Aditya Sangana
12/6/2024

Research Paper: Research into NBA athlete salaries

1. Introduction

Since their performance in their respective professions is publicized nationally and sports culture is such a big part of our society, professional athletes often face the most scrutiny when it comes to their compensation. Their salaries often provide a fascinating lens into economic valuation, market behavior, and the broader economic disparities within team markets across the league. In this study, I will aim to analyze the factors that influence the salaries of professional NBA players in the past 2022-2023 season. In general, the salaries of professional athletes generally tend to reflect the players' performance, the overall team's success, and market dynamics. A highly competitive salary cap system forces teams to balance both rewarding their top-performing players and managing the overall team budgets.

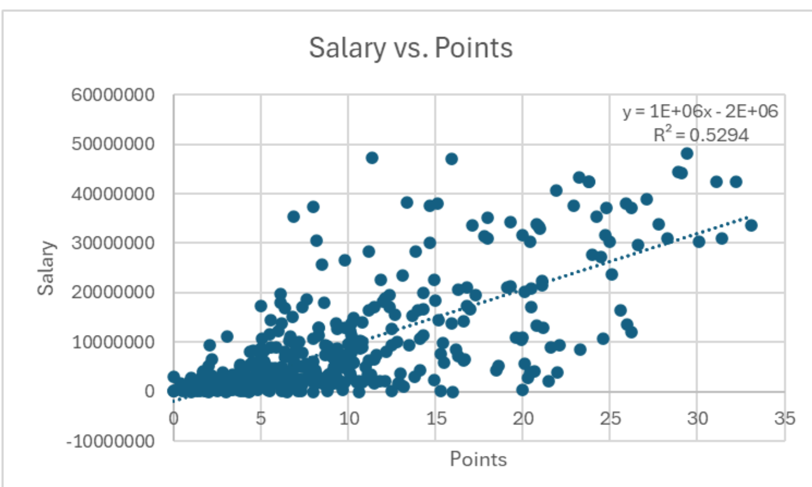
This study explores the predictors of player salaries by focusing on the relationship between individual stats–points, assists, and rebounds per game—and an external factor, the overall team's revenue. By understanding each of these elements and analyzing their impact on how much these athletes are able to earn, we can gain further insight into how talent is valued in arguably one of the most popular sports leagues in the world. This study extends existing literature by analyzing both individual performance metrics and team market sizes to help provide a more comprehensive understanding of determinants of salary. We hypothesize that a better individual performance, specifically scoring and team revenue, will have a strong positive correlation with player salaries: points per game will be the most significant predictor of salary, team revenue will have a smaller but a statistically significant positive effect on salaries, and finally, secondary metrics, such as assists and rebounds per game, will be significant as they will have a moderate but positive effects on salary.

This hypothesis is supported by previous research done on similar topics. In the study "The Salary Cap: Small Market vs. Big Market NBA Teams and the Effects of COVID-19," the researchers highlight that there is a significant disparity between small-market and large-market teams as larger-market teams tend to have much greater financial resources and capital which allows them to pay their top-performing players higher salaries. On the other hand the article "NBA Players' Pay and Performance: What Counts?" centers on the importance of an individual player's performances, more specifically scoring, as determinants for NBA athlete salaries. These studies helped inform us in our approach to selecting variables and developing a hypothesis for our own study. We extended existing literature by using both a player's individual performance with team-level revenue metrics to provide a comprehensive analysis of salary predictors in the NBA. Nearly all of the previous studies that we went through focused solely on player performance but in this study we took a more holistic view by incorporating market factors to give proper insight into salary structures. By analyzing 467 player observations across all 30 NBA teams, we aim to analyze the factors that go into deciding salaries, salary disparities, and how talent is valued in one of the world's most popular sports leagues.

