

1-1-2018

Stay in School, Kids

Caleb Reynolds

Julia Wenndt

jwenndt16@georgefox.edu

Recommended Citation

Reynolds, Caleb and Wenndt, Julia, "Stay in School, Kids" (2018). *Student Scholarship - School of Business*. 10.
https://digitalcommons.georgefox.edu/gfsb_student/10

This Paper is brought to you for free and open access by the School of Business at Digital Commons @ George Fox University. It has been accepted for inclusion in Student Scholarship - School of Business by an authorized administrator of Digital Commons @ George Fox University. For more information, please contact arolf@georgefox.edu.

Stay in School, Kids
Caleb Reynolds and Julia Wenndt
George Fox University

Abstract

This paper presents an analysis of the effects of football wins percentage on universities' retention rates. Using an OLS model we determine the causal relationship between football wins percentages, other independent variables, and universities' retention rates. Throughout the paper we will define all relevant independent variables and present an overview of the data gathered in order to formulate our results. A theoretical framework will be presented followed by a thorough analysis of the OLS estimation results.

Keyword: Football, Retention, Education, Division 1, Sports

JEL Codes: I23, O33, Z20

Introduction

Education has always been a crucial building block for economic prosperity. As populations become higher educated, productivity greatly increases, technological advancement and innovation take place, and overall growth occurs within an economy. Education is thus a pivotal part of the economy and something that should be prioritized. In light of this, we have decided to investigate the retention rates of Division I Universities in the United States through an econometric analysis. Since there is such a high economic value on education, keeping students in school and moving towards graduation is subsequently a highly important economic good.

When looking at American university culture, something that stands out as being particularly important is athletics and especially football. American culture idolizes the game of football and around the country, Division I Universities pour millions of dollars each year into their football programs. We are therefore interested in determining if football performance has a significant effect on retention rates, important information for prospective students and school administrators alike. Football programs can bring national attention to a university as well as help connect the institution to the larger community surrounding it. Teachers, students, parents, alumni, and fans can all come together around the game of football. There is therefore an element of solidarity that football brings to a university and when a team is performing well that solidarity only increases, but is the level of solidarity and school spirit that football brings to universities enough to impact retention rates? Does winning one or two more games mean that more students will continue on with their education at that university? These are the questions we are seeking to answer in this analysis. In order to examine this, we will be using a cross sectional analysis of data to determine the effect football wins has on retention rates, using an

OLS regression and regressing university retention rates on football wins and other university statistics.

Data Overview

In order to build this model, we randomly selected 50 NCAA Division I universities. We automatically ruled out universities that did not have a football program; with the remaining schools, we assigned a number to each school and used a random number generator to randomly select the data set. To determine the independent variables, we brainstormed elements that would have a logical impact on retention rates and then compiled these variables based on what data we could find. We are stating that university retention rates are a function of the university's football wins percentage, undergraduate enrollment, in-state tuition, out of state tuition, acceptance rates, undergraduate male/female ratios, and median debt of students that completed school at the university. Listed below is a brief explanation on all of the independent variables, with values taken from the 2016 football season and the 2016-2017 school year.

| Independent Variables | Explanation |
|-------------------------------|--|
| Football Wins Percentage (W) | This percentage is calculated by dividing a teams amount of wins by their total games played |
| Undergraduate Enrollment (UE) | Total amount of undergraduate students enrolled in the school during the year in question |
| In-State Tuition (IST) | Amount of money a resident of the state in which the school resides must pay |
| Out of State Tuition (OST) | Amount of money a non-resident of the state in which the school resides must pay |

| | |
|--|---|
| Acceptance Rate (AR) | Percentage of applications the school accepts for enrollment |
| Percentage Female (PF) | Percentage of undergraduate population that is female |
| Median Debt of students that completed school (MD) | This median value is an estimate of how much debt a student may have after graduating from the university |

