Examining the Impact of Social Media Following on Player Salary in the National Basketball Association: A Multivariate Statistical Analysis

C. Christopher Lee
Central Connecticut State University

Ruoqing Zhang
Central Connecticut State University

Restinel Lomotey
Central Connecticut State University

Jordan Watkins
Central Connecticut State University

Yuxin Huang
Central Connecticut State University

The National Basketball Association (NBA) has the highest average player salary of any professional sports league that exists. In previous studies, it has been found that on-court performance statistics, such as points per game and rebounds per game, are key determinants of NBA players' salaries. This study contributes to the literature by examining the effects of other on-court performance statistics such as real plus minus, field goal percentage, and free throw percentage. In addition, we test the relationship between NBA players' social media following and their salaries. The multivariate analyses in our study show that points per game, real plus minus, and social media following are the most essential three factors determining NBA players' salaries. The implications for NBA research and practice are discussed.

Keywords: NBA, player salary, social media, sport management

INTRODUCTION

The professional sports industry has been booming over the past years. For example, the global sports market reached a value of nearly $388.3 billion in 2020, having increased at a compound annual growth rate of 3.4% since 2015 (Kumar & Bhalla, 2021). Concurrently, professional athletes turn to be one of the highest paid professions. According to the report of sports leagues' average salaries in the 2019-20 season (Statista, 2021), NBA players, on average, earned the highest salary ($8.32 million), followed by $5.30 million in the Indian Premier League for cricket players, $4.03 million in Major League Baseball, $3.97 million in the English Premier League for soccer players, and $3.26 million in the National Football League.

As managerial decisions determining player salaries are fundamental to sport organizations (Mondello & Maxcy, 2009), it is important and necessary to understand what factors are most critical in determining an athlete's salary.

This study intends to explore the determinants of an NBA player's salary. It is commonly assumed that a player's salary depends on prior on-court performance and other variables (Gius & Johnson, 1998). Prior studies have intensively examined the impacts of various on-court performance statistics on an NBA player's salary (Sigler & Sackley, 2000). These studies have found that some on-court performance statistics, such as points per game (PPG), rebounds per game (RPG), and assists per game, play a vital role in determining the salary, despite the heterogeneity of players' positions (e.g., Ertug & Castellucci, 2013; Sigler & Compton, 2018). However, little or no research has been conducted to investigate the role of other performance statistics, such as field goal percentage (FG%), free throw percentage (FT%), or real plus minus (RPM). In addition, other non-performance statistics, such as the number of social media (e.g., Instagram, Facebook, etc.) followers, may play an important role in determining a player's salary since many NBA players are key influencers on crucial societal and cultural issues through social media (Ertug & Castellucci, 2013; Li & Huang, 2015). However, the impact of social media following on NBA players' pay remains empirically unexamined in the sport management literature. To address the research gaps mentioned above, this study intends to answer the following two research questions:

RQ1: *Do NBA players with more social media followers earn higher salaries?*

RQ2: *Which performance-based statistics are the most crucial in determining an NBA player's salary?*

All in all, the purpose of this study is to test the impacts of both on-court performance (i.e., PPG, FG%, FT%, and RPM) and non-performance statistics (i.e., social media following) on NBA players' pay in an integrated model.

LITERATURE REVIEW

On-Court Performance Statistics and Salary

Of various on-court performance statistics, PPG appears to be the most important determinant. The statistic of points is used in basketball games to keep track of the score. Lyons et al. (2015) analyzed the 2013-14 season salaries of 243 NBA players and their career performance variables. They found that PPG, RPG, and personal fouls contributed significantly to a player's salary. Similarly, Stanek (2016) analyzed the relationship between player performance and team revenues using the data from the 2005-06 season to the 2014-15 season, finding that teams tend to overpay players for their PPG compared to other statistics. Therefore, we hypothesize that:

H1: *There is a positive relationship between the statistic of an NBA player's point per game (PPG) and the player's salary.*

Other on-court performance statistics such as FG%, FT%, and RPM are relatively less investigated in predicting a player's salary in the literature. The measure of FG% refers to the ratio of field goals made to field goals attempted, and the FT% statistic is the ratio of a player's successful free throws in perspective to the total attempts. Besides, RPM is a newly introduced basketball metric by ESPN in 2014. RPM describes a player's average impact on point differential per 100 offensive and defensive possessions. RPM is one of the most cited single number metrics when estimating a player's talent and impact. Limited research studies have examined the impacts of FG%, FT%, or RPM on an NBA player's salary, and the studies (e.g., Louivion & Pettersson, 2017; Lyons et al., 2015) have found mixed relationships. As the three metrics are among the most significant statistics that are representative of a well-rounded basketball player (Bertalotto, 2014; Page et al., 2007), we hypothesize that:

H2: There is a positive relationship between the statistic of an NBA player's FG% and the player's salary.

