Revisiting PSM Analysis of College Athletic Success with Machine Learning: LASSO Regression & Gradient Boosting

Gerard Tetegan, Jacob Lara, Daniel Pate, and Lane Whitehead
Texas A&M University
April 29, 2024

Abstract

This study PSM to determine the causal effects of college football success on donations and applications. We replicated their analysis and then used LASSO regression & (possibly) gradient boosting aiming to improve the robustness of the propensity scores.

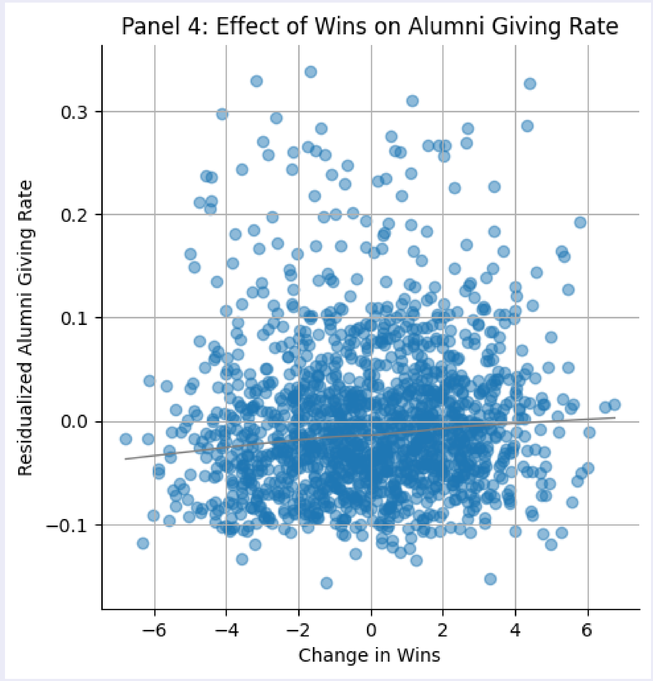
Introduction

The study uses bookmaker spreads to calculate the propensity scores and found that football success leads to better outcomes. We use LASSO and gradient boosting to aiming improve the PSM estimates and variable selection to see their effects.

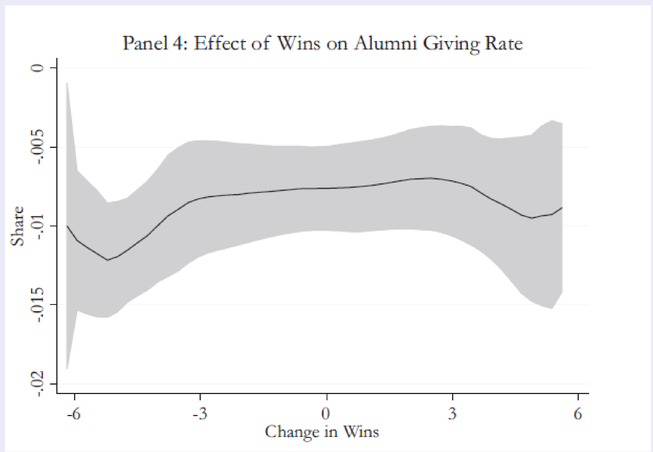
Literature Review

- Pairing PSM and gradient boosting significantly enhances predictive accuracy (Kim et al., 2023)
- Gradient boosting is more robust and preferred over other ML methods in finding average treatment effects (Yang, Chuang, Kuan, 2020)
- LASSO regressions have demonstrated their usefulness in subset selection especially in high-dimensional settings (Tibshirani, 1996).
- While LASSO is not as powerful in comparison to other ML methods, combined with PSM it demonstrates extreme robustness (Pirracchio, Petersen, van der Laan, 2015)

Findings



Figure



Figure

Machine Learning

	coef	std err	z
const	0.6200	0.030	20.782
lead2_pscore_wk10_group_1	−0.2114	0.037	−5.694
lead2_pscore_wk10_group_2	−0.0790	0.046	−1.723
lead2_pscore_wk10_group_3	−0.0775	0.053	−1.461
lead2_pscore_wk10_group_5	0.0691	0.050	1.393
lead2_pscore_wk10_group_9	0.0441	0.094	0.471
lead2_pscore_wk10_group_10	0.0313	0.075	0.416
lead2_pscore_wk10_group_12	0.1960	0.099	1.987

Figure: Lasso Regression Coefficients

While the Lasso Regression chooses fewer variables with smaller confidence intervals it is not statistically different than the original data.

Conclusions

The study finds that success in college football positively affects alumni donations, application rates, and academic reputation, while also increasing in-state enrollment and SAT scores of incoming students, especially in elite conferences. Our machine learning technique shows that with LASSO while some variables are omitted in the earlier stages it does not change the factors that are overall chosen.

Appendix