Math 42 Final Project
How Has NBA Shot Efficiency and Shot Selection Changed Since 1980?

Ian Turner, Cole Hannum, Adam Goldstein, Seunghugh Kim, Hyunchae Kim

June 14, 2023

Contents

1 Abstract

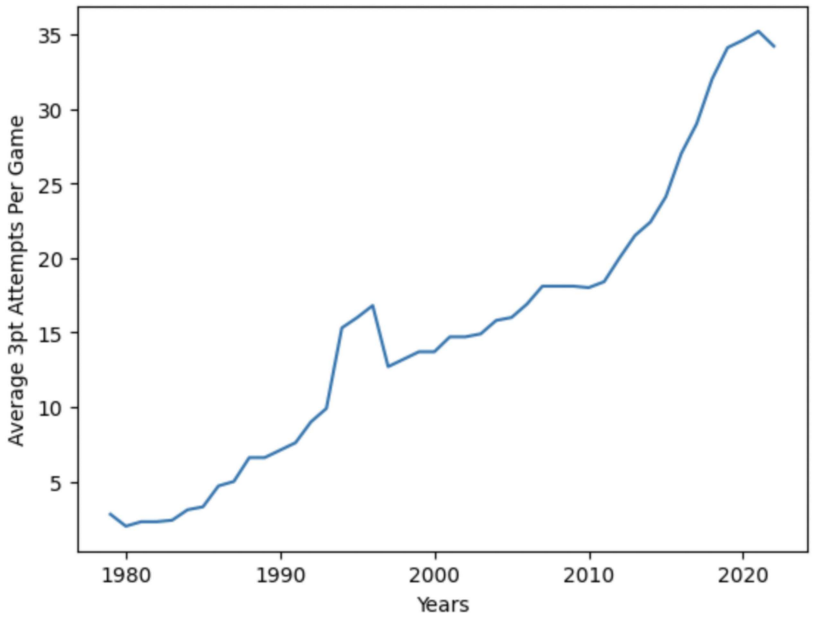
Since the creation of the three point line in 1979-1980 NBA season, the shot has continued to grow in usage and over the last 10 years, it has become arguably the most important shot on the court. The 3pt shot is widely considered the more efficient shot compared to a 2pt shot. But what is the proper balance of 3pt shots and 2pt ones? How has this balance of 3pt and 2pt shots changed over the years? How has it affected offensive output and shot selection?

We will use simulations, modeling, and statistical processes to attempt to answer these questions.

2 Background and Problem Description

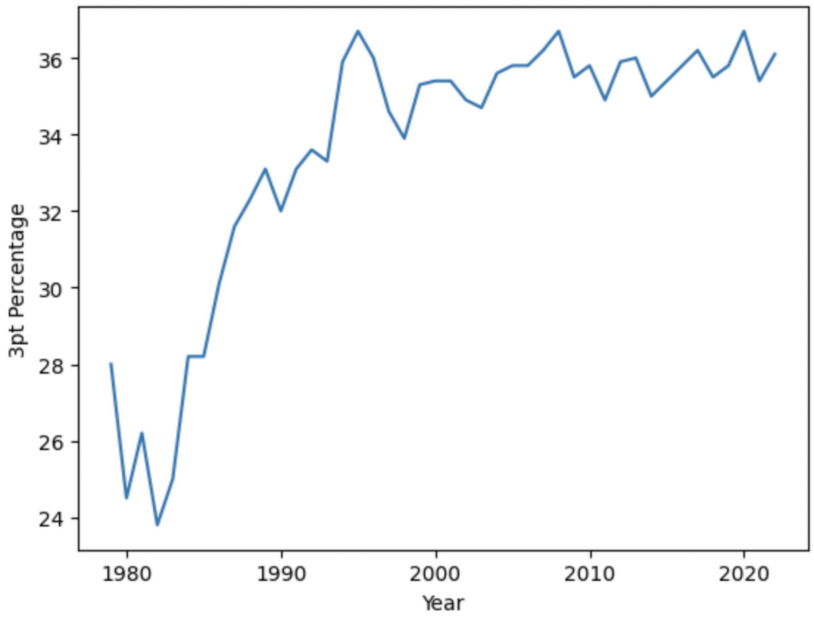
2.1 The Rise of the 3 Point Shot

With the 3pt shot rising to prominence, shot selection in the NBA has changed dramatically. However, the 3pt shot hasn't always had a significant impact on the game. In the first 5 seasons that the 3pt line existed, teams attempted less than 3 3pt shots a game on average, and averaging less than one 3pt make a game. Although 3pt shot attempts continued to increase through the 80's and the early 90's, the shot still didn't have a significant impact on the game, with teams attempting less than 10 threes a game, and making less than 33% of these attempts.



