

# How Much Should Nikola Jokic Get Paid?

Zhaocheng Xu

Stern School of Business, New York University, New York, United States

**Abstract.** This paper talks about NBA player's salary taking into consideration of multiple factors such as points per game (PPG), rebounds per game (RPG), assists per game (APG), steals per game (SPG), blocks per game (BPG), turnovers per game (TPG), box plus-minus (BPM), minutes per game (MPG), players' age, games played (season), and their positions on the court. Along with restrictions of salary cap, this paper finds out players are underpaid; The study result could be applied to players who are subject to the minimum salaries, as the salary prediction logic is the same as for players who are subject to the maximum salaries.

**Keywords:** NBA player, salary cap, PPG.

## 1. Introduction

It has been observed that the famous NBA player Nikola Jokic earns \$588,000 for every basketball game that he plays. In other words, \$204 per second. Despite what seems like excessive compensation, this study will prove the fact that Jokic is currently *underpaid*. My analysis shows that under competitive wage, he could potentially earn \$78.15 million annually based on his performance instead of his current annual salary \$46.6 million—which is half of his potential.

The NBA (National Basketball Association) is a premier basketball league in North America comprising 30 franchise teams (29 in America and 1 in Canada). Founded in 1946, the NBA continued to increase its economic and social impact from local to global. The NBA gained a large fan base through modernization and globalization and maintains its position as the world's most famous basketball association. In the last decade (since 2013), the NBA reported their revenue growth more than doubled, reaching over \$10 billion in the 2022-23 season. At the meanwhile, their impact keep boosting with an astonishing average TV viewer of 1.59 million in the regular season.

The NBA's revenue mainly comes from Basketball-Related Income (BRI), which includes revenue from broadcast rights, arena revenue, sponsorship and media, merchandise sales, and digital media (Manuel 2020). Around 50% of the NBA's revenue is paid out to players. That fraction is determined through the league's Collective Bargaining Agreement (CBA) between the NBA and the National Basketball Players Association (NBPA), rather than through free market competition. This agreement is implemented by imposing salary caps at both the team and the player level.

The concept of the team salary cap was first implemented in the NBA in the 1984-85 season. Essentially, the team salary cap places a limit on a team's total expenditure on players' revenues. A team cannot sign players with contracts that exceed the team salary cap, unless those players are subject to certain exceptions, such as midlevel exceptions, rookie exceptions, and two-way contracts. The team will also be levied a high luxury tax after passing a specific threshold, which disincentivizes overpayment. In the 2023-24 season, the salary cap is \$136 million, and the luxury tax threshold is \$165 million. The luxury tax is \$1.5 per dollar for \$0-5 million above the tax line, \$1.75 per dollar for \$5-10 million above the tax line, \$2.5 per dollar for \$10-15 million above the tax line, and so on.

Apart from the team salary cap, there is also a cap on individual players' salaries, called the maximum salary. This notion was first implemented in the 1998-99 season. Generally, a player's maximum salary differs based on his Years of Experience (YOE). For example, in the 2023-24 season, it is \$32 million for 0-6 YOE, \$39M for 7-9 YOE, and \$45 million for 10+ YOE (*NBA minimum salaries* 2024). Player's maximum salary could also increase if he has won certain achievements (such as making the All-NBA teams). Lastly, the maximum salary and years of the contract are generally higher if a star player decides to sign with his home team (the team they are currently in) instead of signing with another team to incentivize players to stay in their home teams (Ginnitti 2019).

There are several crucial reasons for establishing individual players' maximum salaries—one of them being to maintain the competitive balance in the league. By limiting the amount that any single team can pay, the NBA aims to prevent wealthier teams from amassing all the top talent, which could lead to a few dominant teams and many weaker ones. This helps to keep the league competitive, as teams are more evenly matched in terms of player talent. Another reason, as stated previously, is to retain star players within their original teams, rewarding them for staying rather than moving teams. This helps teams, especially ones with smaller market and less financial strength, to retain star players (Ginnitti 2019). Lastly, the reason ties back to the concept of value shifting. Instead of giving most of its salary to a few superstar players, the team can change the value and redistribute part of its salary to other players. This results in superstars sometimes being underpaid relative to their game performance and popularity (Scolobey 2023).

