How to measure value through NBA player statistics and collect subjective data about people's feelings towards NBA players

Jake Caldwell

For my research on the connection between popularity of NBA players and how statistically good they are, I need three main pieces of information: 1. How to measure the value of a player, 2. How to classify subjects based on their experience level, and 3. What statistical process is best used to compare rankings of items in lists when each list might have different items? To clarify what I mean by good, I can give two definitions that I will focus on: 1. The opinionated stance that comes from expert pundits; this is that a player looks like they are dominating or making things effortless, 2. Based on statistical production, heads and shoulders above the rest of the competition; more easily put as an MVP contender. In this literature review, I will divide my discussions thematically. Research has been done in each of the three categories I listed above, but my specific topic seems to be standalone.

Because my topic is mostly unexplored, it is imperative that I seek out as much information as I can. However, I do find myself pulling a lot of research from domains that aren't just basketball. There isn't much work within NBA research that uses the subjective opinion of human subjects about certain players. My project is looking to try to wrangle this knowledge gap because of its relevance to current public sports conjecture. The reason why using subjective data is hard is because I need to come up with a way to classify my respondents into categories that explain their experience level with the NBA. In simpler terms, how big of a fan is said person. In the main body of my literature review, this section will have most of its sources from other domains. I will attempt to bring these other domains together with the NBA and statistics side of the review to hopefully create a comprehensive overview of the literature required for this project.

Measure of Value:

Measuring the value of an NBA player may seem trivial, as the NBA already awards a player the MVP (most valuable player) award at the end of each season. However, since the early 1990s, the emergence of more detailed statistics that attempt to measure player performance shook up the conversation. The most common measure of a player's performance is the points, rebounds, and assists model (Sigler et al., 2000). This is the model that usually comes up when looking at box scores from previous games. The only problem is that this model doesn't focus on the defensive side of the ball. The basic defensive stats available are blocks, steals, and defensive rebounds. These stats are observable in a game, but might not tell the entire story of a players defensive ability (Franks et al., 2015). Defensive spacing and shot discouragement are factors that can only be noted from the "eye test" or from advanced spatial models (ibid). The database I am planning to use contains these stats, but it also has more advanced defensive stats: ratio of fouls to steals, opponent field goal percentage, and defensive box plus minus (Basketball-Reference.com, 09/21/2022). These might not be able to give me the spatial picture that Miller Franks uses in their paper, but this can give me an idea of defensive value.

Because of the ability to use statistical methods to try and measure player value, many researchers interested in the field have tried to come up with their own method. One such approach is the per minute approach by D. J. Berri. They critiqued IBM's weighted wins model:

$$[(PTS - FGA + RBO + RBD + AST + STL + BLK - PF - TO + (t * Wins * 10)) * 250)]/ [tPTS - tFGA + tRBO + tRBD + tAST + tSTL + tBLK - tPF - tTO]$$

and proposed using more minute data (Berri, 1999). The reason they did this was because of the tempo the game is played at. Each team runs a different pace of game. That means that some players might not put up as many total stats, but still be valuable because their team slows the game down. Berri also uses wins as the end goal for value (Berri, 1999). While I agree with accounting for team pacing, I don't necessarily think that wins is the end all be all when it comes to measuring a players effectiveness. Basketball is a team sport, and the best teams usually prevail as champions. Furthermore, Berri's take using wins only accounts for the games played in the regular season. The playoffs can level the playing field and allow for more upsets (Horowitz, 2018). For my research, I want to look at career statistics. One important distinction to make between my project and Berri's paper is that I am not looking to measure single season team value. I am looking to amass a career value. I can use wins as a part of my criteria, but it is not the end goal in my case.

There are newer methods of quantifying a players in-game and in-season results. One such method, created in the early 2000s, is John Hollinger's gamescore. A study from the university of BYU used this stat to measure for a player's in game production when calculating a players gaussian production curve (Page et al., 2013). Although this statistic has been criticized for lack of defensive production and helping poor shooters, this is a simpler version of the Player Efficiency Rating. It is commonly referred to as an add the good, subtract the bad stat:

$$GmSc = PTS + .4FG - .7FGA - .4FTM + .7ORB + .3DRB + STL + .7AST + .7BLK - .4PF - TOV.$$

It is a linear combination of basic but important in-game stats. In NBA YouTube videos, this stat is often favored over PER because it is easy to break down and understand. My goal is to do something similar: come up with my own stat that is easy to understand but is rooted in fundamental measurements of the game.

Subject Categorization:

The most difficult part of this research design is figuring out the most sensible way to note my participants' knowledge of basketball. This is necessary for my research because tracking expertise level can help me eliminate bias from my study. One of the preliminary ways I thought about classifying subjects was a two tiered question: Are you a fan of the NBA, and if so, how big of one. The only problem with this method is that it is highly subjective. My definition of being a fan is most likely different from someone else. In order to objectify my classification, I am pulling in strategies from other fields. One way to do this is to introduce a pseudo-likert scale variable that has strictly defined parameters. One such example from a study on music except familiarity defines each level as the amount of music education participants have had or are currently attaining (Kinney, 2009). Although this might work for an academic field like music education, there are no specific educational routes for learning about the NBA. Instead, exposure, outside research, dedication, and interest are some of the criteria that come to my mind when thinking about how someone might gain basketball knowledge. This makes study design difficult.