Because of this low 3pt attempt rate and the low scoring in general (with average teams scoring only 101.4 points a game), the NBA decided to make a significant change to the NBA 3pt line in the 1994-1995 season, moving it from 23 feet and 9 inches from the basket (with the exception of the corner which was only 22 feet), to 22 feet from the basket, making it easier to shoot from 3pt range. This dramatic change had a immediate effect on the prominence of the 3pt shot, with 3pt attempts going from a measly 9.9 in 1993-1994 to 15.2 in the subsequent year. In addition, teams were making more of their attempts, with the 3pt percentage changing from 33.3% to 35.9% (which is about as high as it was in the 2022-2023 season). (CBS article)

This major bump in the 3pt shot brought about by the NBA moving the 3pt line line closer to the basket was the first major step in the rise of the 3pt shot. After scoring and 3pt shooting improved, the NBA moved the 3pt line back to its original position, 23 feet and 9 inches from the basket in the 1997-1998 season. Although this decreased the 3pt attempts and 3pt percentage, attempts from behind the arc remained above 10 and percentage never dropped below 34%. From 1997-2012, the three slowly gained prominence, driven by shooters like Steve Kerr, Ray Allen, Reggie Jackson, and Jason Kidd paved the way for one man to change the game for ever.



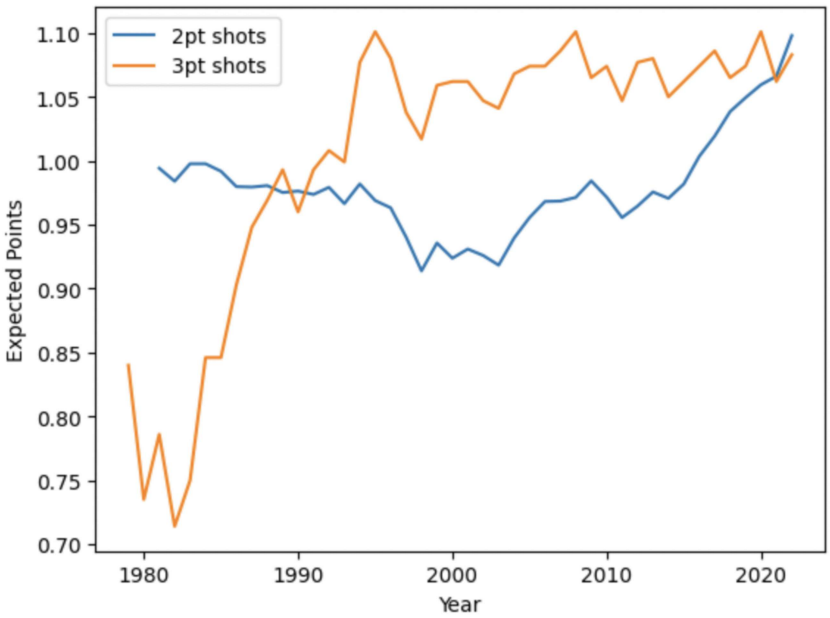
Stephen Curry was drafted by Golden State in 2009 out of Davidson College. Although he was known to be a great shooter, there were concerns regarding his ankles coming out of college - and he struggled with injury his first few

seasons in the league. He was able to overcome this ankle issue and began to emerge as a superstar, starting in the 2012-2013 season, in which he shot 45.1% from three point range, which was a level almost unheard of at the time. In addition to the 2012-2013 season seeing the emergence of Stephen Curry, three point attempts, which had been stagnant at 18 for the last 5 years, saw a jump to over 20 attempts per game, with 3pt percentage jumping from 35% to 36%, the largest jump from season to season in over a decade. In the subsequent seasons, Stephen Curry continued to drive forward the 3pt shot, winning MVP's in the 2014-2015 and 2015-2016 season, while continuing to shoot at a high level. Three point shot attempts continued to increase at the highest prolonged rate in NBA history. This significant increase in the relevance of the 3pt shot changed the way both offenses and defenses ran and players were drafted and developed. (NBA article)