Given the attention and pure economic magnitude of NBA players' competition, an interesting and unanswered question is how large the value shift in a given season is due to the imposition of maximum salaries on individual players. To answer this question, one needs to estimate the salary that top players (those that are subject to the cap) would have received in free market bargaining. To do so, I estimate a model that links player and team attributes to their salaries using data from players that are not subject to the cap. This model allows us to estimate the counterfactual salaries of top players. The value shift is the difference between players' actual salary and the salary they could get based on their performance.

## 2. Literature References

Research and discussions on top player's salaries have been studied for a long while. According to Sigler and Sackley (2000), there is a positive and significant relationship between an NBA player's salary and a player's points per game and rebounds per game for the 1997-98 basketball season. A simple regression and multivariate regression were done in this research to prove the credibility of the positive relationship between salary and the measures of performance. Yet, the data used in this study is over 20 years old, and the match rules changed over the years, calling into question the results. With the fact that we have seen positive correlations before, I still want to conduct my research based on their findings. Furthermore, this study extracts both the player salary and performance data from the same season, which does not account for the logical cause-and-effect relationship between performance and subsequent salary. For my study, I will consider the causal relationship between the two variables by taking the player performance data in the year prior to receiving the new contract in the current (2023-24) season. Furthermore, I will be using the sample of players who renewed their contracts so that the cause-and-effect relationship is constructed.

Another similar research conducted by Kiss (2019) plays a significant role in my research too. By using regression models to produce a prediction model, Kiss shows that traditional metrics, simple statistics like points per game, rebounds per game, and assists per game, play a significant role in understanding the value of a player, instead of advanced statistics such as player efficiency rating and floor coverage impact. However, Kiss's study does not explicitly deal with players who have maximum or minimum salaries. This disrupts the accuracy of the results as salaries can no longer match properly with player performance. In my paper, I will take maximum and minimum individual salary caps into account to address this inaccuracy.

Lastly, in Papadaki (2020) and Tsagris's paper, the authors utilize LASSO variable selection and the RF non-linear algorithm to detect the important statistics that are mostly associated with the NBA player salaries and predict player salaries. The study found that the relationship between player statistics and salary is non-linear. One aspect that this research fails to address is that it does not ask the counterfactual of "what will happen if there is no salary cap?" In my paper, I will address this counterfactual by applying salary prediction model on players with maximum salary.

### 3. Data

Variable selection is crucial for model construction and the accuracy of model prediction. After careful consideration, I selected eleven independent variables: points per game (PPG), rebounds per game (RPG), assists per game (APG), steals per game (SPG), blocks per game (BPG), turnovers per game (TPG), box plus-minus (BPM), minutes per game (MPG), players' age, games played (season), and their positions on the court. Within these eleven variables, the position variable isn't simply a numeric variable, which needs to be treated using dummy variables.

As for the independent variables, I chose simple variables over advanced ones (such as player efficiency rating, win shares/48 minutes, and floor impact counter) because a number of previous papers indicate that simple statistics more straightforwardly reflect players' salaries than advanced statistics do. For example, Kiss's study shows that traditional metrics play a significant role in understanding the value of a player.

As for the dependent variable, I chose  $\ln(\text{salary})$  instead of salary. The reason for this will be discussed later in the Data Collection section.

Data was sourced from multiple authoritative websites. A major one is a website called "Basketball Reference" for performance statistics, which sorts out various NBA data in various seasons (individual and team-wise). I also used a website called "Spotrac" for salary data. Spotrac allows me to apply multiple filters to the salary data in order to acquire the desired group of players.

