Antecedents of Student Retention: A Predictive Modelling Approach

Prathamesh Muzumdar¹ | Ganga Prasad Basyal² | Piyush Vyas³

¹Doctoral Student Dakota State University College of Business and Information Systems Madison, South Dakota, USA
²Doctoral student, Dakota State University
³Doctoral student, Dakota State University

Abstract: Today university and college rankings are the major determinants influencing the enrollment decision of the current student. Universities improving on certain aspects are found to retain a majority of freshman students despite the university tier status. These aspects are seen to determine a student’s retention, which eventually helps the university to attend a higher tier status. The study tries to evaluate the aspects which primarily influence the future enrollment decision, thereby resulting in retaining the student. The evaluated aspects deal with the intrinsic factors responsible to support the academic tenure of a current student. The aspects, mainly revolve around the factors like peer effectiveness, academic progress, and academic advising out of which academic advising is a controlled variable addressed by the university administration. Two other variables are seen to indirectly be controlled by the administration. In the initial stage principal component analysis is used to identify the factors, followed by multiple regression analysis to determine the influence of each factor on the likelihood of retention. To validate the facts logistic regression analysis is used to verify the same factors that influence the actual retention of the student.

Keywords: Student Retention, Design Science, Multiple Linear Regression, Logistic Regression

1 | INTRODUCTION

The higher education institutes are different from the other industries, being a non-profit organization, they depend heavily on government and public funding. The services that higher education institutes provide are intangible and are centered on knowledge and student learning (Adelman, 1999). Student enrollment has been the main source of revenue for the institutes beyond the government and public funding and donations (Light, 2001). The challenge these institutes have been fac-

Supplementary information The online version of this article (https://doi.org/10.15520/ijcrr.v11i11.860) contains supplementary material, which is available to authorized users.

Corresponding Author: Prathamesh Muzumdar
Doctoral Student Dakota State University College of Business and Information Systems Madison, South Dakota, USA
Email: prathameshmuzumdar@gmail.com
ing for the past few years is freshman retention. It is quite evident that freshman prefers to transfer to other institutes for the sophomore year (Asera, 1998). University ranking, reputation, eliteness, and return on investments are the factors openly known to support such decisions (Mangold, Bean, Adams, Schwab, & Lynch, 2003).

Factors can be intrinsic or extrinsic (Noel, Levitz, & Saluri, 1985). This research study investigates factors which are intrinsic for whom the source is the institute itself. Institutes prefer to allocate their budget and invest their capital in resources and activities which can help them increase their enrollment (Bean, 1980 & 1983). This study focuses on the influence and effects of intrinsic factors on freshman retention. It is important to locate and investigate the internal factors which lead to a better student experience (Belgrade & LoRe, 2003 & Pascarella & Terenzini, 2005). Some factors also lead to better student learning, while others contribute to tacit knowledge. Overall, the positive to consider is on the higher side the negatives that lead to dissatisfaction, denial, and turnover require attention (Bean, 1985 & Roach, 1997).

This study solely focuses on the part where institutes control the factors and can further better understand the influence of their decisions control the internal institute-based activities. The study addresses many factors that can be solely discerned of the institute policies that can lead to increased retention. It is the need of the time to investigate the controllable factors which are subjected to institutional policy changes (Braxton, 2000 & Schnell & Doetkott, 2003). Hence, this study contributes to higher education literature for better retention and enrollment. The findings from the study can lead to improved freshman retention, better satisfaction, and better student learning outcomes.

2 | LITERATURE REVIEW AND RESEARCH QUESTION

The intrinsic factors of higher education

This research is a design science study that is methodological and develops mathematical model to predict outcome. The study contributes to the literature of design science method by exploring predictive models that can best fit the data. This study tries to examine the antecedents of student retention considering multiple predictive models and confines itself to multiple linear regression for its predictive superiority over other models. Further, it verifies the validity of this model using a logistic regression model using the outcome variable with verified student retention data. The research question under consideration is “Does multiple linear regression fits to be the best predictive model for student retention”.