H3: There is a positive relationship between the statistic of an NBA player's FT% and the player's salary.

H4: There is a positive relationship between the statistic of an NBA player's RPM and the player's salary.

Social Media Following and Salary

Hill and Jolly (2012) stated that NBA players' salaries have gradually shifted from performance-based metrics to non-performance statistics after analyzing the changes made in the Collective Bargaining Agreement for the NBA. Similarly, Ertug and Castellucci (2013) found that both reputation and status had a positive effect on an NBA player's salary by analyzing the longitudinal data from the 1989-90 season to the 2004-05 season. Therefore, we hypothesize that:

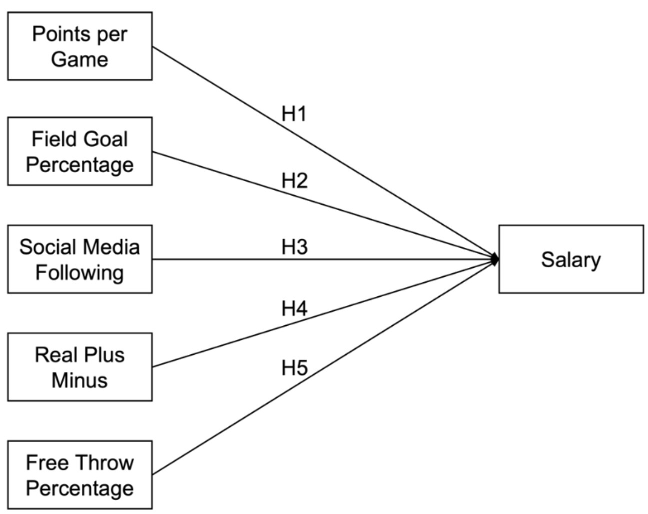
H5: There is a positive relationship between the number of an NBA player's social media followers and the player's salary.

RESEARCH METHODOLOGY

Research Framework

Based on the literature review, we constructed a research model as shown in FIGURE 1 to test the impact of each variable on player salary in the NBA. We used multiple statistical techniques, such as multiple regression analysis, discriminant analysis, and logistic regression analysis, to test and evaluate the hypotheses.

FIGURE 1
RESEARCH FRAMEWORK



Data Collection

In the 2019-20 season, there were 529 registered players in the NBA spread across the 30 NBA teams (https://www.nba.com/). In our study, we collected data for the top 100 highest paid NBA players in the 2019-20 season. Specifically, three statistics such as PPG, FG%, and FT% for the top 100 players were collected from Basketball-Reference (https://www.basketball-reference.com/), a leading website providing comprehensive basketball performance-based statistics; the RPM data of the top 100 players were collected from the database on ESPN (http://www.espn.com/nba/statistics/rpm/_/year/2020); and we manually collected the number of each player's Instagram followers (NIF) as of February 2021.

Regression Analysis

Regression analysis has been widely applied to examine the determinants of professional athletes' salaries (e.g., Lyons et al., 2015; Sigler & Sackley, 2000). Drawing on a theoretical framework from these studies, we proposed the following multiple regression model:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon \quad (1)$$

where Y = Predicted Salary
X_1 = PPG
X_2 = FG%
X_3 = NIF
X_4 = RPM
X_5 = FT%

Discriminant Analysis

Discriminant analysis has been used in numerous sports research studies (e.g., Rubin & Rosser, 2014; Oh et al., 2014; Facca, 2013). In our study, the data was separated into two groups based on the NBA player's salary. The low salary group includes the first quartile of players, and the high salary group includes the fourth quartile of players. The generalized composite canonical discriminant model is expressed in the following mathematical form (Klecka, 1980):

$$\begin{aligned} f_{km} &= \mu_0 + \mu_1 X_{1km} + \mu_2 X_{2km} + \cdots + \mu_p X_{pkm} \\ &= \mu_0 + \sum \mu_j X_{jkm} \end{aligned} \quad (2)$$

where f_{km} = value (score) on the canonical discriminant function for case m in the group k
μ_j = coefficients which produce the desired characteristics in the function
X_{jkm} = the value on discriminating variable X_j for case m in group k
k = 1 for high salary group, and 0 for low salary group
m = the case number of NBA player
X_j = independent variables (j = 1 through p)