### 4. Methodology & Results

I came up with the idea that I could fit a model that links the compensation with observables. In order to control the variables, I only fitted this model using players not subject to the salary cap to minimize the inaccuracy that results from salary set by non-performance considerations, which means they are restricted by maximum or minimum salary. Then, I will apply this model to players with maximum salaries to predict the salary they would have worthed in the absence of individual caps. The difference between the predicted and actual salaries could determine how much their value has been misunderstood.

### 5. Data Selection & Analysis

In this study, I only focused on players who started their new contracts in the 2023-24 season. 11 independent variables were selected, and data was collected from the 2022-23 Season, while the salary as the dependent variable was collected from the following 2023-24 Season. Player's performance from this year will affect their salary in the following year as they sign their contract and review their earnings annually.

Furthermore, there are several filters that I had to apply when selecting the group of players for data collection because some players with particular circumstances could potentially disrupt the accuracy of prediction. First, I need to avoid players with maximum salaries because their performance may not align with their salary due to the salary limit. Since maximum contracts are announced each year and the amount is fixed, plus the players are usually worldwide famous, I pull out the list of players subject to maximum contract and matches them out of the list. The same applies to players with minimum salary because their performance could potentially be below their salary due to the bottom limit of how much they could be paid. For players subject to minimum contracts, since the amount is fixed, I marched the amount and filtered out their name from the list accordingly. Lastly, I need to avoid players whose data in the 2022-23 season is incomplete. This includes but is not limited to rookies who just entered the league and players with severe injuries that cost them an entire season. I then utilized the filtering standards above to produce the list of players eligible for data collection.

Initially, I started with the list of all the NBA players in the league in the 2023-24 season (as of December 21st, 2023), which had 565 data points. Then, out of these 565 players, 179 of them started

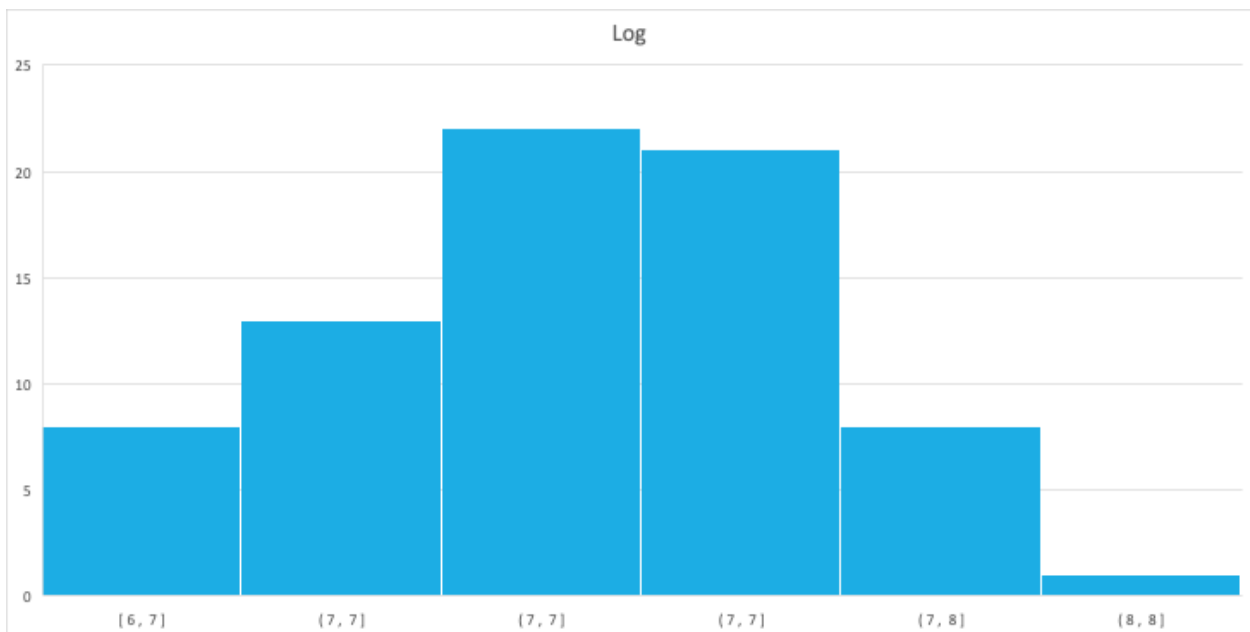
their contract in the current (2023-24) season. Within this smaller pool of players, I found out the ones receiving maximum (6 players) and minimum salaries (157 players) and the new players (12 players). While collecting data, I was able to detect several players whose data was incomplete. This is usually due to severe injuries that cause them to miss the entire season or the fact that they played in some other leagues. After applying filters and removing players who lack statistics, I eventually shortened the final list to 73 players meeting the condition.

