Blue Blood Basketball

2021-12-22

Introduction and Data

Background

Historically, college basketball has been dominated by a select group of universities which are often referred to as Blue Blood schools. They are as follows: Duke University, the University of Kentucky, Kansas University, the University of North Carolina, and the University of California - Los Angeles. Many players from these universities have gone on to be stars at the professional level. Ever since the "one and done" era of modern college and professional basketball, where the most elite players spend only one year in college before entering the NBA draft, Blue Blood schools have become NBA pipelines. But are the players from these schools really the best? We are looking to investigate whether or not these schools live up to their reputation of producing elite NBA talent.

Explanation of the data set

We will be looking at players that have been drafted into the NBA from 2000-2016. The data was collected by Basketball Reference year by year as new players were drafted and as they played. We marked 2000 as the beginning of the modern era in basketball and therefore chose this year to begin our analysis. We chose 2016 as a cutoff to ensure that the players we analyzed had at least 5 years of playing data to analyze, thus getting an accurate picture of their career success. It is important to note that undrafted players are not included in this data set, and will therefore not be a part of our analysis. The observations of this data frame include what colleges the players were drafted from, the year they were drafted, and the pick they were drafted in, as well as their NBA career and per game stats, and other more complex stats that measure how much a player contributed to his team's success. Finally, we created a variable "Blue_Blood" that we used to do a separate analysis for players coming from Blue Blood school and those who do not.

Question and Hypothesis

Research Question: Do NBA players drafted from Blue Blood schools between 2000 and 2016 perform better at the professional level than other players drafted alongside them (whether they attended another university or did not attend an NCAA Division I institution)?

Hypothesis: Players from Blue Blood schools have performed better in the NBA than other players who came into the league during the same time-frame.

Methodology

First, we will compare Blue Blood players to other draftees based on basic box score statistics such as average points, rebounds, assists, and minutes per game. Afterwards, we will use all-encompassing advanced analytics stats to measure the impact and success of these players. Then, we will attempt to conclude whether NBA career success is dependent or independent of Blue Blood status.

Individual Performance Statistics

We'll compare Blue Blood players to other players in minutes played per game (MPG), points scored per game (PPG), assists made per game (APG), and rebounds caught per game (RPG). On average, draftees from Blue Blood schools get more playing time which shows how important it is for their team to have them on the court. Having an average playing time of nearly 20 minutes suggests that these players play around 42% of a game, while other players play around 31% of a game. This difference is also seen in PPG, APG, and RPG with Blue Blood players scoring 43% more, assisting 65% more, and rebounding 34% more (Table 2 & Graph 1).

Contribution to Team Success Using Analytics

So far, the basic and traditional statistics related to the box score have been analyzed to compare the success of a university's draftees in the NBA. However, the prevalence of analytics has increased in determining players' value. We will look at value over replacement player (VORP) per player, box plus-minus (BPM) per player, and Win Shares (WS) per player for each university. We will filter for schools that have had more than 10 players drafted since 2000 so we avoid schools with extremely small sample sizes. (For example, without this filter, Davidson is far and away the best college in terms of win shares per player, as its only drafted player in this time period happens to be Stephen Curry).

First, we will look at value over replacement player (VORP), which measures the value that a player provides to a team over a hypothetical replacement player on minimal salary. Three of the top ten schools are Blue Bloods. UCLA, Kentucky, and Duke in particular are the Blue Bloods who have produced players with a very high VORP. UNC and Kansas missed the top ten with a VORP of 3.5 and 2 respectively. To put these values in perspective, the statistic is designed so that the average bench player has a VORP of zero. Thus, all five Blue Bloods have players that make a positive impact of at least being better than the average bench player with UCLA, Kentucky, and Duke producing players that have starting or even star caliber impact (Graph 2).

Next, we will look at box plus-minus, which measures how many points better a team is per hundred possessions with that particular player on the court. In general, the zero mark is noted as being an average player while a score of less than negative two is viewed as having a real negative impact on the court. It is important to note that even the tops schools have a negative BPM per player. This is because the majority of players in the NBA have a BPM below 0. As shown, four of the top ten schools are Blue Bloods with the other with UCLA, Kentucky, Duke, and UNC making the top ten. Kansas is the only Blue Blood to narrowly miss out with a box plus-minus of approximately -2.1. A box plus-minus of negative two is analogous to what a score of zero is for value over replacement player. Thus, UNC, Duke, Kentucky, and UCLA are producing players with a general positive impact on their team (Graph 3).