Logistic Regression Analysis

Logistic regression models have been used in prior sport studies (Gramacy et al., 2013; Le Crom et al., 2009). We created a new independent variable, Salary2. Salary2 was a (0-1) dummy variable in which 1 = high salary players, 0 = low salary players. For this study, the logistic regression model is expressed in the following mathematical form (Kleinbaum et al., 2002):

$$\begin{aligned} P(X_k) &= P(\text{Salary2} = k \mid X_1, X_2, \cdots, X_p) \\ &= \frac{1}{1+e^{-z}} \end{aligned} \quad (3)$$

where $z = ß_0 + ß_1 X_1 + ß_2 X_2 + \cdots + ß_p X_p = ß_0 + \sum ß_j X_j$
Salary2 = a dichotomous independent variable
k = value of Salary2 (1 or 0)
X_j = independent variables (j = 1 through p)
P(X)= conditional probability of an event k occurring
X = a vector of independent variables

RESULTS

TABLE 1 presents the Pearson correlations between variables for all data (N = 100). The salary variable was significantly correlated with points per game (PPG), the number of Instagram followers (NIF) (in millions), real plus minus (RPM), and free throw percentage (FT%) at the .01 significance level. However, the relationship between field goal percentage (FG%) and salary was insignificant.

TABLE 1
DESCRIPTIVE STATISTICS AND CORRELATION RESULTS

	Mean	*SD*	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
(1) Salary	22,898,965.76	7,641,313.82	1					
(2) PPG	15.458	7.106	.622**	1				
(3) FG%	.4599	.0668	.063	.245*	1			
(4) NIF	3.07069	8.6962	.427**	.263**	.016	1		
(5) RPM	.898	2.642	.613**	.659**	.315**	.463**	1	
(6) FT%	.775	.1067	.328**	.364**	-.388**	.052	.192	1

Note(s): $N = 100$, *$p < .05$, **$p < .01$ (2-tailed).

TABLE 2 summarizes the regression results, in which the dependent variable is the salary of NBA players in the 2019-20 season. The overall model fit was assessed through F statistics. The result of the F test, $F(5,89) = 20.267$, indicated that the regression model was statistically significant at $p < .01$. The variance inflation factors (VIF) was calculated, and the largest VIF was 2.241. Because this value is much lower than the threshold of 10 for the VIF index, multicollinearity was not a problem in our study. For the model regressing an NBA player's salary, three statistically significant variables were PPG ($p = .000$), NIF ($p = .035$), and RPM ($p = .015$). The other two variables, FG% and FT%, did not exert significant effects on salary. Overall, the regression model explained 50.6% (adjusted $R^2 = .506$) of the variance in salary.

TABLE 2
REGRESSION RESULTS

DV = Salary, $R^2 = .532$, Adjusted $R^2 = .506$, $F(5, 89) = 20.267$**						
Variable	B	SE	BETA	*t*	*p*-value	VIF
(Constant)	11,117,717.215	8,001,112.286		1.390	.168	
PPG	410,486.641***	112,578.735	.383	3.646	.000	2.097
FG%	-3,403.849	10,409.157	-.029	-.327	-.327	1.505
NIF	154,395.251*	72,269.262	.183	2.136	.035	1.402
RPM	756,749.638*	306,219.533	.268	2.471	.015	2.241
FT%	6,827.271	6404.471	.097	1.066	.289	1.592

Note(s): $N = 100$, *$p < .05$, **$p < .01$, ***$p < .001$ (2-tailed).

TABLE 3 shows the results of discriminant analyses on the full model and the best fit model, respectively. The analysis on the full model included all the hypothesized variables, while the analysis on the best fit model comprised of the two most significant independent variables of the full model, PPG and RPM. As presented in TABLE 3, the values of the test statistics, such as Eigenvalue and Canonical Correlation, of the full model were greater than those in the best fit model. The full model appeared to be a better goodness of fit for the data. The equation for the discriminant function includes:

$$\text{Discriminant Function} = -.3.815 + .127*X_1 + .002*X_2 + .002*X_3 + .180*X_4 + .001*X_5 \quad (4)$$

where X_1 = PPG
X_2 = FG%
X_3 = NIF
X_4 = RPM
X_5 = FT% * 1000

TABLE 3
DISCRIMINANT ANALYSIS RESULTS

	Full Model		Best Fit Model	
	USCDFC	SCDFC	USCDFC	SCDFC
PPG	.127	.692	.134	.725
FG%	.002	.129		
NIF	.002	.014		
RPM	.180	.396	.194	.428
FT% * 1000	.001	.093		
(Constant)	-3.815		-2.279	
Eigenvalue	.778		.768	
Canonical Correlation	.661		.659	
Wilks' λ	.562		.562	
Chi-square (χ^2)	52.076		52.449	
Df	5		2	
p-value	.000		.000	
Accuracy	77.7		78.1	

Note(s): USCDFC = Unstandardized Canonical Discriminant Function Coefficients; SCDFC = Standardized Canonical Discriminant Function Coefficients.