What does a compelling research design look like then? Consider a Partial Order Knowledge Structure (Desmarais, 1995). This is a user-expertise design structure that uses mathematical theory to determine if a user has mastered a prerequisite of understanding the function of some process/application (Desmarais, 1995). Instead of asking subjects their opinion on their interest in the NBA, I can ask them questions of increasingly niche knowledge with strictly defined parameters (answers). Then, using the results from this preliminary quiz, I can grade them on a likert scale that also has strictly defined values for how much they know about the NBA and the game of basketball.

Statistical methods and ranking:

The last and also difficult part of my project is taking the rankings of players that I have 1. Garnered from my model and 2. Aggregated from my survey, and comparing them using statistical methods. Two methods that I know about already are RBO and Kendall's Tau. These methods can give a variance of rankings between lists, however, they state in theory that the lists must contain the same elements. I envision the list that I am using for the survey will differ slightly from the list that I generate using my model. This presents a problem for the aforementioned methods I proposed. How do I compare lists with different items? That question is too simple and lacks context in ranking NBA players using two different mediums. Thinking about this logically, I infer that both my model and the survey will return similar results in the top 10-20 players. After that, things might vary quite heavily. Luckily, even though I said RBO needs the sameness of items in lists, it is a similarity measure bounded from 0 to 1. It also weights for rankings depending on where they are in the list. I can choose a value p that is also a number between 0 and 1 that determines what rankings carry the most weight. This allows me to choose to weigh the top 20 rankings as more important than the rest. Thus, I can, in theory, use RBO to determine if the lists are dissimilar without worrying about the lists having the same items. Then, because the output value of RBO is between 0 and 1, I can set a significance threshold. If the RBO value is less than that threshold, I can draw some conclusion from my study. This will make interpretability of my study fairly easy. Someone who is interested in my work but doesn't have experience with machine learning can simply look at my results and understand if there is a difference between the computational list and the survey list of players.

The statistics side of my project might seem like it is locked down, however, there is still a problem about using qualitative methods to talk about subjective, non-measurable qualities. What I mean here is that I am using NBA player statistics and counts of how many subjects recognize certain players to discuss if there is bias when pundits talk about how great an NBA player is. Greatness in this sense is a non-quantifiable concept. However, two studies provide examples that other similar concepts, like popularity and reputation, can be modeled by quantitative methods. The first of these two papers uses NBA player season stats and salary to define a players reputation (Ertug, 2013). This jump is not as wide as it may seem. From my point of view, it is easy for me to see why an NBA team would sign a certain player given their history of production and how much money they make. In the other paper, they use TV ratings and total revenue from specific games to measure the popularity of a "superstar" (Hausman et al., 1997). This connection is clear, TV ratings are a preliminary measure of popularity based on viewership. Using these two papers as inspiration, the connection that I will use to test for familiarity and bias is whether or not someone can identify a player based on a picture of said player.

Conclusion:

In my exploration so far, I have learned three main things: 1. Player production is a contested topic and there are many measures that track it, 2. Judging expertise levels requires strict preliminary evaluation of a subject's knowledge, and 3. RBO can be used to compare lists with different items. As my project stands right now, this is enough information for me to begin my initial data aggregation and begin summarizing. However, there are still a few areas of my project that I am unsure about. The end goal of my project is to see if there is a link between recognition and statistics. The connection I am trying to make is that if there is a difference between the ranking of players based on my model and who are the most recognized players, then familiarity influences our opinions on NBA players performance. This hypothesis isn’t as clear as the TV rating or the reputation measures that I discussed earlier in the review. Going forward, I will have to clearly define my hypothesis and make sure that my own bias towards certain players doesn’t influence my research.

Bibliography

Basketball-Reference.com - Basketball Statistics and History. https://www.basketball-reference.com/. (09/21/2022)

Berri, D. J. (1999). Who Is “Most Valuable”? Measuring the Player’s Production of Wins in the National Basketball Association. *Managerial and Decision Economics*, *20*(8), 411–427. http://www.jstor.org/stable/3108257

Desmarais, M. C., Maluf, A., & Liu, J. (1995). User-expertise modeling with empirically derived probabilistic implication networks. *User modeling and user-adapted interaction*, *5*(3), 283-315.

ERTUG, G., & CASTELLUCCI, F. (2013). GETTING WHAT YOU NEED: HOW REPUTATION AND STATUS AFFECT TEAM PERFORMANCE, HIRING, AND SALARIES IN THE NBA. *The Academy of Management Journal*, *56*(2), 407–431. http://www.jstor.org/stable/23412596

Franks, A., Miller, A., Bornn, L., & Goldsberry, K. (2015). CHARACTERIZING THE SPATIAL STRUCTURE OF DEFENSIVE SKILL IN PROFESSIONAL BASKETBALL. *The Annals of Applied Statistics*, *9*(1), 94–121. http://www.jstor.org/stable/24522412

Hausman, J. A., & Leonard, G. K. (1997). Superstars in the National Basketball Association: Economic Value and Policy. *Journal of Labor Economics*, *15*(4), 586–624. https://doi.org/10.1086/209839

Horowitz, I. (2018). Competitive Balance in the NBA Playoffs. *The American Economist*, *63*(2), 215–227. https://www.jstor.org/stable/26754448

Kinney, D. W. (2009). Internal Consistency of Performance Evaluations as a Function of Music Expertise and Excerpt Familiarity. *Journal of Research in Music Education*, *56*(4), 322–337. http://www.jstor.org/stable/40204937

Page, G. L., Barney, B. J., & McGuire, A. T. (2013). Effect of position, usage rate, and per game minutes played on nba player production curves. *Journal of Quantitative Analysis in Sports*, *9*(4), 337-345.

Sigler, K. J., & Sackley, W. H. (2000). NBA players: are they paid for performance?. *Managerial finance*.