The institutional intrinsic factors of higher education can be defined as the factors that are solely controlled by the institute administration and its policies (Braxton & McClendon, 2002). The institute must implement these activities in the assumption that they will lead to the betterment of a student’s progress (Schnell, Seashore, & Doetkott, 2003). Intrinsic factors can be also considered as antecedents of student retention if they positively influence the students’ learning process and satisfies them with the outcome (Braxton, Hirschy, & McClendon, 2004). For this study, 8 factors were considered as potential influencers of retention. Those 8 factors were identified as faculty effectiveness, college costs, academic
progress, peer effectiveness, academic advising, academic course load, and academic success.

The outcome for retention was likeliness to continue which was used as the outcome variable. Faculty effectiveness is defined as the overall quality of instructions delivered through any medium (Cabrera, Casteneda, Nora, & Hengstler, 1992). It is taken for granted that with instructions some amount of knowledge is imparted from the faculty to the students. It is important to understand the role of faculty in leading a student to retain his status in the institute (Cabrera, Nora, & Castaneda, 1993). College cost is one most important factor that fluctuate over time (Tinto, 1987 & 1997). It is completely controlled by the university administration and directly influenced by the running budget (Congos & Schoeps, 2003). It can act as a pivotal turning point when it comes to choosing an affordable option over others.

Academic progress is a factor considered to be guided by multiple aspects. In a research study, it has been proved that academic progress on either side of the scale can lead to a transfer (Carter, 1998 & Tucker, 1999). Students performing better may prefer to upgrade themselves to high ranked universities, while those doing poorly might prefer to choose the universities where they can complete the program with just above the threshold grades (Covington, 2000 & Crissey, 1997). Academic course load is one of those factors that seem to affect knowledge growth at the same time may lead to an increase in stress (Crokston, 1994 & Dervarics & Roach, 2000). It has been reported in studies that excessive course load and unplanned course addition can lead to student turnover in institutes (Fidler, 1991). Academic course load directly affects academic progress and is directly influenced by academic advising (Gilley, 1992).

Academic advising is defined as the structured and guided instructions needed to put forward a plan of study (Glass, 1997). Academic advising leads to better structuring of courses addressing individual needs and skillsets (Graham & Diamond, 1999). It puts forward an idea about the course structure and helps a student to build his e-portfolio around. Academic course load is directly affected by advising (Hensen & Shelley, 2003 & Hurd, 2000). Prompt advising can lead to a better plan, further leading to a feasible course load (Hurte, 2002). A student needs to get the right advice to build a structured plan of study. In the end, peer effectiveness is an important factor (Ishitani & DesJardins, 2002). Fellow students, classmates, colleagues, roommates, etc. are very influential factors when it comes to decision making at the institutional level (Karp & Logue, 2002). Understanding the influence of peer effectiveness can lead to developing a better strategy to influence and retain most freshmen.

**Concurrent Validity**

The concurrent validity model is a framework used in predictive modeling when validating a predictive model like a linear regression model with another outcome model (Kern, Fagely, & Miller, 1998). The outcome variable differs in which the first model is used to understand the influence of the factors on the predicted likelihood of the outcome while the second model is used to validate the factors using the outcome itself. The second model is run once the outcome has occurred. This model helps to validate and verify if the same factors were significantly affecting the outcome variable.

**Data Collection**

Data was collected using an online survey which was distributed among 2000 students at the university. The sample criteria were set to only include freshman and sophomore students. The survey had a total of 30 items (questions) representing 9 constructs. The nine constructs were reassumed through literature review and the items were tested using exploratory factor analysis (EFA) to verify if they contribute towards a particular construct. The survey had a screening section after the instructions. The job of the screening the respondents before the main is very important to get the right responses for the research.

A total of 750 respondents completed the survey. The survey was divided into nine different sections, each section comprising of a construct to be measured. The number of items for each section varied accordingly. The survey was programmed in such a way that no questions were missed and it was mandatory to answer each question before moving on to the next. Overall, the rate of response was 37.5% and all completed the questions in the survey.
The data was collected in 3 batches after sending 15 invitation and reminder emails over 15 weeks to the panel. The data was cleaned and re-coded wherever it was necessary. Further, the data was checked for abnormalities and patterns. The data was thoroughly scrutinized before being further analyzed.