TABLES 4 and 5 summarize the results of logistic regression analyses for both the full model and the best fit model. As shown in Table 4, for the full model, both PPG (p = .018) and RPM (p = .010) were found to be significant predictors of salary; NIF (p = .051) had a marginally significant impact on salary; there were no significant impacts of FG% (p = .424) or FT% (p = .881) on salary. For the best fit model, both PPG (p = .000) and RPM (p = .006) would significantly affect salary. Table 5 indicates that the full model was a better goodness of fit for the data than the best fit model as the former yielded a higher R^2 value of .646, meaning that the full model accounted for 64.6% of the variance in salary.

TABLE 4
LOGISTIC REGRESSION RESULTS

	Variables	*B*	*S.E.*	*Wald*	*df*	*p-value*	*Exp(B)*
Full Model	PPG	.194	.082	5.556	1	.018	1.124
	FG% * 1000	.005	.006	.639	1	.424	1.005
	NIF	.694	356	3.796	1	.051	2.001
	RPM	.569	.221	6.627	1	.010	1.766
	FT% * 1000	-.001	.004	.022	1	.881	.999
	Constant	-6.105	5.320	1.317	1	.251	.002
Best Fit Model	PPG	.236	.066	12.574	1	.000	1.266
	RPM	.528	.190	7.702	1	.006	1.695
	Constant	-4.255	1.077	15.604	1	.000	.014

TABLE 5
LOGISTIC REGRESSION MODEL SUMMARY

	χ^2	*df*	*p-value*	*-2 log likelihood*	*Cox & Snell R^2*	*Nagelkerke R^2*
Full Model	62.866	5	.000	68.569	.484	.646
Best Fit Model	56.690	2	.000	74.745	.449	.600

DISCUSSION

The results show that PPG is significantly related to salary with $p = .000$, supporting H1. The finding is consistent with prior studies examining the relationship between PPG and players' salaries (e.g., Lyons et al., 2015; Sigler & Sackley, 2000; Stanek, 2016). The finding is not surprise given that points in the sport of basketball is one of the most important statistical fields all teams look for when looking for players. Basketball in the NBA has gained popularity by promoting scoring outputs for games. This means that teams are looking for players who can contribute to increasing the number of points scored for the team per game.

The results fail to support H2, that is, if the NBA player's average FG% increases, his annual salary increases. The results show no statistical significance with the relationship between FG% and player salary in the regression models ($p = -.327$). The inconclusiveness of the results for the relationship between FG% and salary may be attributed to the unpopularity of FG% to the average basketball fan. What teams focus on when they are valuing their players are how much money they can bring in, which translates to how well they are able to perform in the most popular and marketable statistical fields.

The results show that social media following has a positive and significant effect on salary in all the models tested, thus, supporting H3. This finding fills the research gap of examining the impact of social media on NBA players' salaries. The evidence shown in this study indicates that RPM has a positive and significant impact on salary with $p = .015$, supporting H4. RPM can be considered as a player's estimated impact on a team's performance. The higher the RPM that NBA players have on a team on average, the higher salary the players would be paid.

The results show no statistical significance with the relationship between FT% and player salary in the regression models ($p = .289$), failing to support H5. The inconclusiveness of the results for the relationship between FT% and salary may also attribute to the unpopularity of this statistical field to the average fan. Nevertheless, it does not mean that FG% is not an important factor that players should ignore if they are looking for a salary increase.

CONCLUSION

In conclusion, this study has shown that PPG, RPM and NIF exert positive and significant impacts on NBA players' salaries. However, FG% and FT% are the statistics that have inconclusive relationships with salary. The major finding in our study shed light to that how well a player promotes himself on social media can directly lead to a higher pay. The relative importance of the five variables, such as PPG, RPM, NIF, FT%, and FG%, in affecting salary for NBA players is presented in TABLE 6.

For NBA front office personnel, this study suggest that they put more of an emphasis on metrics, such as PPG, RMP, and social media following (e.g., NIF), to develop formulas or models that will lead to better pragmatic methos of compensating players. In situations where NBA players negotiate contracts or seek salary increase, they need to demonstrate that they can score more points a game, bring more social media followers to the team with them, as well as prove that they maintain a satisfactory RPM.

