2000

First Year Versus Second Year Retention of College Students: A Case Study

Heather M. O'Neill

Ursinus College, HONEILL@URSINUS.EDU

Follow this and additional works at: https://digitalcommons.ursinus.edu/bus_econ_fac

Part of the Economics Commons, Education Economics Commons, Higher Education Commons, Higher Education Administration Commons, and the Liberal Studies Commons

Click here to let us know how access to this document benefits you.

Recommended Citation


https://digitalcommons.ursinus.edu/bus_econ_fac/9

This Article is brought to you for free and open access by the Business and Economics Department at Digital Commons @ Ursinus College. It has been accepted for inclusion in Business and Economics Faculty Publications by an authorized administrator of Digital Commons @ Ursinus College. For more information, please contact aprock@ursinus.edu.
FIRST YEAR VERSUS SECOND YEAR RETENTION OF COLLEGE STUDENTS: A CASE STUDY

Heather M. O'Neill
Department of Economics and Business Administration
Ursinus College
P.O. Box 1000
Collegeville, PA 19226-1000
e-mail: honeill@ursinus.edu

INTRODUCTION

Students and their families expend much time, effort and money researching which colleges or universities will best suit the students' needs. Simultaneously, institutions desire to find the cohort of students who will succeed at their schools. Recently, faced with more stringent economic constraints, schools are not only seeking students likely to succeed, but are more aware of the financial burden placed on schools if attrition is high. Since the cost of recruiting a class has risen over the years, the cost of losing students has increased. As a result, institutions are more interested in engaging in student retention studies to find ways to reduce attrition.

Some published research explores the determinants of student persistence using institutional data, though most of the literature uses national databases. Using institutional data, this paper examines whether a student is likely to remain or leave school, either voluntarily or involuntarily, during the first versus second year. Sophomore year attrition can exceed first year attrition at many schools, thus necessitating the study of the two years. A better understanding of the different determinants in the two years will assist enrollment administrators enhance student persistence. The data from students attending a small, selective, four-year liberal arts college provide results that are institution-specific, although the model and many results are applicable to other colleges and universities.

LITERATURE REVIEW

Tinto (1993) modeled student retention in a dynamic setting, contending students need to feel connected socially and academically in order to persist. Berger and Milem (1997) used path analysis to explore Tinto's model and found the student behavior-perception-behavior cycle adequately described how students became socially and academically integrated. Using institutional data from a highly selective private university, they examined first year retention. Their dependent variable, persistence, in the multivariate model was based on student intentions, not strictly on actual persistence. They found that females, Caucasians, African Americans, students with above average high school grades and higher income students showed significantly higher degrees of initial commitment to their institutions. Greater commitment, combined with activities and involvement with peers and faculty, increased their intention to enroll.

St. John, Oescher, & Andrieu (1992) used the National Postsecondary Student Aid Study from 1987 (NPSAS-87) to examine the effect of financial aid on within-year persistence for all students. They built a logistic regression model and reported that males were more likely to persist, as were married persons and those attending full time. Students aspiring to complete some years of college were more likely to persist, whereas those looking to advanced degrees were less likely to persist. While B students were more likely to re-enroll than C students, so were those with below C averages in college. Tuition increases reduced persistence, yet increases in loans and work-study income also reduced it.

Somers (1996) used institutional data from a large urban, public university to address the effect of financial aid on first year student persistence. Somers predicted retention using a logistic regression model with independent variables that captured the students' backgrounds, college experiences, achievements and financial aid packages. She found the student's age, grade point average (GPA) and the form of financial aid were statistically significant predictors of retention. Older students were more likely to persist, as were those with GPA's in the lower third of the class. Those receiving financial aid, especially in the form of scholarships, were less likely to remain in school.

MODEL

There are numerous variables that may explain the decision to leave college, regardless if one was asked to leave or if one chooses to leave. These independent variables play a role in describing both the ability and desirability of staying in school. Following Tinto (1993), the independent variables can be classified into three clusters. The first, pre-entry variables, are those that describe the family background, skills and abilities, and prior schooling characteristics of the student. Colleges and universities use high school performance measures and aptitude tests as indicators of academic preparedness in the admissions process. While academic readiness is perceived as a precursor for success at the college level, it may not be as strong an indicator of student persistence. Numerous students can do the work but choose to leave for other reasons.

