Identifying Championship Potential of NBA Franchises

Matt Mullins[*]

April 26, 2018

Abstract

The objective of this research experiment is to determine if every individual team in the NBA has an equal chance of winning a championship. Recognizing that there is a cycle of up and down years that all teams experience, which can be attributed to injuries, retirements, trades and other unforeseen circumstances, we will attempt to look at the long-run likelihood that every team is on equal footing when it comes to the pursuit of a championship. We will attempt to uncover these answers using both a General Linear Model (GLM), as well as a Bayesian approach, in order to get a robust set of results for which we can compare to ensure validity. Depending on our findings, it may make owners think twice as to where they franchise a team. It may make players more cautious of who they choose to sign with. And most importantly, it may make fans more aware of the type of product they are actually paying for.

---

[*]University of Oklahoma. E-mail address: matthew.p.mullins-1@ou.edu

1   Introduction

It has always seemed strange to me that for decades in professional sports some leagues have chosen to have a salary cap, while others have not. Why the disparity? Why does the NFL or the NBA feel that having a salary cap makes their league an equal playing field? And yet, the MLB has continued to operate a successful league with a more hands-off approach, a free market if-have-you. This question has become more intriguing as of late, due to the recent phenomenon of "super teams" in the NBA. As a fan, I have become less and less interested in watching a Cavaliers versus Warriors finals match up every year. Before that, it was nearly just as bad with the Spurs and Heat regularly meeting in the Championship. Is this a product of the coaching, luck, accumulation of talent, or something else? And why are these long streaks of total domination seemingly less frequent in the NFL or MLB? Yes, I will concede that there are instances of domination in the MLB (Giants) and NFL (Patriots), but there also is much more randomness in their opponents and I will objectively confess that as a fan it seems like their is more uncertainty. For those reasons, I have chosen to examine the NBA specifically in an effort to understand if the league is truly setup so that every franchise has the means to build a championship caliber team. My hypothesis is that based on the salary cap as it currently stands in the collective bargaining agreement that is scheduled to last until 2021****. That if teams are each restricted to the same amount of money that they can spend, then players will choose to live in nicer weather, major cities, where they can better enjoy their money and increase their utility. Because of this, it would seem that teams who are not as fortunate geographically will have a harder time convincing players to accept an equal paying contract to live in a less desirable location. This would theoretically put teams at a disadvantage in periods such as free-agency and allow them to acquire less "high-caliber" players

then an opponent. We would expect this to lead to less wins over the course of a season and therefore a smaller likelihood of competing for a championship.

2 Literature

1 - I will be looking first at research done on Urban Development to talk about why people choose to live in specific locations and what researchers think causes and can be done about this

2 - There is some very interesting reserach on the impacts of salary caps in sports and so I will see what effects have been noted and determine if it is having a major impact on the NBA and how research proposes it could be fixed if so.

3 - I have found some articles that talk about the ways other sports work with and without salary caps and so I will see if my findings indicate the NBA is unfair, how removing a cap all together would impact the league

4 - I am continuing to look for articles relating to free agency to look for research on the psychology and or trends of free agency in any sports league

3 Data

The primary two data sources for this research are ESPN and Statista.

ESPN is used heavily because of its accuracy in collecting in-game statistics, the unique polls that they periodically conduct such as management rankings, and the abundance of total data that they have made available publicly. Throughout this research ESPN was used the primary source for collecting information pertaining to wins and loss records for the past 5 seasons. In addition

to that, this source was also used for determining which teams were in the playoffs each year and which seed they were. Finally, ESPN provided the data for each team's Net Ranking. This was a crucial piece of data for the research and is used as the key dependent variable that we are attempting to interact with in order to uncover what other independent variables are impacting it the most. Statista is also an important source throughout this reports research. Another reliable source for NBA data, Statista was used to collect information on some of the more obscure variables including: Facebook popularity of each NBA team, average ticket prices during the regular season for each team in the past 5 seasons and total home game fan attendance for each team in each of the seasons observed.

4   Empirical Methods

The two primary methods that this report will use to conduct experiments will be the frequentist approach, a general linear model, as well as Bayesian inference methods.