TABLE 6
RANKED VARIABLES AFFECTING NBA PLAYERS' SALARIES

Rank of Variables	*Models*			*Conclusion*
	Regression	Discriminant Analysis	Logistic Regression	
(1)	PPG	PPG	PPG	PPG
(2)	RPM	RPM	RPM	RPM
(3)	NIF	FG%	NIF	NIF
(4)	FT%	FT%	FG%	FT%
(5)	FG%	NIF	FT%	FG%

The results of this preliminary research may be improved by accounting for qualitative measures and using additional performance or compensation data from prior years. Future research could investigate the determinants of salary by conducting a longitudinal research method. Another suggestion would be to include additional independent variables in further research to expand on our findings and assess added determinants of NBA players' salaries. Some potential independent variables include rebounds per game, assists per game, minutes played per game, or player efficiency rating, among others.

REFERNCES

Bertalotto, J. (2014). *Buying wins: An analysis of salary and performance in the National Basketball Association*. Retrieved from https://ir.library.oregonstate.edu/downloads/cz30pv89n

Ertug, G., & Castellucci, F. (2013). Getting what you need: how reputation and status affect team performance, hiring, and salaries in the NBA. *Academy of Management Journal, 56*(2), 407–431.

Facca, T.M. (2013). Using discriminant analysis to classify satisfaction data to facilitate planning in a franchise context. *Journal of Strategic Marketing*, *21*(2), 125–139.

Gius, M., & Johnson, D. (1998). An empirical investigation of wage discrimination in professional basketball. *Applied Economics Letters*, *5*(11), 703–705.

Gramacy, R.B., Jensen, S.T., & Taddy, M. (2013). Estimating player contribution in hockey with regularized logistic regression. *Journal of Quantitative Analysis, 9*(1), 97–111.

Hill, J., & Jolly, N. (2012). Salary distribution and collective bargaining agreements: A case study of the NBA. *Industrial Relations*, *51*(2), 342–363.

Klecka, W.R. (1980). *Discriminant Analysis*. London, UK: Sage Printing Press.

Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M., & Klein, M. (2002). *Logistic Regression*. New York, NY: Springer-Verlag.

Kumar, M., & Bhalla, S. (2021). Global sports market today: An overview. *International Journal of Physical Education*, *Sports and Health*, *8*(4), 223–225.

Le Crom, C.L., Warren, B.J., Clark, H.T., Marolla, J., & Gerber, P. (2009). Factors contributing to student-athlete retention. *Journal of Issues in Intercollegiate Athletes*, *2*, 14–24.

Li, Z., & Huang, K. (2014). The monetary value of twitter followers: Evidences from NBA players. In *Proceedings of the 35th International Conference on Information Systems.*

Louivion, S., & Pettersson, F. (2017). *Analysis of performance measures that affect NBA salaries*. Retrieved from https://www.diva-portal.org/smash/get/diva2:1114463/FULLTEXT01.pdf

Lyons, R., Jr., Jackson, E.N., Jr., & Livingston, A. (2015). Determinants of NBA player salaries. *The Sport Journal*, *18*, 1–10.

Mondello, M., & Maxcy, J. (2009). The impact of salary dispersion and performance bonuses in NFL organizations. *Management Decision, 47*(1), 110–123.

Oh, Y., Kim, H., Yun, J., & Lee, J. (2014). Using data mining techniques to predict win-loss in Korean professional baseball games. *Journal of the Korean Institute of Industrial Engineers*, *40*(1), 8–17.

Page, G.L., Fellingham, G.W., & Reese, C.S. (2007). Using box-scores to determine a position's contribution to winning basketball games. *Journal of Quantitative Analysis in Sports*, *3*(4), 1–18.
Rubin, L., & Rosser, V. (2014). Comparing division IA scholarship and non-scholarship student-athletes: A discriminant analysis. *Journal of Issues in Intercollegiate Athletes*, *7*, 43–64.
Sigler, K., & Compton, W. (2018). NBA players' pay and performance: What counts? *Sport Journal*, *2*, 11–19.
Sigler, K., & Sackley, W. (2000). NBA players: Are they paid for performance? *Managerial Finance*, *26*(7), 46–51.
Stanek, T. (2016). *Player performance and team revenues: NBA player salary analysis.* Retrieved from http://scholarship.claremont.edu/cmc_theses/1257%0AThis
Statista. (2021). *Average annual player salary in the sports industry in 2019/20, by league*. Retrieved from https://www.statista.com/statistics/675120/average-sports-salaries-by-league/