Within the Appendix, Table 1.0 displays the summary statistics of each variable. The dependent variable “Retention Rates” has a mean of 86% showing that on average, 86% of first year students return to their university for a second year. Further, with a standard deviation of only 8%, this is a fairly clustered sample. “Football Wins %” is a more spread out variable with a standard deviation of almost 20%. The mean value here demonstrates that on average, teams win 52% of their games. Our hypothesis is that the coefficient on this variable will be small, lacking statistical significance. There are many factors that drive a student’s decision on where to attend school; unless he is planning on playing in the NFL, a school’s football program usually isn’t a deciding factor. “Undergraduate Enrollment” is a rather spread out variable with a standard deviation of almost 12,000 students, and the coefficient on this variable will determine if retention rates differ for large or small schools. We expect this coefficient to be negative; it is typically the case that students seeking a “college experience” go to larger, state schools, while students wanting to quietly focus on their studies go to a smaller private school. Further, small schools are typically more competitive, and thus consist of higher achieving students. Thus, students at small universities are more likely to stay in school. The difference in “In-State Tuition” vs. “Out-of-State Tuition” is expected with out of state having a larger mean. However, not all of the universities in question are state schools, so there is no difference between in state

or out of state tuition for the private universities. This is why the maximum value is the same for both variables; the maximum value is coming from a private school that has a universal tuition rate. We expect the coefficients on both of these variables to be negative, because as schools become more expensive, they are harder to afford and some students are forced to drop out.

“Acceptance Rate” is a very wide variable, ranging from 5% to 96%, with an average of 56%.

We assume that the coefficient on this variable will be negative because as you lower acceptance rate, those that do get in will be less likely to leave since the school is so competitive to get into.

“% Female” is clustered around the 55% value with a standard deviation of 6%, with some outliers. For this variable we are assuming that it does not have a linear relationship, but rather that retention rates would be highest for an evenly split male/female population and would decrease as the population is unequal to either side. Therefore, we are uncertain as to what the coefficient of the “% Female” variable will be because it is pertaining to the percentage of the population that is female exclusively in a linear manor, even though we believe the relationship of this variable on retention rates to be nonlinear. The “Median debt after graduating” variable has a mean of 20,000 dollars. We would expect the coefficient on this variable to be negative, large, and statistically significant because as students accumulate more debt, they are more likely drop out if they can no longer afford school.

The data on Football Wins was gathered from the site TeamRankings, with values taken from the 2016 Fall season. For the rest of the variables, data was pulled from the Department of Education for the 2016-2017 school year. There are two things worthy of noting in regard to the gathered data: First, due to the military affiliation of the US Naval Academy, there is a strong probability of external causes for any correlation presented in the model for that school alone. However, including this observation will increase the external validity of the experiment, so we

have decided to keep the variable included in regressions. Second, median debt for graduates could not be found for three schools; as a result, these schools will not be included in any regression and our number of observations will drop to 47.

Methodology

There is the potential that due to the nature of the research topic, it may be that a panel data set could better explain the relationship between the dependent and independent variables through a fixed effect or time effect model. However, due to time constraints and available data, we have elected to use an OLS model. The model that we will be using for this analysis is the following linear regression model:

$$RR_i = \beta_0 \pm \beta_1 W \pm \beta_2 UE \pm \beta_3 IST \pm \beta_4 OST \pm \beta_5 AR \pm \beta_6 PF \pm \beta_7 MD + u_i$$

Our *a priori* hypothesis is that the coefficient on Football Wins will not be statistically different from 0. It is first important to note that there are plenty of people who do not follow football. During football games, regardless of which university you attend, there are still a large number of people studying in the library, enjoying time with friends, or relaxing at home. For these students, football has absolutely no influence on their decisions regarding whether or not to stay at their university. Next, we will consider students who do follow football. While football has a large emphasis in American culture, a student doesn't go to school for football; a student's focus is on their degree. It is common for students to transfer out of schools with high achieving football teams if the school does not offer what they need in the classroom. While football performance may be a determining factor if a student is torn between two universities, it is rare to see a non-athlete leave a school because the football team isn't performing well. Further, it is true that football is a strong component of the stereotypical "college experience." The sport is idolized, millions of dollars are poured into it every season, and entire cities rally together to

support the local university team. However, that is precisely the reason that football is so popular: football builds community on campus and gives students an opportunity to show their school spirit, regardless of the outcome of the game. Win or lose, students and community members alike can have fun supporting their school at a football game; losing does not lead to transferring, and retention rates are not dependent on the football team.

Results

Our initial regressions, (Appendix Table 2) show that when all variables are included as regressors, undergraduate enrollment, out of state tuition, acceptance rate, and male/female ratio all have strong statistical significance. The only factors with economic significance are acceptance rate and male/female ratio; all other variables have a coefficient of less than 1/2500. Further, three of these four regressions shows that our initial hypothesis is correct, since the coefficient on football wins has neither economic nor statistical significance in models 1, 3, and 4.