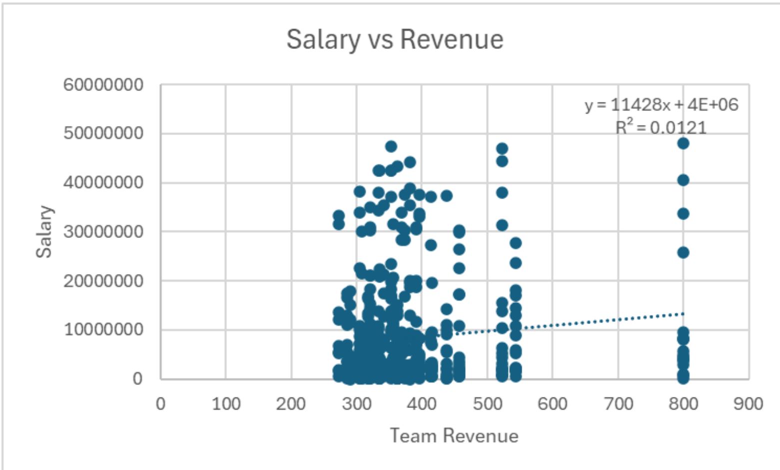
2. Key Results

Our dataset analyses examined the relationship between NBA player salaries and potential determinants, including individual performance metrics—points per game, assists per game, rebounds per game—and an external factor, team revenue. The main goal of these analyses was to find out which determinant had the most significant influence on player compensation. Based on the data and the regression analysis, the findings did support the hypotheses.

After analyzing the data, we were able to observe a strong positive correlation between points per game and player salaries. As you can see in the Salary vs. Points scatter plot, the relationship between the variables is clear. The regression line indicates that for every additional point scored by a player, their respective salary will increase by $869,541.70. The R^2 value of .5294 suggests that 52.94% of the variation in athlete salaries is explained by points scored per game alone, demonstrating how integral this determinant is in influencing player compensation. This analysis aligns with the NBA's emphasis on offensive contributions being critical in valuing talent.



Unlike points scored, team revenue proved to be less impactful but also displayed a positive and statistically significant relationship with player salaries. The Salary vs. Revenue scatterport highlights that there is a modest upward trend with a slope that demonstrates that for every additional million dollars brought in as team revenue, an athlete's salary will increase by $10,636.62. This analysis gave us an R^2 value of .0121 which is quite low, further indicating that though a team's market size/revenue and financial health does in fact have a non-negligible influence on salary, it is much smaller. These results suggest that players that are on teams with larger-market sizes or higher team revenues may see an increase in compensation opportunities even with the salary cap taken into regard.



Next, we looked at secondary performance metrics such as assists and rebounds per game. These statistics did prove to contribute to player salaries, but they were significantly less influential than points scored per game. Our regression analysis revealed that each additional assist per game increases a players salary by $791,318.20, while each additional rebound per game increases a players salary by $456,743.60. These results confirm the idea that it is quite important for athletes to have well-rounded performances when it comes to determining salaries, even though rebounds and assists might have less of an impact compared to scoring.

These results provide several important insights that go into salary structure. Scoring remains the most dominant factor when it comes to determining salary, emphasizing the importance of offensive output in the NBA. However, team revenue also has a meaningful role as there are different salary opportunities for players on large-market teams vs. small-market teams. Finally, while secondary statistics such as assists and rebounds have very moderate impacts on determining salaries, they do contribute to a players overall value, reflecting the importance of well-rounded skills.

3. Data

The data-set that was used for this analysis included comprehensive information on individual player performances and team financial metrics for the 2022-2023 NBA season. This data was sourced from Statista and Gigasheet, specifically a player performance dataset that contained individual statistics and a team revenue dataset. Both sets were manually integrated by linking each player to their respective teams' revenue based on their team name in order to have consistency. By having a unified dataset, we were able to conduct a holistic exploration of how individual performance and market factors tend to influence the salaries of NBA athletes. The performance statistics dataset included 467 observations, covering the NBA players from the 2022-2023 season. It provided detailed statistics such as points per game, rebounds per game, and assists per game. Both datasets are credible with the performance statistics data set being derived from official league sources and the revenue data based on financial reports.

Though the dataset is quite robust, there are potential criticisms. The team revenue dataset represents the aggregate annual revenue but does not account for many factors such as variations within the season or differences in team spending. Furthermore, the performance statistics is unable to capture the important qualitative factors such as leadership and off-court image which also serve as significant influences on salaries. Finally, there might be a possibility of errors due to the merging of the two datasets but cross-validation ensured accuracy. In terms of supplemental data, additional information on individual player endorsements or team payroll distributions might have helped enhance the analysis by providing further insight, unfortunately such data was not found.

The dependent variable in this study is player salary, representing annual compensation for each NBA player. The independent variables were average points scored by a player per game, average assist a player provides per game, average rebounds per game, and the annual revenue of the team to which the player belongs (in millions of USD). The mean salary was $9.13 million with a standard deviation of $6.91 million showcasing a clear disparity in player compensation. The mean for points per game was 9.13 points with a standard deviation of 6.91points indicating a decent variation in offensive output. The mean for assists per game was 2.11 with a standard deviation of 2.28, and the mean for rebounds was 3.53 with a standard deviation of 2.28. Finally the mean revenue was $377.97 million with a standard deviation of $102.26 million showing a clear difference in market size across teams across the league.