2.2 Shot Selection

The rise of the 3pt shot has not only affected scoring, defense, and drafting, but it has also affected the frequency of other shots taken. According to an article written by Dartmouth Sports Analytics, the most popular shot (attempts-wise) from 1997 to 2009 was the mid-range shot (outside of the key but inside the 3pt line). But, with the rise of the 3pt shot, the mid-range shot has dropped to the least attempted shot on the court, while shot attempts 0-3ft away from the basket (the Restricted Area), and shots in the paint have remained constant since 2000. (Dartmouth Sports Analytics article)

This collapse of the mid-range shot can be mainly associated with the rise of the 3 pointer. In the 2022-2023 NBA season, players only made 42% of their mid-range attempts, while player's made 36% of their 3pt attempts. When looking at the expected points per shot, which is the shooting percentage multiplied by the value of a basket made in this region, mid-range has significantly lower points expectation than the 3pt shot, because the mid-range make percentage is only 6% higher, but is worth an entire point less than a 3pt shot.



The plot above shows how the expected points of 3pt and 2pt shots have changed over NBA history. Because of the sharp rise of 3pt percentage and the decline of 2pt shot percentage, the 3pt shot overtook the 2pt shot in the 90's, but in the last couple of years, the 2pt shot has overtaken the 3pt shot once again in terms of efficiency.

Now, let's look into the methods we used to find the optimal balance between 2pt shots and 3pt shots, as well as which locations on the court have the highest efficiency.

3 Simplifications and Assumptions

The data set being used for this paper is from BasketballReference.com and it contains the averages for all several basketball statistics from all the games in each season. However, for this analysis, we will only be examining the seasons following the institution of the 3pt line, so 1979 and beyond.

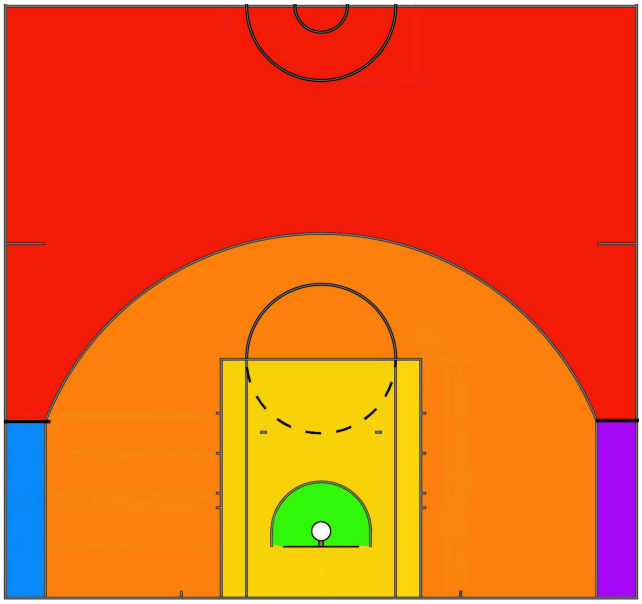
In order to determine the optimal number of 2pt and 3pt shots to attempt, we have to run simulations of each shot being taken and determine whether the expected point value is higher for a 2pt shot or a 3pt shot. To do this, we need to make a distribution assumption regarding the shot percentage of 2s and 3s. Because thousands of shots are attempted every NBA season, these shot percentages should approximately follow the normal distribution, according to the Law of Large Numbers. However, although the mean shot percentage is

recorded over seasons, the standard deviation is not, and it is not practical to attempt to collect all of it over every year. So, we calculated the standard deviation of 3pt% and 2pt% in the 2022-2023 season and used that as our constant standard deviation over all the NBA seasons we simulated.