Now we will look at win shares per player. We believe this statistic is our best look at which players have the best NBA careers because this stat accumulates throughout the player's career. BPM and VORP are efficiency-based stats, so a player who plays limited minutes but performs well can have a very high BPM and VORP. Win Shares takes into account efficiency while also taking into account volume of production and also longevity of sustained performance. Four of the Blue Bloods are in the top 10 with only Kansas not making the rankings, as their shares per player sits at just 15.13. Despite some of the fluctuation in these rankings when comparing different advanced statistics, it is still evident that Blue Blood schools are some of the best NBA pipelines in terms of quality (Graph 4).

The mean shares per player overall for the NBA players drafted since 2000 is 15.6919881. All Blue Blood schools except Kansas clock in over this number. However, the standard deviation of 25.0972087 win shares per player is extremely high due to a heavy right skew and the immense number of potential factors that determine a players NBA success. This high level of variability means there must be a drastic difference in Blue Blood win shares and those of other players to prove statistical significance.

Finally, in a direct comparison of the mean win shares from Blue Blood school players to those who are not we see that the average player from a Blue Blood school has a better career than other drafted players

(Table 3).

Results

Hypothesis Test

To make our final conclusion, we will perform a two sample hypothesis test in order to determine whether NBA success in terms of career Win Shares is independent of Blue Blood status. We know that the mean win shares of Blue Blood players is higher, but this test will give us the probability of this difference occurring by chance (meaning the variables are independent).

Null Hypothesis: Mean win shares are independent of Blue Blood status.

Alternative Hypothesis: Blue Blood Players have a higher mean win shares than other players.

Assuming the null hypothesis is true, that win shares are independent of Blue Blood status, the probability of observing a difference in means of 6.6708116 is 0.0045. Based on a significance level of 0.05, we reject the the null hypothesis since their is sufficient evidence that players from Blue Blood schools have a higher mean win share total (Graph 5).

Conclusion

In our initial exploratory data-wrangling, we found that when comparing basic box score statistics, Blue Blood players have higher per game averages in terms of points, assists, rebounds, and minutes played, indicating that the players individually perform better in any particular game. When diving into advanced analytics, we found Blue Blood schools to be some of the best at producing NBA players with superior VORP, BPM, and career win shares. Players from Blue Blood Schools have a significantly higher mean win shares total, meaning that these players have found more success in their careers, and contributed more to the teams that they have played for. Finally, we prove that this difference is statistically significant through a hypothesis test, in which we found career win shares and Blue Blood status NOT to be independent. In conclusion, we found that NBA players drafted between 2000-2016 who went to Blue Blood Schools performed better than players who went to other schools or did not play for Division 1 Colleges in terms of career success in the NBA.

A flaw in our analysis is that the dataset we used does not include undrafted players. Not including them is certainly limiting in terms of the total amount of data, and including them could have made the perception of players produced by Blue Blood schools better or worse. However, we do not believe this would change our conclusion, because most undrafted players experience limited success and the differences we see in Blue Blood players over non are extremely large. Another potential flaw in our analysis is putting our faith in Win Shares as an all-encompassing encapsulation of player success. While we do believe that this is the most accurate representation of career success we could find, it is always possible that there are things that the formula for win shares overlooks. However, we can be confident in our conclusions because all other statistics we investigated seemed to show Blue Blood superiority as well.

A motivation for potential future analysis would be comparing what picks Blue Blood and non-Blue Blood players are drafted. The picks in which Blue Blood players are drafted are concentrated at the beginning of the draft, often within the first 20 selections. This coincides with our findings that Blue Blood players are superior NBA talents, as scouts value them highly and select them accordingly. In a future analysis, one might investigate whether or not Blue Blood players exceed expectations based on where they are drafted (Graph 6).

Overall, our findings support our hypothesis that the reputation of Blue Blood schools is an accurate one, and their players perform better in the NBA than the other players drafted alongside them.