I collected data on the 2023-24 season salary for the 73 players and put them in summary statistics. Results show that the data has a mean of \$14,600,000 and a median of \$12,402,000 and is right-skewed, as shown below in Figure 1. This right-skewed pattern is not desired as I intend to conduct multivariate regression based on the assumption that the data is approximately normally distributed.



**Figure 1.** 2023-24 Salary Distribution.

Then, I applied the natural log of the salary data. This time, the data displays a more normal distribution, as shown in Figure 2 below. Therefore, I decided to use  $\ln(\text{salary})$  as the dependent variable instead of salary.



**Figure 2.** 2023-24 Log Salary Distribution.

With the finalized list of players for data collection and their respective salaries in the 2023-24 season, I started collecting data for the independent variables. I organized the data into a table named 'P 2022-23' in the Excel spreadsheet, the columns' labels are Player, 2023-24 Salary Formatted Salary (number format), Ln Formatted Salary, PPG, RPG, APG, SPG, BPG, TPG, P&M, MPG, Players' Age, Games Played, Position-C, Position-PF, Position-SF, Position-SG, and Position-PG.

After data collection for players not subject to the salary cap, I then fitted a regression analysis to generate a regression model, the dependent variable being the Ln Formatted Salary and the independent variables being all the ones on the right side of the Ln Formatted Salary on the Excel sheet. Table 1 below depicts the regression results generated.

**Table 1.** Multivariate Regression Results on Player Salary Prediction.

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.8539846							
R Square	0.7292897							
Adjusted R S	0.6639458							
Standard Err	0.1657619							
Observations	73							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	14	4.2933166	0.3066655	11.160798	9.191E-12			
Residual	58	1.5936672	0.027477					
Total	72	5.8869838						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	6.4224164	0.2301438	27.906097	2.615E-35	5.9617335	6.8830994	5.9617335	6.8830994
PPG	0.0247098	0.0093855	2.6327738	0.0108365	0.0059228	0.0434969	0.0059228	0.0434969
RPG	-0.003608	0.0176004	-0.205007	0.8382842	-0.038839	0.0316229	-0.038839	0.0316229
APG	0.0171304	0.0326273	0.5250322	0.6015621	-0.04818	0.082441	-0.04818	0.082441
SPG	0.0290651	0.0848057	0.3427257	0.7330438	-0.140692	0.1988221	-0.140692	0.1988221
BPG	0.1018702	0.0659088	1.5456236	0.1276348	-0.030061	0.233801	-0.030061	0.233801
TPG	0.0359108	0.076329	0.4704736	0.6397812	-0.116878	0.1887	-0.116878	0.1887
P&M	0.0192645	0.0141366	1.3627384	0.1782328	-0.009033	0.047562	-0.009033	0.047562
MPG	0.0096909	0.0082858	1.1695862	0.2469515	-0.006895	0.0262766	-0.006895	0.0262766
Players' Age	-0.002675	0.0065169	-0.410399	0.6830251	-0.015719	0.0103704	-0.015719	0.0103704
Games Played	-0.000346	0.0016057	-0.215616	0.8300437	-0.00356	0.002868	-0.00356	0.002868
Position-C	0.0300608	0.1292344	0.2326064	0.8168859	-0.22863	0.2887515	-0.22863	0.2887515
Position-PF	0.1397692	0.0943028	1.4821325	0.1437182	-0.048998	0.3285368	-0.048998	0.3285368
Position-SF	0.1142979	0.0867506	1.3175459	0.1928367	-0.059352	0.2879481	-0.059352	0.2879481
Position-SG	0.0113672	0.0802202	0.1416998	0.8878082	-0.149211	0.1719454	-0.149211	0.1719454