For our look into the efficiency of shots based on their location, we again only looked at the 2022-2023 season, because collecting data over every single season is too much data to process. For the data we collected, the shots are broken up into 6 zones.

These 6 zones are:

Restricted Area (Green), In the Paint (Yellow), Mid-Range (Orange), Right Corner 3 (Blue), Left Corner 3 (Purple), Above the Break 3 (Red)



4 Mathematical Model

4.1 2pt shots vs 3pt shots

In our model we found the optimal number of two and three point shots by running simulations based on the three point and two point percentages in the NBA. We ran 200 trials with 100 shots simulated in each trial. This was done by pulling a percentage from a normal distribution for the two point shots and pulling a percentage the normal distribution of three point shots and multiplying this number by 1.5 because a three point shot is inherently 1.5 times more

valuable than a two point shot. The number of two and three point shots is then inserted into a list for each trial. We then found the mean of these two lists to find what the optimal amount of three point and two point shots is based on these trials. This allowed us to find and analyze the optimal amount of two and three point shots based on NBA shooting percentages.

Because of the higher percentage shot on free throws as opposed to any other shot, they have a significantly higher point expectancy than all other shots. Due to this, free throws have been removed from observed scoring we are trying to find the optimal amount of three point shots versus two point shots. Also, the amount of shots overall attempted in NBA games has changed overtime. To account for this the number of shots has been scaled to reflect the optimal amount of three and two point shots for that year on average. This means that mean for twos and threes is then changed in accordance with the year for all of 1980 - 2023 and the process is repeated.

4.2 Ideal Shot Selection in Different Areas

In this model, we focused on the most relevant NBA data, the 2022-2023 season. We started by assigning each team's data to an appropriately named variable. We then used these variables to extract each team's field goals made (FGM) and field goals attempted (FGA) for each shooting zone in the 2022-23 season. This data was then added to create an NBAdf that included the total field goals made and those attempted in the NBA for each shooting zone. We could then use these numbers to find a field goal percentage per shooting zone by simply dividing the FGM / FGA. Lastly, we then calculated the "expected points" by shooting zone by multiplying the field goal percentages of the shooting zones worth 2 points (Restricted Area, In the Paint (Non-RA), and Mid-Range) by 2 and multiplying the field goal percentages of the shooting zones worth 3 points (Left Corner 3, Right Corner 3, Above the Break 3) by 3. These variables were then constructed into a data frame.

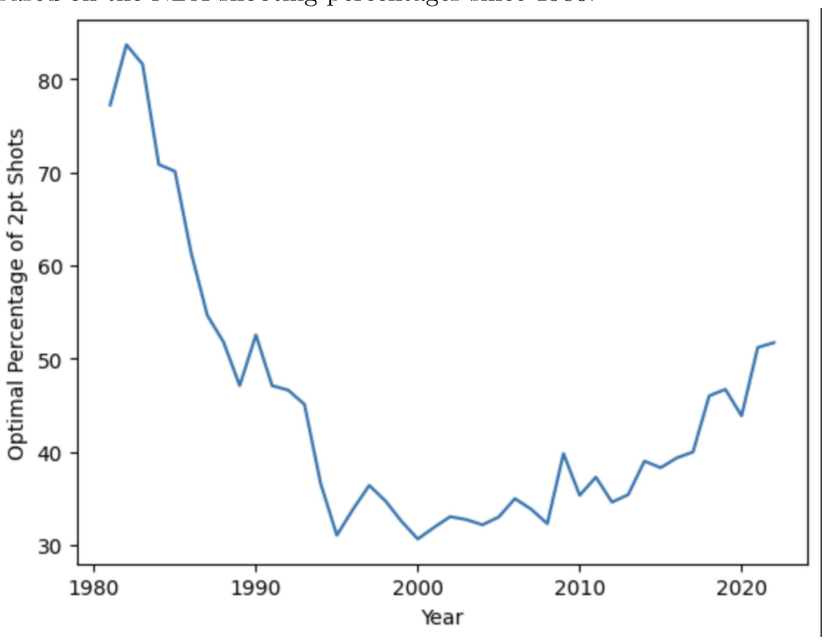
Next, we repeated the process above; however, instead of extracting each team's field goals made (FMG) and field goals attempted (FGA) for each shooting zone, we classified these variables by distance. Thus we created an NBA AvgDistance variable that had FGM and FGA organized by feet (ft) from the center of the hoop for all of the NBA. Once again, we used these numbers to find a field goal percentage per foot from the basket by dividing the FGM / FGA. After all the variables were made, we created a scatter plot to eventually add a regression line. For the scatter plot, we decided to plot points for every 2.5 feet from the basket, starting at 0 feet to and ending at 42.5 feet. We then made a regression line with scale one that related our x variable (Distance in feet) to our y variable (field goal percentage). This line of regression had the equation: y = 0.622 - 0.012x. Thus this model allowed us to see the negative correlation between distance and field goal percentage in the most recent NBA

