Tyler Terbrusch

ECON 5495

Professor Kim

Final Project

Analysis of NCAA’s Academic Progress Rate Scoring System

Collegiate athletics play an important role on the campuses of United States colleges and universities. Rooting for these athletic teams supply opportunities for students to show school pride and for the community to come together. Collegiate athletics also provide schools with areas to generate revenue from ticket sales, selling merchandise, and broadcasting deals. Collegiate athletics also equip thousands of student-athletes with the opportunity to attend an institution of higher education at a reduced cost or at no cost altogether through scholarships while also being a member of one of the school’s varsity athletic teams. While the topic of fairly compensating these student-athletes with paychecks on top of their scholarships is a hotly debated one (“Should NCAA”), this report will not focus on that issue. This analysis instead examines the National Collegiate Athletic Association’s (NCAA) Academic Progress Rate (APR) scoring system that was instituted in August of 2003. In response to concerns that student-athletes were not being prepared adequately for post-education life or being held to the same academic standard as their non-athlete peers, the NCAA installed the APR scoring system. This system assigns scores to NCAA teams to hold them accountable to make sure their student-athletes are succeeding in their schoolwork. This analysis will examine the effects that certain characteristics of schools, sports seasons, and student-athletes have on their academic outcomes and their teams’ APR scores and whether this policy was effective in raising the academic performance of student-athletes. This report will identify where statistically significant differences exist and will recommend areas of need to even out academic outcomes so that all student-athletes can be set up for success.

Before the analysis of the APR Score policy, some more background on this rule will be provided. The APR is a measure of how well student-athletes are progressing through their academic coursework. This measure is calculated by taking the total number of student athletes on a team who are receiving athletically awarded financial aid, and awarding one point for each student who is still in school from the previous year, and another point to the student if he or she is academically eligible to participate that year. The teams’ number of points accumulated divided by their total number of points possible multiplied by 1000 gives them their APR score (“Academic Progress Rate Explained”). The APR was put into place before the start of the 2003-2004 academic and collegiate athletic years. While scores were kept track of over the next 3 years, sanctions and penalties were not administered until after the end of the 2007-2008 school year, which allowed for an entire cohort of student-athletes to matriculate through their four years of college. From that point on, NCAA teams needed to maintain a four-year average APR score of 930 in order to be eligible for postseason tournaments. Teams that repeatedly fall below the 930 APR score mark will also be subject to other penalties that include diminished scholarships available to award and restrictions on competition (“Academic Progress Rate Timeline”). Many schools look to their teams to participate in competitions, tournaments, and championships to provide not just a monetary boost to the school, but also to provide an indirect benefit of notoriety with potential applicants on a regional or even a national scale. Success by NCAA teams has been linked to increased application numbers (Glatter) (Mosendz et al.). There is a penalty worth avoiding that comes with not meeting the 930 APR score, and these penalties were created as incentives for schools to raise their standards for student-athletes. In the following section, the success of this policy on raising student-athlete academic outcomes will be analyzed.

The dataset being used for the following analysis is comprised of all Division I athletic teams and their APR Scores, their rates of Eligibility, and rates of Retention for every year since 2004, the year the APR policy was implemented. This dataset has 6511 observations from teams across 38 sports and from 385 Division I schools (“Academic Scores”). Also of note, there are 48 conferences of schools represented in this dataset (“NCAA Conferences”). All data analysis was done using R. Commands, code, and explanations on different codes and processes can be found in the accompanying code file. The following graphs demonstrate how student-athlete academic outcomes have changed over time since the implementation of the APR score in 2004. As we can see from the following graph that measures APR scores over time, average APR scores have risen consistently since their introduction in 2004. The punishments for not meeting the 930 score benchmark which became effective in 2008 had an easily seen effect as average APR scores across the entire NCAA jumped by over 6 points from 2007 to 2008.



The graphs below also show how the advent of the APR score policy had positive effects on student-athlete eligibility rates and retention rates. We see steady climbs in these rates for every year since 2004 when the policy was first enacted.