The general linear model can be depicted in the following equation:

$$y = \text{beta0} + \text{beta1} * \text{bigCity[i]} + \text{beta2} * \text{teamSpending[i]} + \text{beta3} * \text{taxRate[i]}$$

$$+ \text{beta4} * \text{avgTemp[i]} + \text{beta5} * \text{managementRanking[i]} + \text{beta6} * \text{luxuryTax[i]} + \text{beta7} *$$

$$\text{milesTraveled[i]}$$

$$+ \text{beta8} * \text{attendance[i]} + \text{beta9} * \text{ticketPrice[i]} + \text{beta10} * \text{fbFans[i]} + \text{beta11} * \text{crimeIndex[i]}$$

The Bayesian analysis will consist of using a Gibbs Sampling approach. Using this technique, we will obtain samples from the posterior distribution by sampling each variables conditional distribution while holding all others fixed at their current values. This sampling will continue until the samples have the same distribution, like they are a joint distribution. Mathematically:

$$\text{Initialize: } x^{(0)} \sim q(x)$$

$$\text{for: } i = 1, 2, \ldots \text{do}$$

$$x_1^{(i)} \sim p(X_1 = x_1 | X_2 = x_2^{(i-1)}, X_3 = x_3^{(i-1)}, \ldots, X_D = x_D^{(i-1)})$$

$$x_2^{(i)} \sim p(X_2 = x_2 | X_1 = x_1^{(i-1)}, X_3 = x_3^{(i-1)}, \ldots, X_D = x_D^{(i-1)})$$

$$x_D^{(i)} \sim p(X_D = x_D | X_1 = x_1^{(i-1)}, X_2 = x_2^{(i-1)}, \ldots, X_D = x_D^{(i-1)})$$

The model that will be used for Our Gibbs Sampling Approach will be:

$$\text{netRTG}[i] \sim dnorm(\mu[i], \tau)$$

$$\mu[i] \, beta0 \, + \, beta1 * bigCity[i] \, + \, beta2 * teamSpending[i] \, + \, beta3 * taxRate[i] + \, beta4 *$$

$$avgTemp[i] \, + \, beta5 * managementRanking[i] \, + \, beta6 * luxuryTax[i] \, + \, beta7 *$$

$$milesTraveled[i] \, + \, beta8 * attendance[i] \, + \, beta9 * ticketPrice[i] \, + \, beta10 * fbFans[i] \, +$$

$$beta11 * crimeIndex[i]$$

Where Tau is the precision term:

$$\tau = \frac{1}{\mu^2}$$

For this model we will assume apriori independence and will assign low impact, stochastic priors as to avoid any subtle bias:

$$beta0 \sim dnorm(0.0, 1.0^{-6})$$

$$beta[i] \sim dnorm(0.0, 1.0^{-6})$$

$$\text{sigma} \sim dunif(0, 1000)$$

****I will continue my explanation with how many chains will be used for the Gibbs Sampler and then further explain if we will center the data, how the linear model is used, etc...

5   Findings

1 - I anticipate that many of the findings between both the bayesian and linear approach will be fairly similiar depending on how the sampling is conducted for the Bayesian approach.

2 - I am then anxious to see which of the variables is really impacting a teams netRating over the past decade. I am unsure if it will be due to weather, spending, city size, or maybe none at all... but I am sure it is going to paint a picture of what it is that helps teams compete more effectively than others.

3 - Furthermore, I am curious as to how much of netRating will be caused by the variables I have selected. Wanting this to be a fair experiment, all variables have been selected ahead of time and will be the only ones used throughout the duration of the model. Therefore, it will be cool to see if I am missing key variables, and which ones I have selected do not matter at all. Also, I will look to see any key interactions between different variables.

6   Conclusion

1 - Based on my experiment I have now discovered or not discovered that some variable or variables is/are giving a group of teams a long-term competitive advantage that other NBA clubs cannot replicate

2 - Depending on the first conclusion, I will then make a recommendation that the cap limits should not be equal across all teams and that the NBA would actually make the league more competitive by rethinking a way to allocate funds that helps teams compensate for their regional short-comings.

3 - Finally, I think depending on the findings I can then make an argument for specific cities

that teams should relocate to or cities the NBA could expand to in order to keep the league fair.

