **0** | 1 |
| **1** | -1 |