The year 2008 also acts as a threshold of when the punishments were first mediated for low APR scores. To test that APR scores, retention rates, and eligibility rates are significantly different on either side of this year, t-tests were conducted between the 2007 measures and the 2009 measures. These tests yielded results rejecting the null hypothesis that the values in 2007 were the same as in 2009. A t-test on the average number of athletes per team in 2007 and 2009 was also conducted to see if the samples in 2007 and 2009 were comparable and that t-test gave a p-value of 0.481, which is much higher than the p-value of 0.05 for statistical significance at the 95% confidence level. This high p-value means that we are confident that the average number of athletes per team is the same in 2007 and 2009 and these samples are comparable. There is a higher rate of student athletes who are maintaining their eligibility and staying in school now than when this policy was first started in 2004. It is also noteworthy that since the implementation of the APR score policy, the number of teams who have had APR scores below this 930-point benchmark that makes a team subject to punishment has fallen dramatically. In the first year that APR scores were calculated, there were 1099 athletic teams across the NCAA that had APR scores less than 930 points. In 2014, there were only 345 teams that had APR scores lower than the necessary 930 points. The graph below shows how the number of teams not meeting the 930 points has continuously dropped every year since this policy was first started.



By observing the graphs above and the results of the t-tests, it can be said that the implementation of the APR policy had a positive effect in raising student-athlete academic outcomes.

The above figures look at means across all NCAA collegiate teams and we can see that this policy has made a positive impact as on average, APR scores, retention rates, and eligibility rates have all been climbing. However, there are still areas where improvements can be made. The following sections will look at differences in student-athlete academic outcomes across observable characteristics to see where there are still discrepancies. We will first look at the academic outcomes of men’s teams versus women’s teams. In the dataset, there are 2,945 men’s teams comprised from over 66,000 athletes and 3,544 women’s teams comprised from over 61,000 athletes. Using a t-test to determine whether there are statistically significant differences in APR scores, eligibility rates, and retention rates all over the past four years, I find that teams of female athletes will have higher APR scores, rates of eligibility, and rates of retention. The average APR score for female teams (984.7) was over 11 points higher than the average for male teams (973.0). I will now turn to look at the differences in academic outcomes between “revenue sports” and “non-revenue sports.” There are often three collegiate teams at universities that bring in substantial revenue to the school through ticket sales, merchandise sales, and money from contracts on broadcasting the games on television. These sports are football and men’s and women’s basketball. In testing the difference of academic outcomes between teams that are in this revenue sport category and teams that are not considered a part of this group, it was found that non-revenue sports were found to have significantly higher averages of APR scores, retention rates, and eligibility rates. Non-revenue sports have an average APR score over the past four years of 981.8 while revenue sports have average scores of 965.4. While these average scores are above the 930-point threshold to avoid punishment, it is still found that revenue sports and male teams have lower rates of eligibility and retention meaning there are some athletes on these teams who are having negative academic outcomes. As an administrator looking at these differences, I would try to direct more academic resources to these teams so that male teams and revenue sports can raise their eligibility and retention rates.

I was also interested in examining if there were any effects of when the sport is in season on academic outcomes. Conducting pairwise t-tests between sports who compete only in the fall semester, only in the spring semester, and the sports who compete during the “winter season” which spans both semesters (Burrell), I found that there were significant differences in academic outcomes. From pairwise t-test results using Bonferroni adjusted p-values, it was determined that APR scores and eligibility rates of sports that are played in the spring are expected to be significantly higher than those measures of sports played in the fall and winter. There was not a significant difference between fall and spring sports’ retention rates, but fall and spring sports both had significantly higher measures of eligibility, retention, and APR scores than those sports played in the winter season that spans both semesters. Winter season sports, which begin in the fall semester and end in the spring semester, have significantly lower academic measures than their counterparts who play their entire season in one semester. This indicates that student-athletes who are playing these winter sports are negatively affected by having the rigors of their season affect two semesters and do not have a semester where they do not need to practice, play, and travel to games. These results show winter sports teams need to have more academic resources and options, such as summer classes, in order to raise their academic outcomes because both semesters are affected by having their sport in season. The differences in APR scores across sports by season can be seen in the boxplot below.