3 | METHODOLOGY

The research was divided into 3 different studies. Study 1 was used to investigate the number of factors responsible for the likelihood of future enrollment of a student. The steps are taken to investigate this phenomenon initially consisted of breaking down 32 variables into a few selected factors using principal component analysis. Then in study 2, the influence of each factor on the dependent variable was determined using multiple regression. The sole purpose of this analytics was to segregate and evaluate significant factors contributing to the dependent variable.

Study 3 was used to validate the factors by testing them on real retention of the freshman student. Retention being a categorical variable in nature and using it as a dependent variable, logistic regression stands out to be the method that can be used in such a condition, keeping in mind the objective of the study to evaluate the influence of each factor on the freshman retention. Discriminant analysis is another method that can be used to do the same function is also used to group factors in a particular function. Functions in the Discriminant analysis represent two different associated with the dependent variable. The main purpose of the Discriminant analysis has segregated the factors into different groups while logistic regression analysis helps to understand the influence of each factor on the actual retention.

Overall, both the sections are used to compare the results so that the factors explored in study 2 can be validated and cross-checked from study 3. Study 1 helps us to understand the factors responsible for the likely hood retention, improving these factors would help a university secure, high retention before the students transfer or drop out of the school. These factors also indirectly affect the tier ranking of the university where mandatory criteria like GPA are strongly affected by peer effectiveness and academic advising. Study 3 inspects each of these factors on the real retention of students and clarifies the hypothetical relationship between retention and the inspected factors. Overall, the methodology tries to clarify the importance of each factor in freshman retention.

4 | RESULTS

Study 1: Dimension Reduction using Exploratory Factor Analysis

Study 1 was used to identify and determine the influential factors affecting the likelihood of future enrollment and not actual retention. Results of the principal component analysis indicate that 47 metric variables are reduced into 9 factors. Further out of these 9 factors, 8 factors were found to pass a reliability test with Cronbach’s alpha value over 0.8. These 8 factors were named as follows faculty effectiveness,
college costs, academic progress, peer effectiveness, major, academic advising, likely to continue, academic course load, and academic success.

**Table 02: Multiple Regression Analysis**

| Model                        | Unstandardized Coefficients | Standardized Coefficients | t     | Sig |
|------------------------------|-----------------------------|---------------------------|-------|-----|
| 1 (Constant)                 |                             |                           | 5.035 | 0.00|
| Peer Effectiveness           | -0.67                       | -0.461                    | -9.839| 0.00|
| Academic Progress            | -0.95                       | -0.518                    | -13.348| 0.00|
| Academic Advising            | 2.798                       | 1.312                     | 23.441| 0.00|
| R                             | 0.785                       |                           |       |     |
| R Square                     | 0.621                       |                           |       |     |
| Adj R Square                 | 0.618                       |                           |       |     |
| Std Error of Estimate        | 1.087                       |                           |       |     |
| F Value                      | 234.772                     |                           |       |     |
| Sig                          | 0.008                       |                           |       |     |

a. Predictors: (Constant), Academic Advising, Academic Progress, Peer Effectiveness

**Study 2: Multiple Linear Regression Equation Model**

\[ y = b_0 + b_1 \times \text{Peer Effectiveness} + b_2 \times \text{Academic Progress} + b_3 \times \text{Academic Advising} + \epsilon \]

Study 2 was used to evaluate the significance of each factor, all factors were entered simultaneously in the regression analysis of the dependent variable likelihood of enrollment. The results of regression analysis indicate that three factors have a significant effect on the dependent variable and these 3 factors influence the likelihood of enrollment in either a positive or negative way. From figure 2, we can conclude that peer effectiveness, academic progress, and academic advising are responsible for affecting the likelihood of continuing at the university for the next semester. It is clear that peer effectiveness and academic progress negatively affect the likelihood to continue while academic advising positively affects the dependent variable. Considering the beta coefficients, academic advising (+1.312) has a stronger positive effect compared to peer effectiveness (-0.461) and academic progress (-0.518) in their negative forms. Thus, this study overall supports hypothesis 1.

