

THE BENEFITS OF COLLEGE ATHLETIC SUCCESS: AN APPLICATION OF THE PROPENSITY SCORE DESIGN

PSM Group: Ty McClure, Natalie Mikosz, Anshuman Puppala, Ethan Tabiang, and Juan Rodriguez
Texas A&M University
April 24, 2024

Abstract

In this study, Anderson investigates the claim that successful athletic programs lead to institution success as well. Anderson uses bookmaker data to estimate who is a favorite in college football and then applies a propensity score design to estimate the effects of winning on institutional health. Institutional health using variables like acceptance rate, alumni nonathletic operating donations and total alumni donations. We used Random Forest regression on Anderson’s study to better enhance the decision trees and make it a better predictive model.

Introduction

College sports is all encompassing and effects more than just each schools specific athletic program. Anderson’s study investigates how team wins change throughout the season and the propensity score depends for each win the previous game which made for significant challenges. Our code refined the implementation of propensity score to better account for confounding factors. We then extended a comparative analysis between BCS and non BCS schools to see if there was a difference. We then investigated the long term effects of athletic success’s and financial metrics of the institution.

Literature Review

Random forest regression is an ensemble learning method which utilizes multiple, randomly selected and constructed decision trees to reach a more nuanced and adaptive predictive model. Much of the literature regarding the random forest regression and its use focus on its ability to significantly reduce overfitting in a predictive model. The study “Propensity Score and Proximity Matching Using Random Forest” highlights the specific effect of using the random forest method on the calculation of propensity scores in a medical study. The study found that, while using the propensity score matching method by itself produced adequate results, the random forest addition mitigated the negative results of missing data by about 40 percent, and also improved the accuracy of the propensity score by balancing out the covariate distributions between different groups. In fact, random forest regressions are especially useful for the propensity score matching method because they can produce more robust and accurate probabilities due to the repetitiveness of their training.

Methodology

In this study, I applied machine learning techniques to explore the impact of collegiate athletic success on alumni donations. By utilizing Random Forest and Logistic Regression models, I calculated propensity scores to adjust for potential confounders, enhancing the accuracy of our causal inferences. Nearest neighbors matching was then employed to compare schools with similar profiles but varying levels of athletic success, isolating the effect of additional season wins on alumni giving. The inclusion of these ML techniques allowed for a more robust analysis, revealing nuanced insights into how changes in sports performance influence donation behavior.

Findings

Our analysis reveals compelling insights into how athletic success influences alumni generosity. Contrary to initial expectations, the propensity score matching (PSM) method suggested that an additional football win had a non-significant increase in athletic donations and a non-significant decrease in non-athletic donations, with substantial growth in overall alumni donations. However, when employing advanced machine learning (ML) techniques, we observed a smaller, yet more precise adjustment in the donation patterns: athletic donations showed a non-significant change, whereas non-athletic donations decreased significantly. Most notably, the total alumni donations were significantly positively affected, albeit the magnitude of the impact was smaller than that indicated by PSM. These nuanced results highlight the value of integrating machine learning methodologies to refine our understanding of the dynamics between collegiate sports success and alumni donation behaviors.

Table: Effects of Football Wins on Outcomes

	PSM		ML	
	Coefficient	P-value	Coefficient	P-value
Alumni Athletic Operating Donations	181.5	0.450	21.9	0.423
Alumni Nonathletic Operating Donations	-124.2	0.865	-258.8	0.014
Total Alumni Donations	3973.0	0.007	688.5	0.000

Discussion

This study investigates the causal impact of college football team success on various outcomes, including donations, applications, and academic reputation. By applying machine learning techniques, specifically Random Forest, we enhanced the robustness of our model. This choice was motivated by Random Forest’s proficiency in handling high-dimensional and complex non-linear interactions, which offers a substantial improvement over the traditional IV method used previously. Random Forest has provided a nuanced view compared to traditional Propensity Score Matching (PSM). For instance, while PSM suggested a significant increase in total alumni donations (coefficient of 3973.0, p-value=0.007), Random Forest adjustments yielded a more conservative yet statistically significant increase (coefficient of 688.5, p-value < 0.001). Interestingly, the impact on alumni nonathletic operating donations was initially deemed non-significant with PSM (coefficient of -124.2, p-value=0.865), but with Random Forest, we observed a significant decrease (coefficient of -258.8, p-value=0.014). This indicates that athletic success might lead to a reallocation of donation types, with possible reductions in non-athletic donations in favor of athletic ones.

Conclusions

In conclusion, we used Random Forrest and Logistic Regression Models in tandem with Anderson’s study on how collegiate athletic success affects non-athletic factors such as alumni and operating donations. Through our analysis of (Propensity Score Matching) PSM we didn’t find any significant increase in donations due to collegiate success, but through advanced machine learning (ML) we did find a precise pattern between success and donations. With collegiate success, non-athletic donations decreased slightly, whereas alumni donations increased significantly.

References

Anderson, Michael L. “The Benefits of College Athletic Success: An Application of the Propensity Score Design.” The Review of Economics and Statistics 93, 2011, 112-120.