To finish this section on t-tests, the differences between public and private schools were looked at and it was found that private schools had significantly higher academic measures than the outcomes of public schools. Private schools’ teams averaged APR scores close to ten points higher than the scores of teams from public schools. While this is an expected result, I was unsure what the magnitude of this difference would be. The above t-test results were very informative on the differences that exist between academic outcomes on measurable and observable characteristics. These results will be used to build regression models that will be discussed in the following section. These regression models will help show how there are disparities that exist between groups of schools and the academic outcomes for their student-athletes.

One of the largest topics of discussion in college athletics that influences schools’ administrators’ decisions and education policies is that of which other schools to align with. In college athletics, teams join conferences of similar schools that share certain characteristics, academic standings, and geography. However, a recent trend in collegiate athletics has been for schools to change conferences trying to climb the ladder and join conferences that will provide more lucrative payouts in television and media contracts. While schools look to escalate their standing in the sports landscape by jumping to new conferences, this movement is not always what is best for the student-athletes. Often when teams join new conferences, the geographic fit is not the same and thus during the season, teams will need to travel further from campus for longer stretches of time. An example of this is the University of Connecticut’s athletic teams. UConn used to compete in the Big East Conference that was made up of schools in New York, Rhode Island, New Jersey, and Pennsylvania. However, UConn’s teams now compete in the American Athletic Conference that has these teams routinely travel to Texas, Oklahoma, Louisiana, Kansas, and Florida. I am hypothesizing that this effect of more rigorous travel due to changing conferences will have a negative effect on APR scores, eligibility rates, and retention rates as student-athletes are away from campus and missing class for an extended period of time. To estimate this effect, I needed to create a dummy variable that equaled one if a school changed athletic conferences since 2004. From creating this indicator variable, it was shown that 128 schools out of the 385 in the dataset changed conferences and this affected 1908 teams. This creates a treatment group of 1908 observations and a control group of 4603 observations. The analysis is started with simple linear regressions regressing the latest four-year academic measures of APR score, eligibility rates, and retention rates on the dummy variable of whether a school changed conferences or not. These simple linear regressions reported results indicating that there was a statistically significant negative impact on all three of these academic success measures for schools that had changed conferences. From the simple linear regression model estimating the impact of conference change on APR scores, there was an estimated negative impact of 1.7173. This estimate shows that relative to schools that did not change conferences, schools that did had a four year APR score from the 2011 season to the 2014 season 1.7173 points lower. The p-value on this estimate was 0.00107, showing that this estimate is significant at the 99% confidence level. In addition, the estimated effects of a school changing conferences on their eligibility rates and retention rates over the past four years were also statistically significant at the 99% level both with p-values that were essentially equal to zero. The statistically significant effects estimated for the impact on eligibility rates (-0.00325) and on retention rates (-0.00347) showed that relative to schools who did not change conferences, those schools that did had eligibility and retention rates that were lower by 0.3%. While these estimated effects are small, it is important to note that there is a statistically significant impact on student-athlete academic measures when schools change the conferences that their athletic teams compete in. These models reported R-squared measures that were considerably low, but these models did exhibit that there is a relation between these factors. The next section will discuss methods to improve the explanatory power of these regression models in order to bolster confidence in these results.