**Study 3: Logistic Regression Equation Model**

\[ \log \frac{p}{1-p} = b_0 + b_1 \times \text{Peer Effectiveness} + b_2 \times \text{Academic Progress} + b_3 \times \text{Academic Advising} + \epsilon \]

Study 3 was used to externally validate the results of study 2. As a part of concurrent validity, the data enrolled for the sophomore year are retained and thus to further test the three factors, peer effectiveness, academic progress, and academic advising for their significant influence on the enrollment of students. The dependent variable was a categorical variable consisting of only two outcomes, yes or no, logistic regression was used to evaluate and validate the influence of each factor on the dependent variable. From figure 3, it is clear that all three factors are found to be significant at a 1% significance level.

### Table 03: Logistics Regression Analysis

|                      | B     | S.E. | Wald | df  | Sig  | Exp(B) |
|----------------------|-------|------|------|-----|------|--------|
| Step 1* Peer Effectiveness | -1.52 | 0.22 | 48.35 | 1   | 0.000 | 0.217  |
|          Academic Progress | -1.67 | 0.22 | 53.61 | 1   | 0.000 | 0.188  |
|          Academic Advising | 4.95  | 0.44 | 112.83 | 1   | 0.000 | 141.342|
|          Constant      | -3.54 | 0.88 | 15.88 | 1   | 0.000 | 0.029  |
| Chi Square           | 302.57|      |      |     |      |        |
| Sig                  | 0.000 |      |      |     |      |        |

a. Variable(s) entered on Step 1: Peer effectiveness, Academic Progress, Academic Advising
b. Estimation terminated at iteration number 7 because parameter estimates changed by less than 0.001
c. The Cut value is .500

5 | DISCUSSION

This study has several contributions to the field of higher education. Student retention in higher education has not been predictively investigated before. The use of concurrent validity model with regression analysis as a predictive instrument answers the call of many researchers to discuss the need for improvements in student retention in higher education.

Higher education institutes are considered to be a non-profit organization. Our findings can lead to better academic advising which can further lead to increased academic progress. A positive influence of peer effects on student retention suggests that the mentor programs and group therapy can lead to improving student confidence in others. All this constitutes to develop a better student retention strategy the university can focus on a better allocation of funds.

Peer effectiveness should be encouraged in the academic structure, due to its numerous benefits, as
proved in this study. In a time where higher education institutes face increased criticism by the people and face challenges such as budget cuts, enrollment decline, faculty turnover, etc. One main takeaway from this research is that the higher education institute should adopt a better academic advising approach.

6 | LIMITATIONS AND FUTURE RESEARCH

This study has few limitations as common with any study using empirical analysis. The data was collected from a single source (panel) from an institute ranked and accredited on a particular level. There is a chance that there was a response bias leading to causal inference. Future research can use longitudinal data collected over some time on certain intervals which will enable the researchers to correct the response bias effect in a causal relationship.

REFERENCES

1. Adelman, C. (1999). Answers in the toolbox: Academic intensity, attendance patterns, and bachelor’s degree attainment. Washington, DC: U.S. Department of Education.

2. Asera, R. (1998). Supporting student persistence. Black Issues in Higher Education, 15(10), 104.

3. Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. Research in Higher Education, 12(2), 155-187.

4. Bean, J. P. (1983). The application of a model of turnover in work organizations to the student attrition process. Review of Higher Education, 6(2), 129-148.

5. Bean, J. P. (1985). Interaction effects based on class level in an explanatory model of college student dropout syndrome. American Educational Research Journal, 22(1), 35-64.

6. Belgarde M. J., & LoRe, R. K. (2003). The retention/intervention study of Native American undergraduates at the University of New Mexico. Journal of College Student Retention: Research, Theory & Practice, 5(2), 175-203.

7. Braxton, J. M. (Ed.). (2000). Reworking the student departure puzzle. Nashville, TN: Vanderbilt University Press.

8. Braxton, J. M., & McClendon, S. A. (2002). The fostering of student integration and retention through institutional practice. Journal of College Student Retention: Research, Theory & Practice, 3(1), 57-71.

9. Braxton, J. M., Hirschy, A. S., & McClendon, S. A. (2004). Understanding and reducing college student departure. ASHE-ERIC Higher Education Report, 30(3).

10. Cabrera, A. F., Casteneda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. Journal of College Student Development, 63(2), 143-164.

11. Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. Journal of Higher Education, 64(2), 123-139.