The regression equation has an adjusted R-square of 0.664 and an F-value of 11.161, which corresponds to a significance of 9.191E-12. This is significantly lower than the alpha at 95% significance level (alpha = 0.05), meaning that the regression equation is meaningful in estimating players' compensation based on statistical performance. In terms of the individual independent variables, other than PPG, all other variables have a P-value that is larger than 5%. This may suggest that, individually, variables other than PPG are statistically insignificant. This may also be the result of the study's relatively small sample size (73 players).

I then applied the regression equation to players with maximum salary (those who are subject to the maximum salary cap) to find out what amount they should actually get paid if there was no salary limitation. Table 2 below illustrates the summary statistics of players who are subject to the maximum salary cap in terms of predicted salary, actual salary, difference, and percent difference.

**Table 2.** Regression Summary Statistics of Players Subject to the Max. Cap.

Name	▼ Predicted Salary▼	Actual Salary▼	Difference	▼ Percent Difference▼
Nikola Jokic	79,154,313	47,607,350	31,546,963	
Joel Embiid	110,668,579	47,607,350	63,061,229	86.13%
LeBron James	76,017,487	47,607,350	28,410,137	
Ja Morant	50,389,519	34,005,250	16,384,269	
Zion Williamson	79,458,266	34,005,250	45,453,016	61.04%
Darius Garland	34,440,287	34,005,250	435,037	

The regression results show that all six players are currently underpaid, Joel Embiid being underpaid the most and Darius Garland being underpaid the least. The total extent of the underpayment for the six players is over 185 million dollars (the sum of the six players' difference in predicted and actual salary). I also calculated the percent difference between the actual and the predicted salaries for each of the three players. The table captures two types of players, the first three in light blue (Nikolas Jokic, Joel Embiid, and LeBron James) are the ones who have been in the league for a relatively longer time, whereas the three in light green (Ja Morant, Zion Williamson, and Darius Garland) are the rising stars who just completed their rookie contracts, which are more restricted to receiving a higher salary. However, even though the first three players receive higher salaries than the second three, the percent difference between the predicted and actual salary of the first three players (86.13%) is still higher than that of the second three players (61.04%). This shows that although the first three players are paid the most, they are also underpaid the most.

## 6. Discussion & Evaluation

Therefore, by generating a regression model using player performance data, I was able to estimate players' salaries based on their performance. By applying the regression model to the six players whose salaries are subject to the maximum caps and began in the 2023-24 season, I found that the total extent of the value shifting is over 185 million dollars. To answer the question raised initially "How Much Should Nikolas Jokic Get Paid", this study proves that Nikolas Jokic should get paid \$78.15 million based on his performance, which is \$31.55 million more than his actual salary.

The fact that all six players with maximum salaries and began their contracts in the 2023-24 season are underpaid according to the model highlights the importance of the individual salary cap in creating value shifts in the NBA, which allows teams to better redistribute portions of their total payment to other players, reducing the salary inequality between players. It is worth noting that the value shift of over 185 million dollars accounts for over 4% of the entire league's expenditure on players' salaries, which is underestimated because there are a total of 29 players with salaries subject to the maximum cap, but the above calculation takes into account only 6 of them. If this value shift is assumed to be equal for the remaining 23 players, then the total comes to 19.33% of the league's expenditure on players' salaries.