Cope and Hannah (1975) reviewed several papers and noted the relationship between academic preparedness and persistence was not always positive and statistically significant. Family socio-economic background may help predict persistence. Children with college educated parents may be raised in an environment where graduating from college is expected. If parents and their children form a commitment to graduating, persistence is more likely.1

The second cluster contains goals and commitments on the part of the student and the institution. Students’ aspirations and expectations change over time due to college experiences. What the college has to offer the student becomes an important factor in the decision to persist. This interconnectedness between the student and the college’s goals and commitments is a central tenet of Tinto’s work. The choice of major and minor, both at the application and matriculation stages, are examples of student intentions or goals. If an institution fulfills students’ expectations regarding the course of study, persistence increases.

The third cluster, institutional experiences, also pertains to what happens academically and socially once the student arrives. Success in the classroom, as measured by grades, participation in student activities such as sports, clubs and greek organizations and financial aid provide information for this cluster.

---

1 Hackman and Dysinger (1970) studied three mid-west liberal arts schools. They built a summary measure of commitment and found a positive correlation between parents’ educations and commitment, and commitment and persistence. Astin (1975) found mixed results regarding higher parental income and persistence.
Different factors affect attrition in the first versus second year due to the dynamic process of students re-evaluating their college experiences. Though pre-entry attributes may yield predictive power for students who do not persist through the first year, these are unlikely to have a strong bearing once a student is acclimated to the college scene. For instance, academic preparedness and the high school curriculum may play a role with first year persistence, these effects may become less relevant by the sophomore year when institutional experiences play a more dominant role.

The dependent variable is a discrete variable denoting the probability of retention. If a student leaves, the dependent variable takes on the value of zero. If the student persists, it equals one.

**METHOD**

A logistic regression model is used. Using a logistic model allows for changes in the independent variables to have non-linear effects on the probability of retention. Therefore, a 100 point higher SAT score will affect the probability of retention differently if looking at SAT scores going from 900 to 1000 versus 1300 to 1400. A logistic model is also useful to discern differences in retention rates for discrete independent variables, such as gender or ethnicity, when all other student and institutional characteristics are the same. For example, one can find the difference in retention rates for students who are alike in all values of the independent variables except gender or race.

**DATA**

The data span four years with the annual number of observations varying by the size of the entering classes of 1991 through 1994. The total number of observations is 1,204, although missing data reduce the sample sizes in the models. Virtually all students are full-time and matriculate immediately after high school, therefore only full-time students are used in the data set. Students declare majors no later than the end of sophomore year, though most declare at the end of freshmen year. The school offers need-based and merit-based financial aid. Students may also receive government funds. Students applying for aid are assessed a measure of need by the financial aid office based on income and assets relative to tuition. The difference between the need and the sum of all loans, scholarships, grants and work-study funds, defines the financial net need amount. Students who do not apply for aid are given a need value of zero.

Students who leave sometime during freshmen (sophomore) year and prior to the start of sophomore (junior) year and never return to the institution are given a first year (sophomore) retention value of zero. If they remain in school, the value is one. Ninety two percent of students persist through the freshmen year. Of those remaining, 88% persist through the sophomore year.

**LOGISTIC REGRESSION RESULTS FOR FIRST YEAR RETENTION AND SOME IMPLICATIONS**

Table I gives the results for two models. The models differ due to the choice of independent variables.\(^2\) Model Two incorporates high school sizes and class ranks. The sample

\(^2\)Only seven students left prior to the beginning of their second semester so changes in GPA between the first and second semesters are used as an independent variable.
size is significantly smaller in Model Two than in Model One because not all private schools publish class ranks or class sizes.

The parameter estimates indicate the direction and level of significance for the effects of independent variables on persistence. Only statistically significant results are discussed. Predicting the marginal effects on the probability of persistence, which are non-linear, will be done after examining the significant factors.