We believe that football wins percentage does not have a statistically significant impact on retention rates because for the majority of students, there are so many other things impacting their decision to stay enrolled at a university. It is interesting to see that in model 2 football wins percentage was found to be statistically significant to the 10% level, but this can be explained as the presence of omitted variable bias. We know this was omitted variable bias because as other variables were added the significance on football wins percentage dissipated and the other variables added were seen to be consistently significant across models 3 and 4. Although football programs do have large impacts on the communities of universities and even the surrounding communities of the towns in which they are in, they have no impact on student retention. This result then begs the question of whether or not the funds that are being put towards football

programs across the nation could be better put towards programs that would help keep students in school. As discussed earlier, keeping students in school and working towards completion is an economic good for the greater community and from this analysis it can be seen that football programs are not aiding in that good.

Some interesting results that these regressions give are that in state tuition never achieved statistical significance, but out of state always did. A logical explanation for this is the distribution of the variables: out of state tuition has a wider range than in state tuition does, and therefore the extreme values will alter a decision to attend a school out of state more than in state's extreme values will affect the decision for that school. Further, the out of state coefficient was positive in all 3 regressions it was included in. We believe that this happens due to the student's strong desire to attend that out of state school. Consider, for example, a student who travels halfway across the country to attend a school that's been their dream college for years; if that school charges a lot for out of state students to attend, then we have further confirmation that this student is very passionate about the university and will not want to transfer to a different school. Therefore, students that are paying large out of state tuitions are more likely to stay enrolled in the university than students that are paying lower in state tuition. Further, students who stay in state may be less committed to the entire college experience than someone that has left everything they know behind in order to attend a school. Those that stay in state may see college as more of a test run and if home is easy to return to, they are more likely to leave when things start to become something that that student did not ask for. We would expect a student unsure about the college experience to pick a university close to home, leading to the in-state variable having a negative coefficient.

The coefficient on acceptance rate is also one that has particular economic significance as well as statistical significance. The coefficient on this variable is -0.158 , this means that for every percent a school raises its acceptance rate they will decrease their retention rate by approximately $1/6$ th of a percentage point. This makes sense because schools that have a lower acceptance rate are seen as more prestigious and therefore if a student is accepted into one of these prestigious schools then they are less likely to want to drop out. This builds an interesting conflict for schools because often schools will want to attain high levels of prestige, but will not yet have the reputation developed. Therefore, they will not have the ability to maintain a low acceptance rate because they will need the revenue generated by bringing in more students. This high acceptance rate, however, will contribute to lower retention rates which will reflect poorly on the school and ultimately keep them from achieving their goal of prestige.

Another interesting finding within the results is that the coefficient on percentage female is negative, statistically significant, and economically significant. The coefficient on percentage female is -0.333 , this means that for every percent increase a school has in the percentage of the schools population that is female, the retention rate of that school will go down a third of a percentage point. That is a fairly large impact on retention rates for the small one percent change of the school's female population. So what could be causing this to be the case? One hypothesis we have are that it is possible that women drop out of college at a higher rate than men because they are starting a family or get married and decide to not finish school. This would mean that schools that have a higher percentage of their student body being female would experience lower retention rates than a school that has a lower percentage of their student body being female. This is merely a hypothesis however, as our research does not have the scope to definitively make a

claim as to what the model is picking up here. Ultimately, this is an interesting find but not one that should inform policy.

Limitations

One potential source of error within this experiment occurs with the inclusion of the Naval Academy as a data point. This school has a special case for some variables: due to its military affiliation, tuition is free for all students in return for their service and it is far more uncommon to drop out of the Naval Academy than a public state university. However, we believe that these special circumstances do not jeopardize the assumptions that OLS runs on. The Naval Academy meets the criteria of a DI university, and its football program makes it eligible for this study. Not including it in the study would raise concerns of a nonrandom sample: by only selecting schools that have nice variable values, you are creating a nonrandom sample and are no longer running regressions that will give you externally valid results. Including the Naval Academy in our study, despite its uniqueness, strengthened our results. After conducting this research, some questions that were raised are: what effect does a football program have on a university, and whether a time series model would have better represented this study? Football programs clearly play a large role on college campuses, so even though they have no significant impact on retention rates they will have impact on something. I think further research could investigate a different dependent variable and use a similar model to try and discover what impact football actually has on university campuses. Also, after working through this research we have wondered if time series data would have told a different story, and potentially have been a stronger model. This would be an interesting train of thought to investigate in the future by gathering the same variables that we have gathered except over an expanse of time and then running the model again to see if the results are the same.