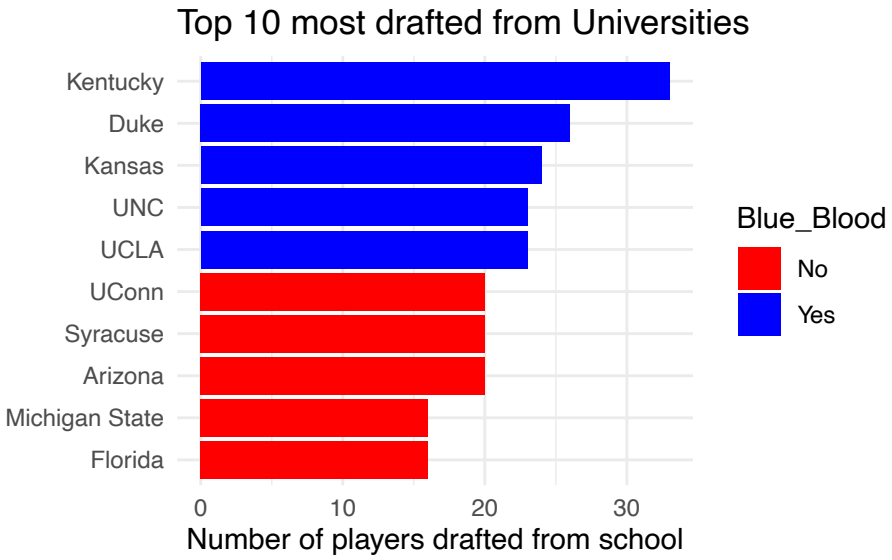
Citation: Data taken from https://www.basketball-reference.com/draft/

Apendix

The number of players drafted from each school between 2000 and 2016 proves the strong representation of Blue Blood players in the NBA. From the top 10 most drafted from universities, the five Blue Blood schools unsurprisingly take up the top five spots with Kentucky being the most drafted from with over 30 players. This is even more interesting if we take into account the proportion of total draftees from Blue Blood schools. Considering that only these 5 universities represent 12.76% of total draftees, this denotes a significant over-representation.

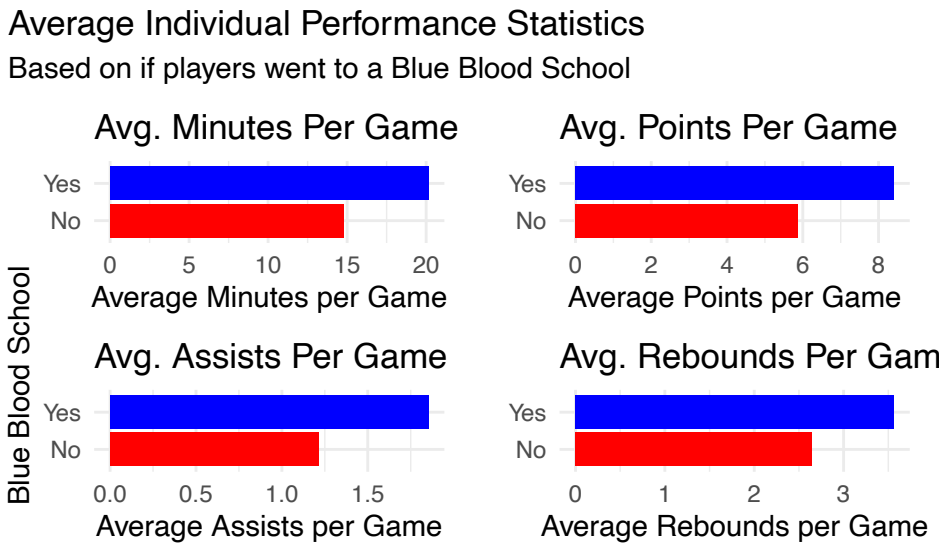
Table 1: Number of Blue Blood Players Drafted

Blue Blood	Number of Draftees	Proportion of Draftees	Percentage of Draftees
No	882	0.8724	87.2404
Yes	129	0.1276	12.7596