The data is internally consistent based on the visualizations and descriptive statistics. This dataset provides a reliable foundation for analyzing NBA player salaries and reveals meaningful patterns validating the use for regression modeling.

4. Modeling

In this study we utilized multiple linear regression (MLR) for further examination of the relationship between NBA player salaries (the dependent variable) and key performance statistics such as points, assists, and rebounds per game, and an external factor of team revenue (independent variables). We looked at many models such as Logistic Regression which could have been used if dependent variables were binary but the MLR model was selected as it provided the ability to model the effects of multiple independent variables on one dependent variable simultaneously, which suited this dataset the best given that salary is continuous.

The regression model provided us with an R-squared value of .554 indicating that 55.4% of the variation in player salaries is able to be explained by the independent variables. Across all the determinants, it was observed that points per game (ppg) had the strongest relationship with salary, with a coefficient of $869541.70. The assists per game (apg) also had a significant positive effect, with a coefficient of $791318.2 while rebounds per game (rpg) had a coefficient of $456744.60. Finally, the revenue was smaller but still a statistically significant factor as it had a coefficient of $10636.62.

We also completed residual analysis in order to ensure the accuracy of the assumptions. The residual vs. predicted plot did not have any discernible patterns, indicating homoscedasticity. However, the histogram and quantile plot of the residuals did reveal some abnormalities. Through the Anderson-Darling normality test, we got a p-value of less than .001 meaning the residuals were not normally distributed. Despite all this, the dataset proved to be proper for identifying relationships between salary and statistics/revenue. Not only did we complete residual analysis, but multicollinearity was also evaluated using VIF values. All of the VIF values were below 4 indicating that multicollinearity was not an issue at all. Additionally, we evaluated outliers through studentized residuals and found no server outliers.

The MLR model was quite effective in examining the key determinants of player salaries with ppg being the most impactful predictor. Though there were some slight non-normalities, the primary findings were valid.

5. Summary

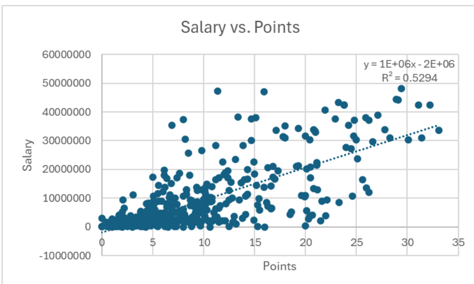
This study took an in-depth dive into the factors that go into influencing NBA athlete salaries during the 2022-2023 season, focusing on performance metrics such as ppg, apg, rpg, as well as external factors such as team revenue. By using MLR, we were able to find that these variables are significant determinants in salary structure, with an R^2 value of .554. The analysis confirmed that the ppg was the most critical influence of salary, followed by assists and rebounds. Team revenue was the least influential and but still did have a meaningful role in shaping salary structures.

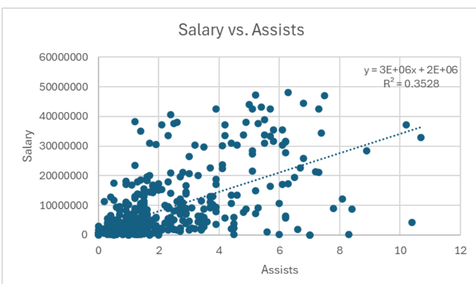
These findings provide actionable insights for various individuals such as team stakeholders, team managers, and even the players themselves. Team managers can use the results in order to identify undervalued players, helping maneuver the salary cap when attempting to form winning teams. For players, they are able to see what they need to work on the most when it comes to trying to scale up their earning potential. Not only that but using this analysis league, officials can inflict significant change regarding the concerns about salary disparities between large and small market teams.

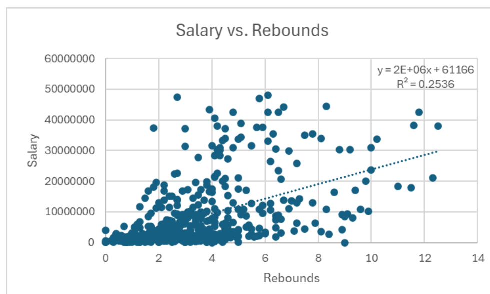
The study does emphasize the utility of regression modeling when it comes to determine drivers behind salary. There are some limitations such as lack of qualitative factors like leadership and public image which were left out of the analysis. Not only that, but the team revenue predictor did not account for in-season fluctuations or differences in spending. Overall, this analysis provided a holistic and comprehensive view of how individual statistics and market dynamics shape salaries.