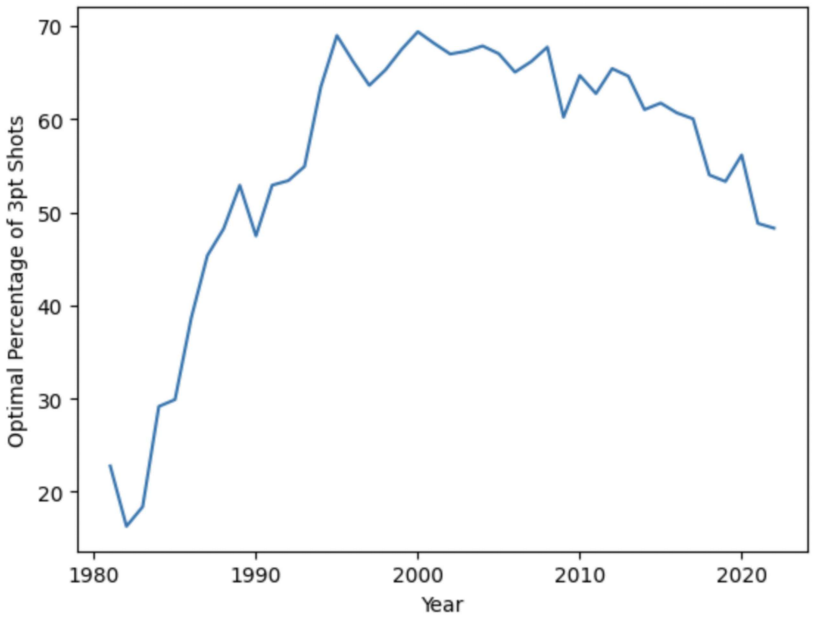
season. Approximately for every foot further away from the basket, the chances of the ball going in drop by 1.2 percent .

5 Results

5.1 2pt shots vs 3pt shots

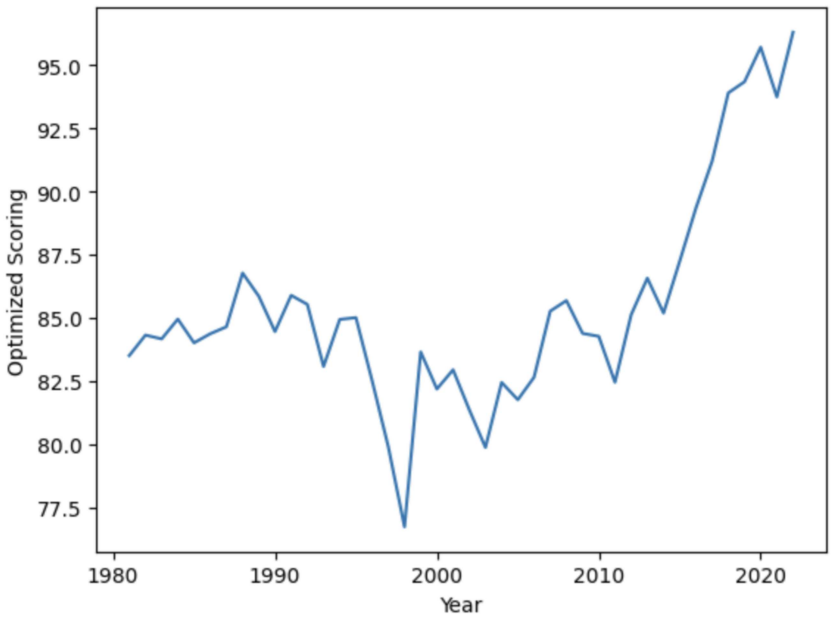
These were the results for the optimal amount of three and two point shots based on the NBA shooting percentages since 1980.