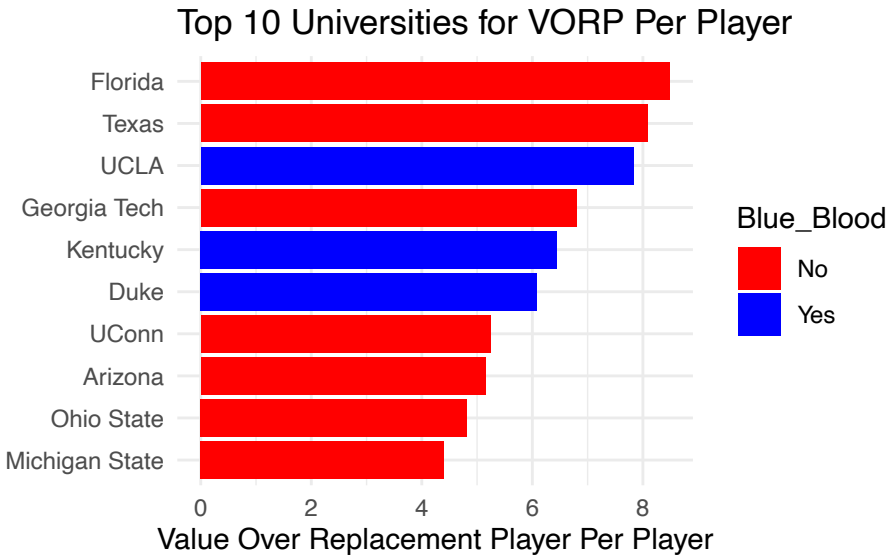


Graphs

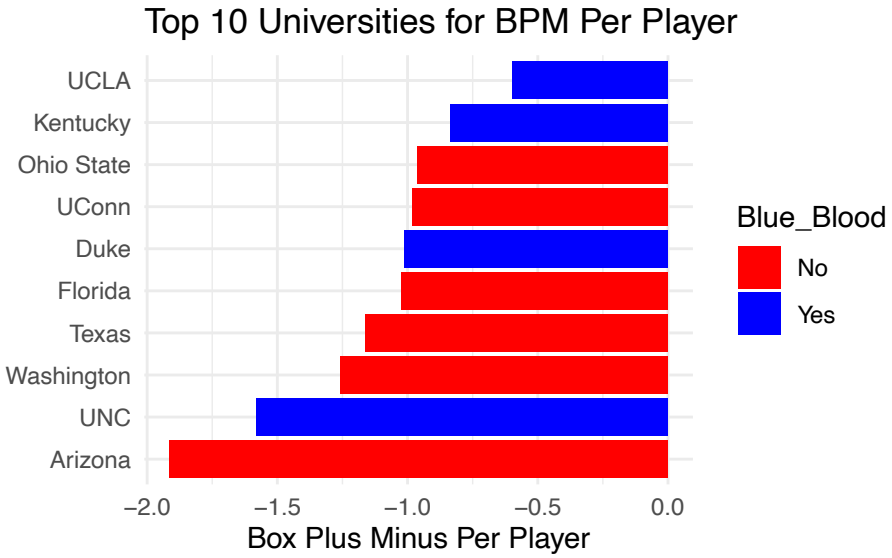
Graph 1



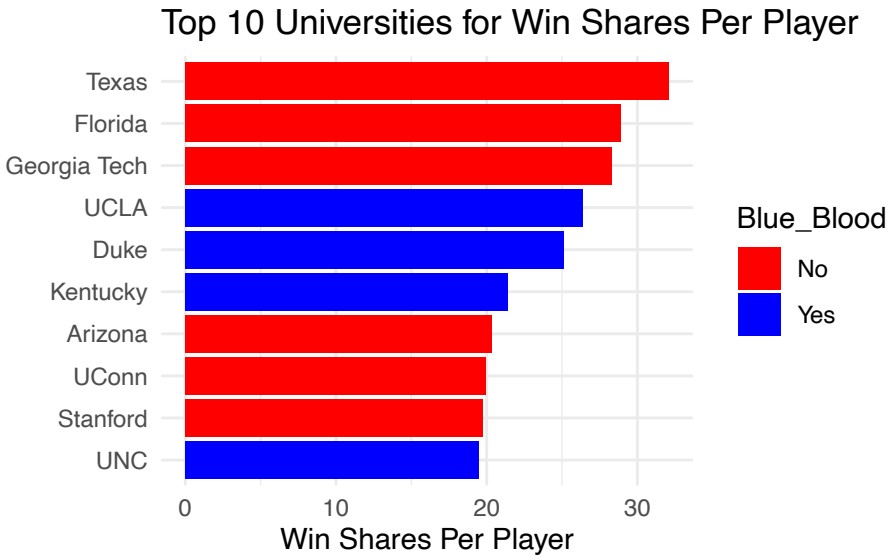
Graph 2



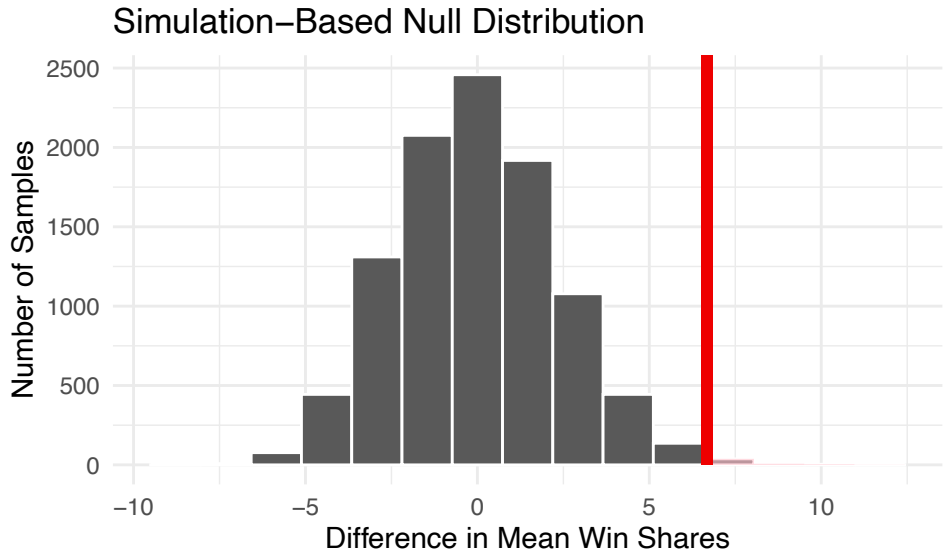
Graph 3



Graph 4

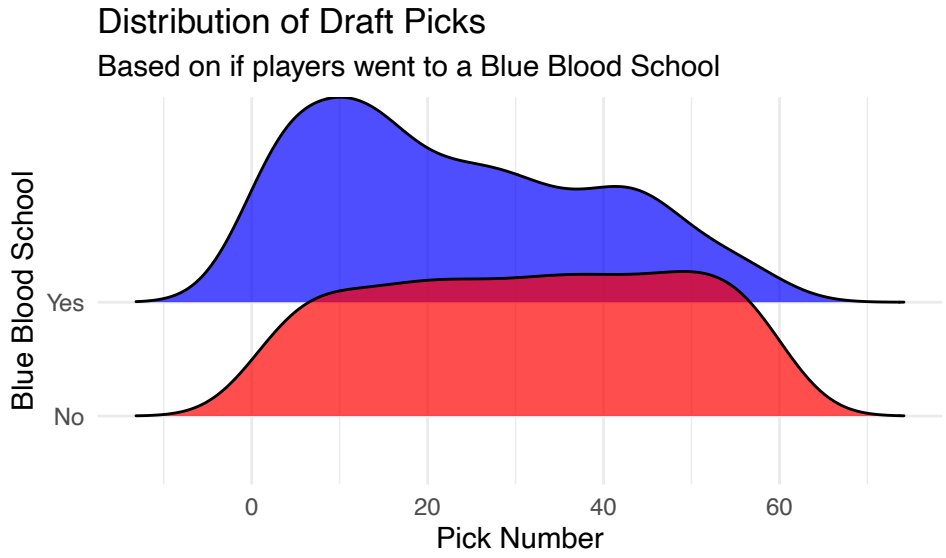


Graph 5




```
## # A tibble: 1 x 1
##   p_value
##     <dbl>
## 1  0.0045
```

Graph 6



Tables

Table 2

Table 2: Per Game Statistics of Draftees

Blue Blood	Mean MPG	Mean PPG	Mean APG	Mean RPG
No	14.7880	5.8743	1.2121	2.645
Yes	20.1473	8.4031	1.8535	3.562

Table 3

Table 3: Mean Win Shares Per Draftee

Blue Blood	Mean Win Shares
No	14.84
Yes	21.51