Table I: Logistic Regression Results for First Year Retention

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter Estimate</th>
<th>Parameter Estimate</th>
<th>(Male=1, Female=0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>0.5694**</td>
<td>0.7416**</td>
<td>(African American or Hispanic=1)</td>
</tr>
<tr>
<td>BLACK/HISPANIC</td>
<td>-0.0892</td>
<td>0.1919</td>
<td>(Private School =1, Public=0)</td>
</tr>
<tr>
<td>PRIVATE</td>
<td>-0.0877</td>
<td>-0.0459</td>
<td>(Parochial School =1, Public=0)</td>
</tr>
<tr>
<td>PAROCH</td>
<td>0.6163*</td>
<td>0.6513*</td>
<td>(Math SAT score)</td>
</tr>
<tr>
<td>SATMATH</td>
<td>-0.00290*</td>
<td>-0.00303</td>
<td>(Verbal SAT score)</td>
</tr>
<tr>
<td>SATVERBAL</td>
<td>0.00111</td>
<td>0.000716</td>
<td>(Social Science Intended Major=1, Humanities=0)</td>
</tr>
<tr>
<td>SOCS CIENCE</td>
<td>0.7118</td>
<td>0.5679</td>
<td>(Science Intended Major=1, Humanities=0)</td>
</tr>
<tr>
<td>SCIENCE</td>
<td>0.4671</td>
<td>0.3069</td>
<td>(Undecided Intended Major=1, Humanities=0)</td>
</tr>
<tr>
<td>UNDECIDED</td>
<td>0.0835</td>
<td>-0.1725</td>
<td>(First Year Need-Award)</td>
</tr>
<tr>
<td>FROSHNEED</td>
<td>-0.000024</td>
<td>-0.000017</td>
<td>(First Year GPA)</td>
</tr>
<tr>
<td>FROSHGPA</td>
<td>0.0149**</td>
<td>0.0169**</td>
<td>(Spring-Fall Semester GPA)</td>
</tr>
<tr>
<td>CHANGEGPAI</td>
<td>0.0115**</td>
<td>0.00823**</td>
<td>(HS Graduating Class&gt;342)</td>
</tr>
<tr>
<td>LARGE HS</td>
<td>0.9041**</td>
<td>0.0876</td>
<td>(Class Rank &gt; 94th percentile)</td>
</tr>
</tbody>
</table>

***Level of significance<.05    **Level of Significance<.10    *Level of Significance<.15

* Likelihood-Ratio Test Significant at .0001
Model One indicates higher first year GPA’s and improvements in first year grades between the first and second semester significantly increase retention. Parochial school students are more likely to persist. Higher math SAT scores reduce retention, whereas verbal SAT scores are not significant. Males are more likely to persist.

Model Two indicates the level of significance and size of the parameter estimates for first year grades, changes in them, parochial school students and gender are similar to those in the Model One. In addition, students coming from large high schools are more likely to persist. Ethnicity, SAT scores, unmet need and intended majors are not significant indicators of retention.

Using Model Two, the predicted effects on the probability of retention is evaluated at the means of the independent variables. Although any combination of values on independent variables can be used, a few examples are given. Observing a non-black or Hispanic female from a public high school with less than 342 in the graduating class and average SAT math and verbal scores, while intending to major in the sciences instead of humanities, yields a 93.68% probability of retention. Given identical characteristics except gender, we find a 96.89% predicted retention rate. If the same female had attended a large high school, her predicted persistence rate would be 97.34%. For an increase in the average female’s first year GPA from 2.8 to 2.9, her predicted retention increases by 1%. If her GPA falls from 2.8 to 2.7 between the first and second semesters, her retention rate falls by .5%.

LOGISTIC REGRESSION RESULTS FOR SOPHOMORE RETENTION AND SOME IMPLICATIONS

The determinants of first year and second year retention can be very different. High school experience factors wane after the first year; SAT scores, high school ranks and sizes, etc. are not part of the decision-making process that occurs in the sophomore year. Likewise, first year financial aid decisions are no longer relevant, whereas those in sophomore year are. Given the significant amount of time spent on campus by sophomore year, students have explored social and academic niches. The current decision to persist resides with the announced major, not the intended one from the admissions application.

Table II gives the regression results for sophomore year retention. High school variables are dropped and greek organization membership is included, since membership can occur beginning in the sophomore year. The difference in unmet need between the sophomore and freshmen year is included. Declared majors are used in place of intended majors. Gender and ethnicity remain in the model. The richness of the model is apparent with the numerous highly statistically significant results.

Not surprisingly, higher GPA’s and improvements in GPA’s between semesters increase the likelihood of sophomore retention. If the gap in unmet need increases, students are less likely to persist. Using humanities majors as the base, students who declare either social sciences or sciences relative to humanities are more likely to persist. Males are more likely to remain through the sophomore year. The only currently available measure of social connectedness is greek-organization membership. Not only is such membership highly significant, but the size of the coefficient is very large.