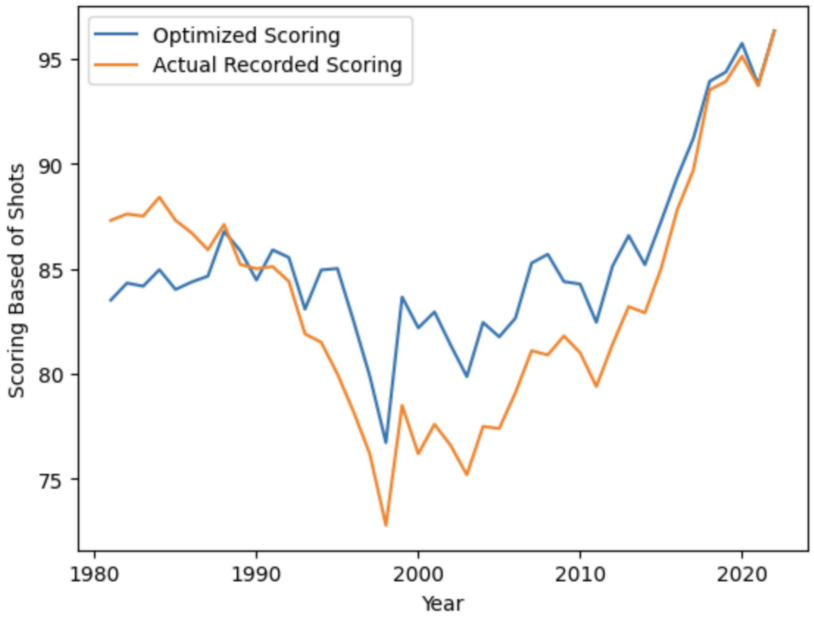


These graphs display that following the institution of the 3pt line in 1980, the 3pt shot continued to become the more efficient and optimal shot, reaching its maximum efficiency, in the early 2000's.

Using these amount of shots we can also find the optimal amount of scoring throughout the years.

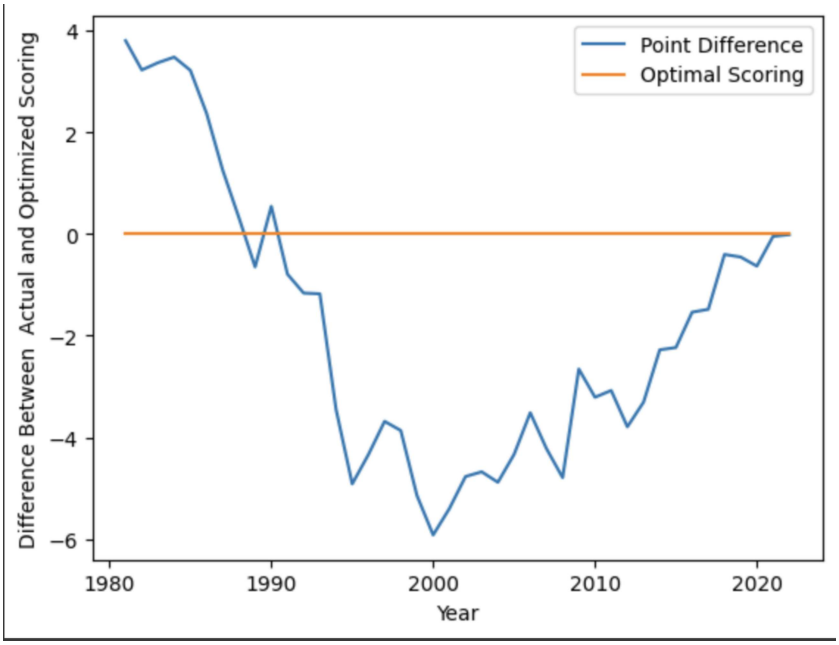


Comparing these values to the scoring changes over the NBA we can see a relatively accurate model.