The simple linear regression models were able to display that there was a negative relationship between a school changing conference and student athletes’ academic measures of APR scores, eligibility rates, and retention rates, but now pairwise bootstrapped regression will be used to estimate these possible effects. Pairwise bootstrap sampling regressions were used in order to estimate stronger results and reinforce the statistical significance found from the simple linear regressions. For each dependent variable, one thousand iterations of a pairwise bootstrap sampled simple linear regression on the conference change dummy was conducted and the estimated results reinforced confidence in the statistical significance of the impact of changing conferences on student athlete academic outcomes. For the bootstrapped regressions for APR scores, the estimated effect of changing conferences dropping expected APR scores over the latest four-year period of 1.717 points matched the result found from the simple linear regressions. More importantly, the estimated t-stat from this bootstrap model was -3.194, which exceeded the estimated 95% confidence level critical values of -2.178 and 2.017 found by the bootstrap method. Additionally, for the bootstrap model for eligibility rates, an estimated effect of a decrease in eligibility rates of 0.3% for schools that changed conferences relative to those that did not change conferences matched what was found in the simple linear regression models. The t-stat on this estimate was -4.265 that was also well outside of the range set by the bootstrapped 95% level critical values of -2.123 and 1.824. Likewise, the model for retention rates also estimated lower rates of 0.3% for schools who changed conferences relative to those who did not, again matching the estimate from the simple linear regression models. The t-stat from this model was -5.04, which was clear of the estimated critical values of -2.192 and 1.984 at the 95% level. The results from the pairwise bootstrapped regressions adjusted the critical values for confidence at the 95% level, but the t-statistics from these regressions were outside of these ranges reinforcing the statistical significance of the estimated effects of changing conference on student athletes academic measures.

In order to improve the explanatory power of regression models estimating the effect of a school changing conferences on their student athlete’s academic outcomes measured by eligibility rates, retention rates, and the NCAA standard of APR scores, covariates will be added to the models. The following models will be estimated to see how the impact of a school’s new conference affiliation has on their student athlete’s academic outcomes:

1. Four year APR Score = B0 + B1(Conference Change Indicator) + B2(Private) + B3(Sport) + B4(Athletes) + B5(Conference)
2. Four Year Eligibility Rates = B0 + B1(Conference Change Indicator) + B2(Private) + B3(Sport) + B4(Athletes) + B5(Conference)
3. Four Year Retention Rates = B0 + B1(Conference Change Indicator) + B2(Private) + B3(Sport) + B4(Athletes) + B5(Conference)
4. Four Year APR Score = B0 + B1(Conference Change Indicator) + B2(Private) + B3(Sport Season) + B4(Gender) + B5(Revenue Sport) + B6(Athletes) + B7(Conference)
5. Four Year Eligibility Rates = B0 + B1(Conference Change Indicator) + B2(Private) + B3(Sport Season) + B4(Gender) + B5(Revenue Sport) + B6(Athletes) + B7(Conference)
6. Four Year Retention Rates = B0 + B1(Conference Change Indicator) + B2(Private) + B3(Sport Season) + B4(Gender) + B5(Revenue Sport) + B6(Athletes) + B7(Conference)

There are two different types of models that are being estimated. One model uses the controls for whether a team is from a private school or not, what sport the team plays (this variable is a vector of dummy variables for each sport), the number of athletes on the team, and the conference the team plays in (this variable is a vector of dummy variables for each conference). The second model uses control variables for whether a team is from a private school or not, dummy variables for what season the sport takes place in, the gender of the team (male, female or coed), an indicator variable for whether the team plays a “high revenue earning sport,” the number of student-athletes on the team and a vector of dummy variables for each NCAA conference. The results from these regression models provide informative results that consistently show that when schools change conferences, their athletic teams will experience lower scores in measures of student-athlete academic success. Estimates on the conference change dummy variable are consistently statistically significant at the 99% confidence level across all six of the above regressions. All of these estimates are negative and show that schools that have changed their conference that they compete in can expect for their athletic team’s APR scores, eligibility rates, and retention rates to decline. Looking at the other statistics that these regressions provide, it is seen that the explanatory power of these models are all fairly large with adjusted R-squared measures of all above 0.34 except for one model, regression 6 that had an adjusted R-squared of 0.2878. Looking at models 1 and 4, that have APR scores as the dependent variable, model 1 has an adjusted R-squared of 0.4139 and model 4 has an adjusted R-squared of 0.3786. These measures show that the estimated models explain 41.39% and 37.86% of the variation in APR scores respectively. All six of the estimated models have F-statistics testing the overall regression that are all greater than 40 with all p-values being very small decimals that are essentially zero showing that the variables used in the model have a statistically significant impact on explaining the variation in their dependent variables. There are also interesting results found by observing the estimated coefficients on the control variables used in these regressions. In line with the t-tests that were discussed earlier, it is found that teams from private schools will have higher measures of APR scores, eligibility rates, and retention rates. Teams that play in the winter are expected to perform worse on these measures as are teams that are considered “high revenue” sports like football and men’s and women’s basketball. There is some statistical significance as well that female sports perform better than men’s sports and the coed teams. These regression results are informative as to how different characteristics of a team, the school they are from, and whether or not they have changed from one conference to another that requires a more rigorous travel schedule and competition has had any effect on academic measures. These models help show that there are statistically significant differences in academic outcomes across many characteristics of an NCAA team.