Conclusion

Using a linear OLS regression to study the impact of Division 1 universities football programs on universities retention rates, we have confirmed our hypothesis and shown that a football team's win percentage has no statistical significance in determining retention rates. Football is a sport that is idolized within American culture, but even though it is given so much attention it is not something that is impacting the decisions of students to stay in school. The strongest factors that are influencing retention rates are the variables out of state tuition and acceptance rate; as students are willing to travel far and pay a lot to attend college, they are less likely to drop out. Also if a school is seen as extremely competitive, meaning that they have a lower acceptance rate, then students will be less likely to drop out. With an adjusted R^2 of 0.702, this model is fairly strong in explaining the causation of retention rates, however, there is obviously more room for explanation. Within that missing 0.3 there could be a variable that has enormous impact on retention rates, but due to the consistency of our coefficients I trust the robustness of this model. We believe that our model is internally valid because it was built upon the basis of random selection and we gathered data from the Department of Education. We also believe our model to have external validity because it presents a logical framework regarding retention rates and its results could be applied to the larger population of universities, given that they match the population standard of being a division 1 university. When studying other D1 universities, we would expect to find that their retention rates are independent of the football team's performance, but are affected by undergraduate enrollment, out of state tuition, acceptance rate, and male/female ratios. If universities wanted to raise retention rates, they should raise enrollment, raise out of state tuition, lower acceptance rate, or admit less females. However, since all of these actions except enrollment could result in excluding some individuals

from a college education, maximizing retention rates might not be the best way to achieve the economic good that education provides after all.

Appendix

Table 1.0: Summary Statistics

| | Retention Rate | Football Wins % | Undergrad Enrollment | In-State Tuition | Out of State Tuition | Acceptance Rate | % Female | Median Debt after graduating |
|----------|----------------|-----------------|----------------------|------------------|----------------------|-----------------|----------|------------------------------|
| Mean | 0.86 | 52.36 | 23787.22 | 15727.18 | 28512.54 | 0.56 | 0.55 | 19982 |
| Median | 0.88 | 53.8 | 23118 | 10270 | 25851 | 0.58 | 0.55 | 20250 |
| Min. | 0.68 | 16.7 | 4525 | 0 | 0 | 0.05 | 0.28 | 0 |
| Max. | 0.98 | 93.3 | 55113 | 51010 | 51010 | 0.96 | 0.64 | 27000 |
| St. Dev. | 0.08 | 19.93 | 11911.57 | 13717 | 11398 | 0.25 | 0.06 | 4585 |

Table 2.0: OLS regression results.

| Dependent Variable: Retention Rates | | | | |
|-------------------------------------|----------------------|---------------------------|---------------------------|----------------------------|
| Regressor | Model 1 | Model 2 | Model 3 | Model 4 |
| Football Wins | 0.0005 (0.0005) | 0.0007 * (0.0004) | 0.0004 (0.0003) | 0.0004 (0.0004) |
| Undergrad Enrollment | | | -6.24e-08 (1.21e-06) | 1.81e-06 ** (7.98e-07) |
| In State Tuition | | -8.73e-07 (1.07e-06) | -1.50e-06 (1.03e-06) | -1.03e-06 (7.89e-07) |
| Out of State Tuition | | 5.03e-06 ** (2.07e-06) | 3.17e-06 ** (1.29e-06) | 4.02e-06 *** (1.04e-06) |
| Acceptance Rate | | | -0.198 *** (0.052) | -0.158 *** (0.048) |
| % Female | | | | -0.333 ** (0.145) |
| Median Debt | | | | -1.30e-06 (2.42e-06) |
| Intercept | 0.833 *** (0.032) | 0.693 *** (0.050) | 0.882 *** (0.087) | 1.004 *** (0.083) |
| Adjusted R ² | -0.005 | 0.307 | 0.534 | 0.702 |
| Notes: n = 47 | | | | |

References

TeamRankings. (n.d.). College Football Team Win Trends - All Games, 2016. Retrieved

November 1, 2018, from

https://www.teamrankings.com/ncf/trends/win_trends/?range=yearly_2016

College Scorecard. (2017, September 29). Retrieved November 1, 2018, from

<https://catalog.data.gov/dataset/college-scorecard>