Name: Kyle Salgado-Gouker

Date: June 3, 2023

Class: DSC530 - Professor Metzger

Week 12

**Exploratory Data Analysis**

Statistical/Hypothetical Question: Should an NBA team deliberately lose so they can draft earlier and get a player who will contribute more future wins?

Related Question: How well can draft rank predict a player’s future win contribution?

Bonus Observation: A large payroll expense is also no guarantee of NBA success.

EDA Outcome: Significant variables aside from a player’s draft rank affect their ability to win games over their career and even during the first seasons that most benefit their drafting team.

Summary: I used two strategies. First, I compared a team’s draft power to their season performance. Drafts occur at the end of the season, so I first introduced a lag of one season and studied the correlation. While ‌a negative correlation exists between season performance and draft power—an expected outcome because draft picks are usually awarded to the worst performing teams—there is still a wild variation, mostly the outcome of trades or bad luck in the draft lottery, itself a league invention to minimize the rewards of deliberately losing. I introduced a second year of lagging as well, but observed little benefits.

Second, I used multivariate regression to determine whether the draft rank, coupled with a player’s experience, is a good predictor of their ability to win games over their career. I scraped several sites and built an SQL database that described team seasons and player seasons and then constructed a view specifically for determining a player’s winning impact, highly weighted to their postseason success. I built a model using ordinary least squares and analyzed the results and concluded both for players and the teams themselves, there are too many other variables besides draft rank. These include coaching, team strategy, and culture.

Therefore, I reject the hypothesis that teams should deliberately lose, because there are no guarantees even top picks will contribute to the drafting team’s success. In fact, there have only been 3 number one picks who have won an NBA championship with the team that drafted them: LeBron James, Kyrie Irving (one each), and Tim Duncan (five).

Other Variables: I would like to add information from sites that study draft picks as well to determine factors that might lead to better predictions. I am aware, however, that teams have access to all such data and yet still have an uneven record of draft success. Likewise, adding player contract and team revenue information may provide alternative measures of success.

Assumptions: My primary assumption is to view an NBA franchise’s success by how many wins it achieves, strongly weighted for the postseason. There are other ways of judging franchise success, including its income, valuation, and audience.

Challenges: A great deal of NBA information is proprietary. The available data needed to be organized and merged into a coherent, useful dataset.

**Acknowledgements:**

The following sites were scraped to assemble the SQL Database:

Basketball-reference.com (For player and team stats)

"https://www.basketball-reference.com/"

USA Today’s hoopshype.com (For team salary data. Much of the original data is from Patricia

Bender's site: https://www.eskimo.com/~pbender/ which is generally accepted as the most reliable source pre-2010.

"https://hoopshype.com/salaries/"

Wkipedia - for Forbes valuations of franchises.

“https://en.wikipedia.org/wiki/Forbes\_list\_of\_the\_most\_valuable\_NBA\_clubs#Historical\_valuations”

Spotrac - sports contract authority, but only after 2011.

“https://www.spotrac.com/nba/cba/”

"https://www.thehoopsgeek.com/average-nba-salary/