This report aimed to analyze the effects that the NCAA’s policy of Academic Progress Rate scores had on the academic outcomes of student athletes across all sports. The analysis of this policy indicates that since APR scores were first calculated in 2004, student-athlete’s rates of eligibility and retention have increased while APR scores for these teams have continued to increase as well. This policy seems to have created an incentive system that has been effective in having schools maintain high academic standards for their student-athletes as only 345 athletic teams out of over 6500 had APR scores below 930 in 2014 when one sixth (1099) of the dataset had failing APR scores in the first year these scores were calculated in 2004. The creation of this policy also allowed for a measurable and quantifiable score to be given to athletic teams that can be used to observe how teams perform academically. Analyzing these measures it was determined that teams that compete in the winter season, male athletic teams, and high revenue earning sports like football and men’s and women’s basketball teams all can be expected to have lower academic measures. This is interesting from a policy perspective because these teams, especially the high revenue earning sports, are the teams that create the largest opportunity cost for the university or college when they fail to reach the 930-point threshold. These teams missing out on opportunities to play in nationally televised events cost the university exposure as well as revenue streams. Administrators at schools should attempt to provide more academic resources to these athletes. It was also shown in this report that when a school changes conferences, this has a statistically significant impact on the academic performance of student-athletes due to a myriad of factors; the one factor cited in this report being a more rigorous travel schedule. It is important for all student-athletes to succeed both in the classroom and in competition and the above analysis highlights areas where there are statistically significant differences in academic performances between groups. These differences should be addressed so that academic performance can be maximized for all student-athletes.

Works Cited

"Academic Progress Rate Explained." *NCAA*, www.ncaa.org/aboutresources/research/academic-progress-rate-explained. Accessed 20 Apr. 2018.

"Academic Progress Rate Timeline." *NCAA*, www.ncaa.org/about/resources/research/academic-progress-rate-timeline. Accessed 20 Apr. 2018.

"Academic Scores for NCAA Athletic Programs." *Kaggle*, www.kaggle.com/ncaa/academic-scores/data. Accessed 20 Apr. 2018. Raw data.

Burrell, Jackie. "Sports and Seasons of The National Collegiate Athletic Association." *ThoughtCo.*, 16 July 2017, www.thoughtco.com/sports-and-seasons-of-the-ncaa-3570382. Accessed 20 Apr. 2018.

"Division I Academic Eligibility." *NCAA*, www.ncaa.org/about/division-i-academic-eligibility. Accessed 20 Apr. 2018.

Glatter, Hayley. "The March Madness Application Bump." *The Atlantic*, 16 Mar. 2017. *The Atlantic*, www.theatlantic.com/education/archive/2017/03/the-march-madness-application-bump/519846/. Accessed 20 Apr. 2018.

Mosendz, Polly, et al. "March Madness Cinderella Stories Send Applications Soaring." *Bloomberg*, 13 Mar. 2017, www.bloomberg.com/news/features/2017-03-13/march-madness-more-students-apply-to-schools-that-break-brackets. Accessed 20 Apr. 2018.

"NCAA Conferences." *Kaggle*, www.kaggle.com/jorgan/ncaa-conferences/data. Accessed 20 Apr. 2018. Raw data.

"Should NCAA Athletes Be Paid?" *US News and World Report*, 2 Apr. 2013. *US News and World Report*, www.usnews.com/debate-club/should-ncaa-athletes-be-paid. Accessed 20 Apr. 2018