Based on the results, it shows that Joel Embiid should be paid the most based on his performance. This complies with reality as Joel Embiid is currently a candidate for the Most Valuable Player Award (MVP Award), and the MVP Award is largely determined by a player's statistical performance. His good performance is recognized and reflected in his predicted salary in my regression model. This validity further adds confidence to my research.

The strength of the research lies in its methodology. Unlike other research papers that focus their analysis on one single season, I focused on consecutive seasons and discovered the cause-and-effect relationship between player performance and salary. As a result, I collected data on player performance in the season prior to receiving their new contracts.

However, in the real-world situation, the determinants of players' salaries are complex, and this research fails to account for all factors. For example, I did not explicitly take into account players' popularity, the popularity of the teams they are in, and the college/institution they were part of before joining the NBA, although these all can impact how much players earn. Furthermore, in the ANOVA analysis in the regression model, only the independent variable PPG has a P-value lower than alpha (0.05). All other variables have P-values greater than alpha, meaning that, individually, each variable is insignificant in estimating players' compensation.

The study result could be applied to players who are subject to the minimum salaries, as the salary prediction logic is the same as for players who are subject to the maximum salaries. However, such an approach may be more difficult to conduct as players with minimum salaries may lack sufficient performance statistics (they have less chance to play) for the regression model to be applied appropriately and fairly, and thus is left for future research.

## References

- [1] Adrien, Bouchet, Michael, et al. Dynamic Pricing Usage in Sports for Revenue Management[J]. *Managerial Finance*, 2016.
- [2] David, J., Berri, et al. Stars at the Gate[J]. *Journal of Sports Economics*, 2016. DOI:10.1177/1527002503254051.
- [3] Burke, Monte. Average Player Salaries In The Four Major American Sports Leagues.[J]. *Forbes.com*, 2012:3-3.
- [4] Owan L M. The Effects of Recessions and Market Sizes on NBA Player Salary[J]. 2017.
- [5] Zhao-Lin W, Qing-Yuan Y. An Assessment on Non-market Value of Rural Residential Land in Chongqing Area of Qinba Mountain[J]. *China Land Sciences*, 2017.
- [6] Rishé P J, Mondello M J. Ticket price determination in professional sports: An empirical analysis of the NBA, NFL, NHL, and Major League Baseball[J]. *Sport Marketing Quarterly*, 2004, 13(2): 104-112. DOI:10.4103/0019-5278.40819.
- [7] Paul R, Weinbach A, Losak J. Market Efficiency and the Setting of Salaries for NBA Daily Fantasy on FanDuel and Draft Kings[J]. 2020. DOI:10.5750/jpm. v14i2.1813.
- [8] Santora S D, Scheetz R, Eaton J, et al. Monetizing Fandom: A Look into the Future of Sports and Sponsorship[J]. 2017.
- [9] Gibbs J. Point Shaving in the NBA: An Economic Analysis of the National Basketball Association's Point Spread Betting Market[J]. 2007.
- [10] Gramacy R B, Taddy M, Tian S. Hockey Player Performance via Regularized Logistic Regression[J]. *Stats*, 2016. DOI:10.48550/arXiv.1510.02172.
- [11] Segal A E. Basketball Now! The Stars and Stories of the Nba[J]. 2015.
- [12] Khoo J. PODCAST: Asia NBA market sees glimmer of hope for a rebound[J]. *ICIS News*, 2020(Sep.1).
- [13] Xue H, Alberta U O. Managing Guanxi: A Case Study of the NBA in China[J]. 2015.
- [14] Oberhofer H, Schwinner M. Do individual salaries depend on the performance of the peers? Prototype heuristic and wage bargaining in the NBA[J]. *Department of Economics Working Paper Series*, 2017.