Model Three indicates that a non-black or Hispanic female, science major, who is not a member of a sorority, shows an 87.5% probability of persisting. Even decreasing her unmet need, i.e., over-awarding her by $1000 from first to second year, increases her predicted retention rate by a mere .69%. Thus, it would take an approximate over-award of $1500 to increase the persistence rate by one percentage point. If the same student joins a sorority, her persistence

---

3No statistical significance occurs in models with these variables.
probability rises from 87.5% to 98.4%. The choice of major influences the predicted retention rate significantly. If the same non-sorority member declares humanities as her major instead of science, her predicted probability of remaining falls from 87.5% to 77.3%.

Table II: Logistic Regression Results for Sophomore Year

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter Estimate</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>0.7700***</td>
<td>Model Three</td>
</tr>
<tr>
<td>BLACK/HISPANIC</td>
<td>0.00456</td>
<td>Sample Size 1,097</td>
</tr>
<tr>
<td>CHANEGPA1</td>
<td>0.0106***</td>
<td>Variable</td>
</tr>
<tr>
<td>CHANEGPA2</td>
<td>0.0172***</td>
<td>Parameter</td>
</tr>
<tr>
<td>CHANEGPA3</td>
<td>0.0127***</td>
<td>Estimate</td>
</tr>
<tr>
<td>SOCSCIENCE</td>
<td>0.9070***</td>
<td>(2nd Fall Semester-1st Spring Semester GPA)</td>
</tr>
<tr>
<td>SCIENCE</td>
<td>0.7219***</td>
<td>(2nd Spring Semester-2nd Fall Semester GPA)</td>
</tr>
<tr>
<td>GREEK ORG.</td>
<td>2.1713***</td>
<td>(Social Science Major=1, Humanities=0)</td>
</tr>
<tr>
<td>NEED DIFFERENCE</td>
<td>-0.000064*</td>
<td>(Science Major=1, Humanities=0)</td>
</tr>
<tr>
<td>FROSHGPA</td>
<td>0.0191**</td>
<td>(Greek Member=1, Non-member=0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Unmet Need 2nd Year-Unmet Need 1st Year)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sophomore GPA can be interchanged here.</td>
</tr>
</tbody>
</table>

***Level of Significance<.05   **Level of Significance<.10   *Level of Significance<.15

Choosing average characteristics for all independent variables, but choosing a non-black or Hispanic male, instead of female, science major, the predicted retention rate rises from 87.5% to 94.8%. The effect of unmet need on this student is half that on the female, implying that a $3000 over-award would be necessary to increase retention by one percentage point. If the male, science major joins a fraternity, his predicted retention rate is 99.3% compared to 94.8%. As in the case of females, male humanities majors are at risk of leaving. The non-fraternity humanities major’s retention rate is only 88%.

CONCLUSION

Incorporating Tinto’s model with logistic regression analysis, this study finds that for first year students, higher GPA’s and increases in grades between semesters significantly increase retention, which is slightly different than the findings by St. John et al (1992). Schools can easily
First Year vs. Second Year Retention of College Students

monitor grades to see the effect on retention and be mindful of the consequences. Males and students from large high school also are more likely to persist. On the other hand, high school rank, SAT scores, ethnicity and financial aid are not driving first year retention at this institution. These findings, though institution specific, can easily be modeled at other schools.

Monitoring second year developments is key to retaining sophomores. Grades, changes in grades, gender, choice of major, financial aid and fraternity/sorority membership are all significant determinants of retention. Several points are noteworthy. First, ethnicity does not play a role in sophomore retention. Second, humanities majors are at risk of leaving. Bi-directional causality may be at work. It could be that humanities majors are more ambiguous about remaining in school compared to students in other disciplines and this phenomenon may be true nationwide. On the other hand, perhaps humanities majors seek to transfer to a program with more to offer humanities students, which would be an idiosyncratic result.

Third, greek-organization membership, unfortunately the only available measure of social connectedness, not only enhances retention, but also does so dramatically. The result does not necessarily imply uni-directional causation of joining a greek organization causing retention to rise. It may be true that students planning to stay decide to join a fraternity or sorority. Fourth, "buying" the sophomore class can increase retention, but it is very costly. Lastly, females are more likely to leave than males. This last result is perplexing and requires further study by breaking down the data by gender.

There are several routes for further research. Capturing additional data about students would help expand the number of variables in Tinto’s clusters. For example, athletic and club participation may prove significant. Along another vein, examining voluntary withdrawals versus dismissals would shed more insight on how Tinto’s institutional experiences affect retention.

REFERENCES