As seen above, our model was fairly accurate. This displays that the 2000-

2013 era, before the massive growth of the 3pt shot brought about by Stephen Curry, was the least efficient era of basketball, as the differential between optimal scoring and actual scoring is the greatest. As the 3pt shot has risen, the actual scoring has been approaching the optimal scoring, meaning that offenses has become more efficient as the 3pt shot continued to rise. This can be seen more clearly in the graph below.

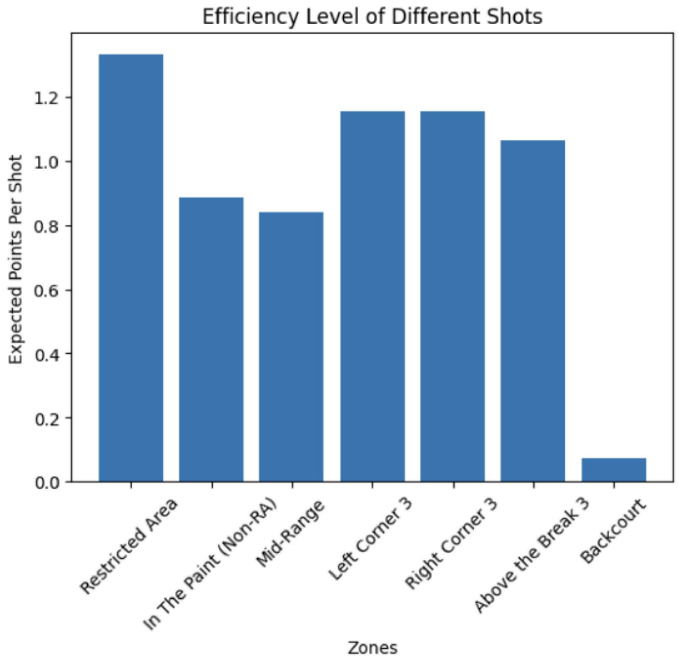


5.2 Ideal Shot Selection in Different Areas

These were the results for the expected points and field goal percentage by zone for the 2022-23 NBA season.

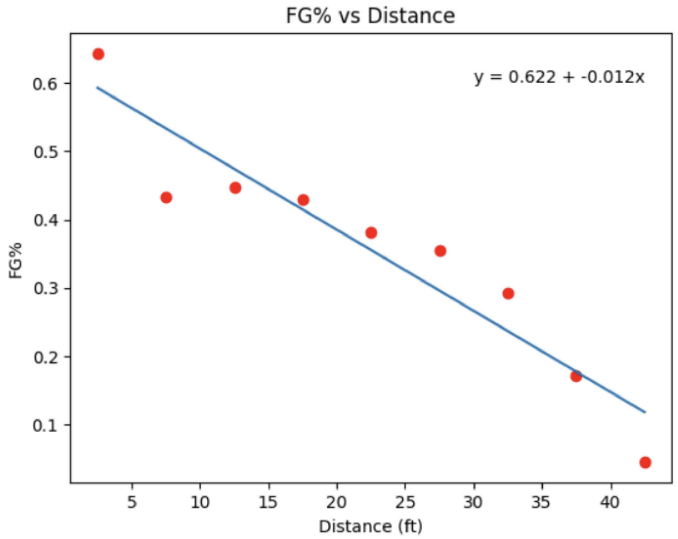
Zone	FGM	FGA	FG%	Expected Points
Restricted Area	43298	64983	0.666297	1.332595
In The Paint (Non-RA)	18827	42441	0.443604	0.887208
Mid-Range	10831	25760	0.420458	0.840916
Left Corner 3	4262	11066	0.385144	1.155431
Right Corner 3	3831	9939	0.385451	1.156354
Above the Break 3	22201	62625	0.354507	1.063521
Backcourt	10	406	0.024631	0.073892

This table can be shown as a histogram to visualize the efficiency level of different shot selections.



This histogram shows that the mid-range and paint shot are the least efficient, while the restricted area and 3pt shots are the most efficient.