12. Carter, T. (1998). “Rankled by the Rankings,” ABA Journal, Vol. 84. 46-53.

13. Congos, D. H., & Schoeps, N. (2003). Inside Supplemental Instruction (SI): One model of what happens that improves grades and retention revisited. Journal of Student Centered Learning, 1(13), 159-170.

14. Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. Annual Review of Psychology, 51, 171-200.
15. Crissey, M. (1997). “Changes in Annual College Guides Fail to Quell Criticisms on their Validity,” The Chronicle of Higher Education, Vol. 44, No. 2, A67.

16. Crookston, B. B. (1994). A developmental view of academic advising as teaching. NACADA Journal, 14(2), 5–9.

17. Dervarics, C., & Roach, R. (2000). Fortifying the federal presence in retention. Black Issues in Higher Education, 17(3), 20-25.

18. Fidler, P. P. (1991). Relationship of freshman orientation seminars to sophomore return rates. Journal of the Freshman Year Experience, 3(1), 7-38.

19. Gilley, J.W. (1992). “Faust Goes to College,” Academe: Bulletin of the AAUP, Vol. 78, No. 3, May/June, 9-11.

20. Glass, S. (1997). “The College Ranking Scam,” Rolling Stone, 93- 94.

21. Graham, H.D. & Diamond, N. (1999). “Academic Departments and the Rating Game,” The Chronicle of Higher Education, Vol. 45, No. 41, B6.

22. Hensen, K. A., & Shelley, M. C., III. (2003). The impact of Supplemental Instruction: Results from a large, public midwestern university. Journal of College Student Development, 44(2), 250-259.

23. Hurd, H. (2000). Staying power: Colleges work to improve retention rates. Black Issues in Higher Education, 17(18), 42-46.

24. Hurte, V. J. (2002). Mentoring: The forgotten retention tool. Black Issues in Higher Education, 19(18), 49.

25. Ishitani, T., & DesJardins, S. (2002). A longitudinal investigation of dropout from college in the United States. Journal of College Student Retention: Research, Theory & Practice, 4(2), 173-201.

26. Karp, R., & Logue, R. (2002). Retention initiative for unscheduled sophomores and unscheduled readmits. Journal of College Student Retention: Research, Theory & Practice, 4(2), 147-172.

27. Kern, C. W., Fagley, N. S., & Miller, P. M. (1998). Correlates of college retention and GPA: Learning and study strategies, testwiseness, attitudes, and ACT. Journal of College Counseling, 1(1), 26-34.

28. Light, R. J. (2001). Making the most of college: Students speak their minds. Cambridge, MA: Harvard University Press.

29. Mangold, W. D., Bean, L. G., Adams, D. J., Schwab, W. A., & Lynch, S. M. (2003). Who goes who stays: An assessment of the effect of a freshman mentoring and unit registration program on college persistence. Journal of College Student Retention: Research, Theory & Practice, 4(2), 95-122.

30. Noel, L., Levitz, R., & Saluri, D. (1985). Increasing student retention: Effective programs and practices for reducing dropout rate. San Francisco: Jossey-Bass.

31. Pascarella, E. T., & Terenzini, P. T. (2005). How college affects students: A third decade of research. San Francisco, CA: Jossey-Bass.

32. Roach, R. (1997). Retention pacesetter: University of Virginia. Black Issues in Higher Education, 14(20), 35.

33. Schnell, C. A., & Doetkott, C. D. (2003). First year seminars produce long-term impact. Journal of College Student Retention: Research, Theory & Practice, 4(4), 377-391.

34. Schnell, C. A., Seashore Louis, K., & Doetkott, C. (2003). The first-year seminar as a means of improving college graduation rates. Journal of the First-Year Experience and Students in Transition, 15(1), 53-75.
35. Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45(1), 89–125.

36. Tinto, V. (1987). Leaving college: Rethinking the causes and cures of student attrition. Chicago: The University of Chicago Press.

37. Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. Journal of Higher Education, 68(6), 599-623.

38. Tucker, J. (1999). Tinto’s model and successful college transitions. Journal of College Student Retention: Research, Theory & Practice, 1(2), 163-175.

How to cite this article: Muzumdar P., Basyal G.P., Vyas P. Antecedents of Student Retention: A Predictive Modelling Approach. International Journal of Contemporary Research and Review. 2020;21906–21913. https://doi.org/10.15520/ijcrr.v11i11.860