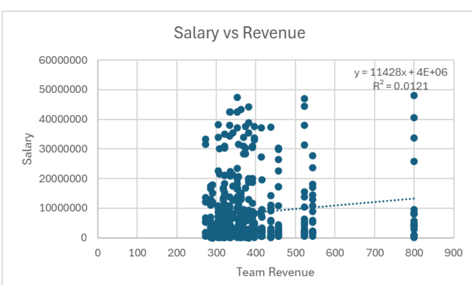
6. Appendix

Scatter Plots -









Regression Output -

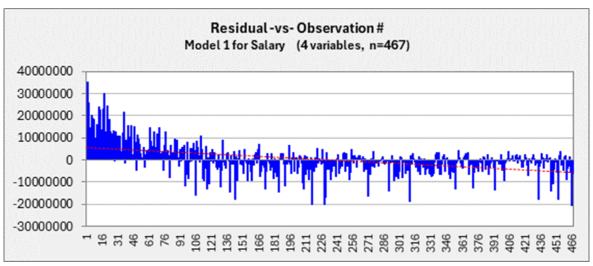
Summary of Regression Model Results	
Linear Model For Salary	Model 1
Run Time	12/6/24 8:34 PM
# Fitted	467
Mean	8,416,598.747
Standard Deviation	10,708,118.047
Number Of Variables	4
Standard Error of Regression	7,180,801.785
R-squared	0.554
Adjusted R-squared	0.550

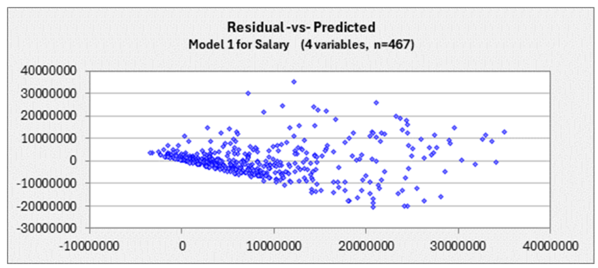
Model: Model 1
Dependent Variable: Salary

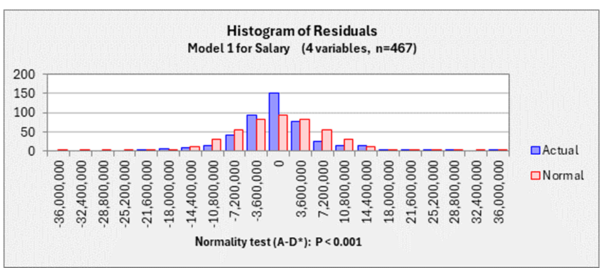
	R-Squared	Adj.R-Sqr.	Std.Err.Reg.	Std.Dep.Var.	# Fitted	# Missing	Critical t	Confidence
	0.554	0.550	7,180,802	10,708,118	467	0	1.965	95.0%

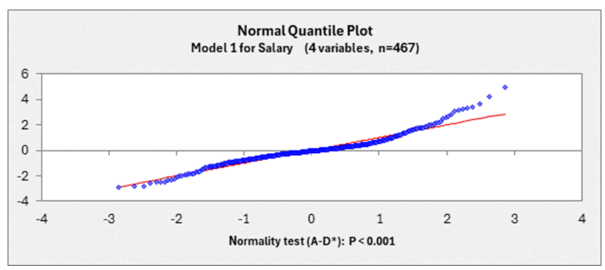
Variable	Coefficient	Std.Err.	t-Statistic	P-value	Lower95%	Upper95%	VIF	Std. Coeff.
Constant	-6,860,853	1,374,048	-4.993	0.000	-9,561,011	-4,160,695	0.000	0.000
AST	791,318	251,529	3.146	0.002	297,036	1,285,600	2.191	0.145
PTS	869,542	83,521	10.411	0.000	705,414	1,033,669	3.007	0.561
Revenue	10,636	3,225	3.298	0.001	4,298	16,974	1.001	0.103
TRB	456,744	187,270	2.439	0.015	88,738	824,750	1.643	0.097

