**Question 2:**

**Table 1.** Balance table for universities that were ranked (Treatment) and not ranked (Control) top 50 in 2017

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) |  |  |
|  |  |  |  |
|  | Control | Treatment | Difference |
| main |  |  |  |
| Academic.Quality | 0.515 | 0.466 | 0.049 |
| Athletic.Quality | 0.424 | 0.551 | -0.127\* |
| Near.Big.Market | 0.360 | 0.700 | -0.340\*\*\* |
| Ranked.2017 |  |  |  |
| Constant |  |  |  |
| Observations | 100 |  |  |

Table 1 displays the mean of academic quality, athletic quality, and whether the school is near a big market (indicator variable) for both schools that are ranked and not ranked as top-50 basketball in 2017. Here, we can see that the distribution is not balanced. This is especially for athletic quality and whether the school is near a big market, which differ significantly across the two groups. This motivates us to generate propensity scores.

**Question 3:**

Propensity score methods are more credible when they are generated based on the variables which assignment was made using. Unfortunately, this isn’t the case. There are way more variables (unobserved in this dataset) that determined whether the school managed to rank top 50. As such, using propensity score is unlikely to eliminate selection bias.

**Question 4:**

**Table 2.** logistic regression predicting logged likelihood of making top 50 in 2017

|  |  |
| --- | --- |
|  | Ranked.2017 |
| Ranked.2017 |  |
| Academic.Quality | -.88 |
|  | (.78) |
| Athletic.Quality | 2\*\* |
|  | (.81) |
| Near.Big.Market | 1.6\*\*\* |
|  | (.46) |
| Constant | -1.4\*\* |
|  | (.65) |
| Observations | 100 |
| *R*2 |  |

Standard errors in parentheses

\* p<0.1, \*\* p<0.05, \*\*\* p<0.01

Table 2 shows the output of the logistic regression model that examines the effect of the three covariates on the university’s likelihood of making to top 50. The model suggests that athletic quality and whether the university is near a big market both significantly predict their odds of making top 50.

**Question 5:**

Chart

Description automatically generated

**Figure 1.** Distribution of propensity score for universities that were and were not ranked top 50

Most of the data points overlapped in their propensity score. Those that lie beyond the boundaries of the overlap (n = 22) were excluded from the final analysis.

**Question 7:**

**Table 3.** OLS model on the effect of ranking top 50 on alumni donation. Fixed effects for blocks were included in the model but omitted in this table.

|  |  |
| --- | --- |
|  | Alumni.Donations.2018 |
| Ranked.2017 | 500\*\*\* |
|  | (.26) |
| Academic.Quality | 102\*\*\* |
|  | (2.4) |
| Athletic.Quality | 46\*\*\* |
|  | (5.2) |
| Near.Big.Market | 997\*\*\* |
|  | (4.2) |
| Constant | 2.7 |
|  | (3.5) |
| Observations | 78 |
| *R*2 | 1.000 |

Standard errors in parentheses

\* p<0.1, \*\* p<0.05, \*\*\* p<0.01

From Table 3, we can see that after controlling for academic quality, athletic quality, near big market, and propensity scores generated using these three covariates (fixed effect, omitted from Table 3), Being ranked top 50 still has a significant positive effect on alumni donation. This provides some moderate support for a positive association between ranked top 50 and alumni donation in the following year. However, I am hesitant to make causal inferences due to the vast number of potential confounders that are omitted from the analysis.

I do notice, however, that the model has an R-squared value of 1 (I assume this is attributed to the data simulation process). This adds more confidence to the causal inference, since it suggests that all variances have been accounted for using these variables. Unless there exist variables that perfectly aligns with the treatment, I would say that there is ample evidence supporting the causal claim.