4 - I may make a suggestion that if the NBA is unwilling to make the league more competitive, then ticket prices should be more reflective of a teams chances to win. I have not fully thought out this idea. More to come.

4 - For fun, perhaps in the presentation, I will take this years NBA data and make a prediction of who is going to win the championship.

7   Charts and Data

Preliminary Graphs from observing data collected:
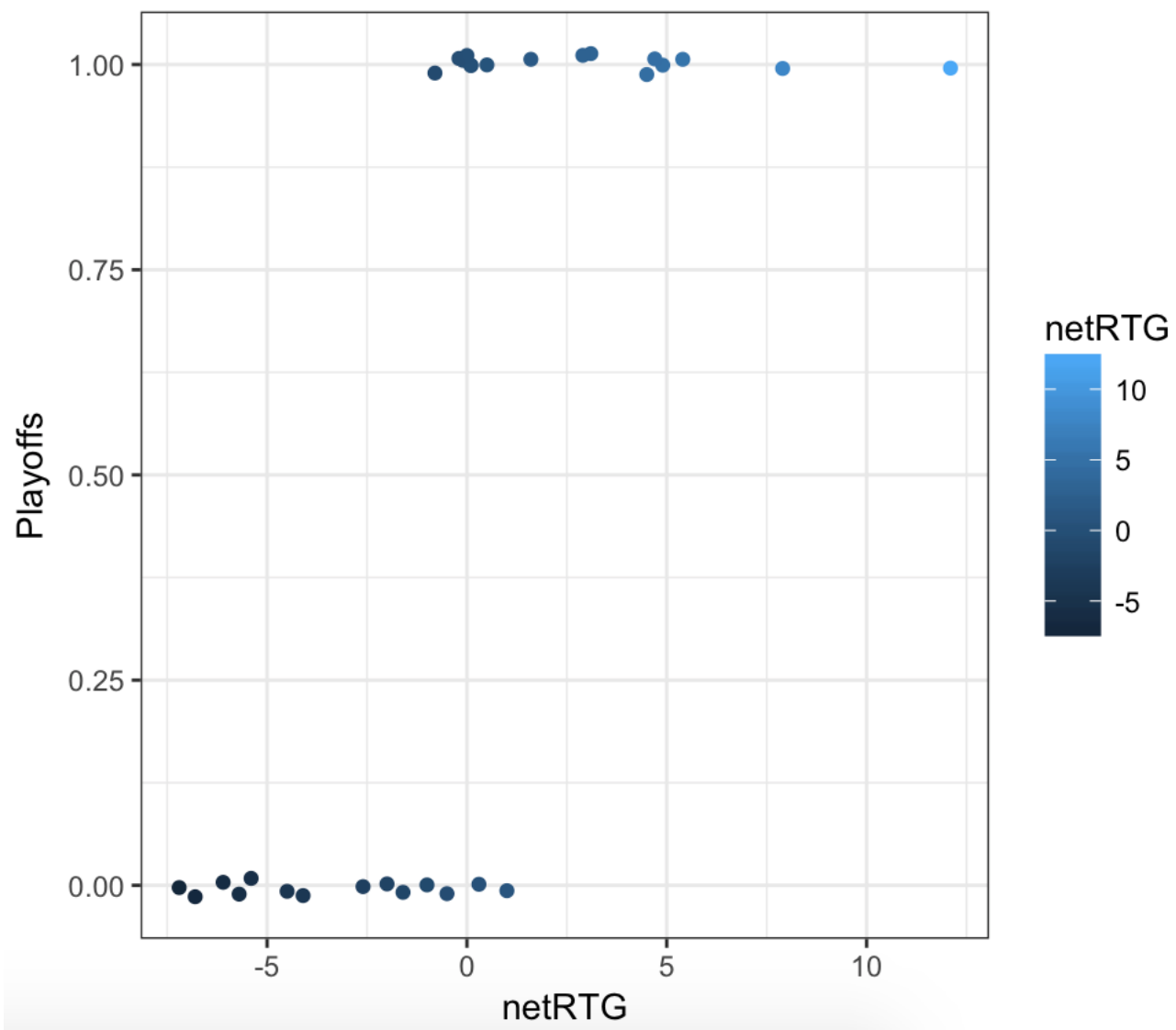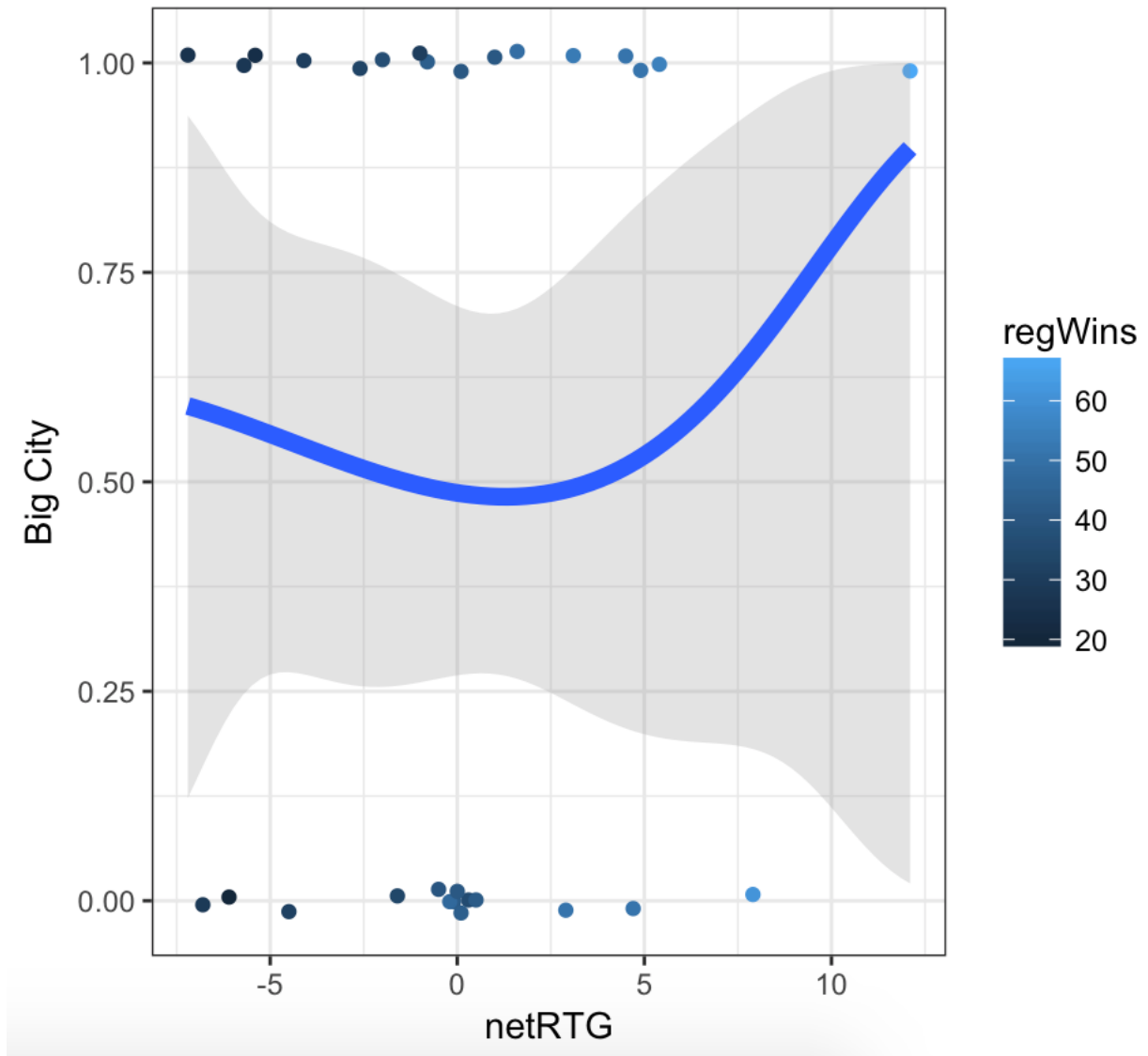
8   References

*** warning - I could not get this to load yet

References