	Mean Error	RMSE	MAE	Minimum	Maximum	MAPE	A-D* stat
Fitted (n=467)	4.946E-10	7,142,257	4,887,823	-20,748,324	35,191,154	1378.3%	9.07 (P=0.000)









K	L	M	N	O
10636.32	456743.6	791318.2	869541.7	-6860853
3225.348	187269.7	251528.6	83520.66	1374048
0.554163	7180802	#N/A	#N/A	#N/A
143.5636	462	#N/A	#N/A	#N/A
2.96E+16	2.38E+16	#N/A	#N/A	#N/A
B4 Revenu	B3 TRB	B2 AST	B1 PTS	B0 Intercept

For every additional rebount per game the salary increases by 456,743.6

For every additional assist per game the salary increases by 791,318,2

For every additional point scored per game the salary increases by 869,541.7

For each additional million dollars in team revenue the salary of players increases by 10,636

The base salary, -6,860,853 when all the other variables are 0 this value is mostly theoretical since player stats are never all zero

Salary = -6680853 + 869541.7 * PTS + 791318.2 * AST + 456743.6 * TRB + 10636.32 * Revenue

R^2 = .554163

This shows us that 55.42% of the variation in salary is explained by the model while the remaining variation is likely due to other factors

Statisitcal Significance -

PTS = 869541.7/83520.66 = 10.41 which is highly significant

AST = 791318.2/251528.6 = 3.15 which is significant

TRB = 456743.6/187269.7 = 2.44 which is significant

Revenue = 10636.32/3225.348 = 3.3 which is significant

Descriptive Statistics -

TRB		*AST*		*PTS*		*Team Revenue*	
Mean	3.528051392	Mean	2.107708779	Mean	9.129979	Mean	377.9667
Standard Error	0.105374382	Standard Error	0.090583072	Standard Error	0.319572	Standard E	18.67068
Median	3	Median	1.4	Median	7.1	Median	354.5
Mode	1.7	Mode	0.5	Mode	5.6	Mode	320
Standard Deviation	2.277159649	Standard Deviation	1.957516738	Standard Deviation	6.906006	Standard D	102.2635
Sample Variance	5.185456066	Sample Variance	3.831871778	Sample Variance	47.69292	Sample Va	10457.83
Kurtosis	1.825123818	Kurtosis	2.418144674	Kurtosis	0.979718	Kurtosis	9.596875
Skewness	1.236664914	Skewness	1.593481687	Skewness	1.220883	Skewness	2.716044
Range	12.5	Range	10.7	Range	33.1	Range	528
Minimum	0	Minimum	0	Minimum	0	Minimum	272
Maximum	12.5	Maximum	10.7	Maximum	33.1	Maximum	800
Sum	1647.6	Sum	984.3	Sum	4263.7	Sum	11339
Count	467	Count	467	Count	467	Count	30

Variable	Mean	Std. Dev.	Median	Min	Max	Count
TRB	3.53	2.28	1.7	0	12.5	467
AST	2.11	1.96	1.4	0	10.7	467
PTS	9.13	6.91	7.1	0	33.1	467
Revenue	377.97	102.26	354.5	272	800	30

CORRELATION TABLE					
	TRB	AST	PTS	Revenue	Salary
Salary	0.503633	0.593971	0.727597	0.110136	1
TRB	1	0.39038	0.618264	0.032311	0.503633
AST	0.39038	1	0.732933	0.00361	0.593971
PTS	0.618264	0.732933	1	0.007071	0.727597
Revenue	0.032311	0.00361	0.007071	1	0.727597

Works Cited

"NBA Players' Pay and Performance: What Counts?" *The Sport Journal*, 19 June 2018, thesportjournal.org/article/nba-players-pay-and-performance-what-counts/.

Steele, Kendal. "The Salary Cap, Small Market vs. Big Market NBA Teams, and the Effects of Covid-19." *The Park Journal*, The Park Journal, 11 Aug. 2020, www.theparkjournal.com/park-insights/the-salary-cap-small-market-vs-big-market-nba-teams-and-the-effects-of-covid-19.

NBA Player Salaries – 2022–23 Season (Gigasheet Dataset)
Gigasheet. Retrieved from
https://app.gigasheet.com/spreadsheet/NBA-Player-Salaries--2022-23-Season--COPY/14d4e324_617f_4257_b625_0128fa74664c

Revenue of National Basketball Association Teams (Statista Dataset)
Statista. Retrieved from
https://www.statista.com/statistics/193704/revenue-of-national-basketball-association-teams-in-2010/