Lastly, we can find the trend of field goal vs distance by adding a linear regression line to a scatter plot of the data.



This displays why restricted area shots are so efficient. Even though they are only worth two points not three, because their make percentage is so high, they are more efficient than the 3pt shot.

6 Improvements

There are several possible improvements to both the 2pt vs 3pt model and the shot selection model that could have been implemented to provide more insight, into both how modern basketball shot selection and 3pt usage will continue to change in the future, as how shot selection has changed historically.

First off, in order for the simulation to optimize the number of 2pt and 3pt shots to take, we assumed a constant standard deviation from year to year because it would have taken a great deal of data to find the standard for every single year. However, if this were to be done, the model would likely be more accurate, because earlier in the 3pt shot's life, there was likely a higher standard deviation because there were less 3pt attempts, where as in recent years, since there were more 3pt attempts, there was likely a smaller standard deviation, but the only way to tell is to comb through 40 years of data and calculate. This improvement could have changed the optimal number of 3pt and 2pt shots by year.

When looking at possible improvements for our shot selection model, there are several possible improvements as well. Like the 2pt and 3pt optimization

model, we could have looked at how shot selection has changed over time, but with the data set we chose that was not a practical option, and would have required downloaded 40 data sets. Another possible improvement is attempting to optimize the number of shots in each zone like we did with the previous model. The reason this was not attempted was because if we were to run a similar simulation model, it would have likely suggested to take upwards of 90% of the shots in the restricted area, which is simply not a practical way to run an NBA offense. However, this could have been altered by adding a sort of defensive metric in each zone, with points in the paint having the highest, and 3pt shots having the lowest. But, there was no defensive metric we could find per court zone.

7 Conclusion

This paper set out to explore how the three point shot has been utilized and how it should be optimally utilized in NBA offenses. In addition, we looked into shot selection in modern NBA basketball, and where the most efficient shots are taken. We employed both statistical methods by running simulations drawing from the normal distribution, and mathematical optimization methods to answer these questions.

When looking at the optimal number of 3pt shots to attempt and how it has changed over the last 43 years, we discovered that early in the 3pt shot's life, both optimized three point shooting and actual three point shooting were low. But, following its slow start, the efficiency of the three point shot continued to rise, and it became the more optimal shot in the early 90's. The three pointer reached it's peak due to 2pt efficiency dropping in the early 2000's. However, in the last 10 years, the two point shot as made a comeback. So although today's NBA game has more 3pt shots than ever, the 2pt shot is slowly becoming the more efficient shot. This is likely due to the 3pt shot's ability to spread the floor, but more research and statistical analysis is needed to explain why the 2pt shot has retaken the 3pt shot as the most efficient shot in basketball.

After looking into the different efficiencies of 2pt and 3pt shots, we wanted to look deeper into what types of 2s and 3s were ideal. We collected data of the most recent season and found the most optimal shot by looking at the expected points of each shot. By doing this, we found that shots in the restricted area and shots behind the 3pt line are the most efficient shots, whereas the mid-range is the least efficient shot on the court, based on the NBA 2022-2023 season's data.

However, NBA offensive and shot selection is more complicated than numbers can represent. Several teams in NBA history have found great success by focusing on 2pt shots and points in the paint rather than 3pt shooting, even when the statistics point towards their type of basketball being inefficient. In

addition, although the mid-range shot is the least efficient shot on the court, players such as Kevin Durant, Chris Paul, Tim Duncan who excel at the mid-range are considered among the NBA's all time greats. So while statistical analysis and mathematical modeling can provide valuable insights, there are elements of basketball that statistics can't evaluate properly, or at least, not yet.

8 References

https://www.basketball-reference.com

https://www.nba.com/news/stephen-curry-changed-game-long-before-3-point-milestone

https://www.cbssports.com/nba/news/what-the-shortened-nba-3-point-line-of-the-mid-1990s-says-about-the-future-of-long-range-shooting/

https://sites.dartmouth.edu/sportsanalytics/2022/01/25/nba-shot-selection-how-have-players-changed-their-shooting-in-the-last-twenty-five-years/