Is There Salary Discrimination by Race and Nationality in the NBA? A New Approach

JiaYou Liang(✉), ShuaiJie Zhao, and HaoYuan Zhu

Beijing Normal University - Hong Kong Baptist University United International College, Zhu Hai, China
q030018038@mail.uic.edu.cn

Abstract. This study seeks to investigate the existence of salary discrimination in the National Basketball Association (NBA). According to Gary Becker's theory of discrimination in labor economics, this study focuses on employer discrimination. While previous research has examined discrimination based on race and nationality separately, this paper combines them to analyze the issue comprehensively. Specifically, the study focuses on whether salary discrimination occurs among American-white, American-nonwhite, other country-white, and other country-nonwhite individuals. Therefore, based on the data collected from Basketball Reference for the 2021–2022 season, a Mincer equation has been constructed to predict the effect of experience, races, nationality, and players' performance on earnings. The OLS regression result provides the evidence that there was no such race or nationality discriminations in NBA.

Keywords: NBA · Salary Discrimination · Mincer Equation · OLS

1 Introduction

As one of the most popular sports in the world, the National Basketball Association (NBA) has been plagued by discrimination, particularly in terms of salaries. Previous research has looked at the differences between white and black players. According to Kahn and Sherer [1], customer discrimination led to a racial gap in salaries for black players in the 1980s, with some teams making discriminatory salary offers because fans preferred white athletes. In the 1990s, Matthew [2] found that the NBA had become a racially equal labor market, and there was no longer a difference between white and black fans. Similarly, when Hisahiro and Yu [3] revisited the issue of a racial gap in the NBA, they found that the racial salary gap had disappeared in the 1990s and 2000s, but another racial gap began to emerge in the 2000s and reached about 20% in the 2010s. They suggested that this emergence may have been caused by the globalization of teams, the introduction of luxury taxes, and minor leagues.

After the 2000s, more and more international players entered the NBA, leading to a new salary discrimination between American and international players. During the 2000s

Liang, J., Zhao, S., Zhu, H—These authors contributed equally.

X. Li et al. (Eds.): ICEMGD 2023, AEPS, pp. 391–399, 2024.
https://doi.org/10.1007/978-981-97-0523-8_38

and 2010s, international players received lower salaries than their U.S.-born counterparts [4]. However, the tide has quickly turned, with Hoffer and Freidel [5] finding that not only have wages for foreign players caught up to those of their American counterparts, but foreign-born players have actually received higher average wages.

We believe that previous research on salary discrimination has not adequately addressed the issue in today's context. Most studies have analyzed discrimination separately based on nationality and race, but our idea is to combine them in order to get a more comprehensive understanding of the issue. Specially, our work focuses on examining whether salary discrimination occurs among American-white, American-nonwhite, other country-white, and other country-nonwhite individuals.

2 Conceptual Framework

Discrimination is a longstanding and significant topic in economics and society. In The Economics of Discrimination [6], Gary Becker discusses discrimination from an economic perspective. The author identifies three main types of discrimination in the labor market: employer discrimination, employee discrimination, and consumer discrimination. Becker's theory suggests that discrimination occurs when employers or workers have a preference for or against a certain group of people, leading to unequal treatment in the labor market. Discrimination can arise from various factors, including prejudice, stereotypes, and beliefs about the abilities and characteristics of different groups. Becker also notes that discrimination can be rational or irrational, and can occur in both competitive and non-competitive labor markets.

Our research focuses on employer discrimination, specifically whether there is a salary discrimination in the sports league. The Mincer equation is a well-known economic model that describes the relationship between an individual's earnings and their education and experience. The equation was developed by economist Jacob Mincer in the 1970s [7] and is widely used in labor economics to predict the effect of education and experience on earnings [8]. To investigate this issue, we use the Mincer equation to build a model and perform OLS regression. Our data comes from the 2021–2022 season and includes player productivity statistics as well as individual characteristics such as height, weight, race, and nationality. We create four new dummy variables by combining race and nationality to further analyze the data: American-white, American-nonwhite, other country-white, and other country-nonwhite. By using the Mincer equation and regression analysis, it is easily for us analyses to determine whether employer discrimination exists in the NBA. Our research is significant because it sheds light on the issue of employer discrimination and provides evidence of its existence or lack thereof in a specific context. By understanding the factors that contribute to discrimination in the labor market, we can better address and combat it.

3 Data

The data for this research project was collected from Basketball Reference, a website that maintains comprehensive statistics on NBA players. In order to obtain information on players' race, we used a combination of methods, including looking at photos, examining

players' family backgrounds, and visiting their individual pages on Basketball Reference. The data pertains to the 2022–2023 season and the dependent variable in the analysis is player salary, which was converted to the natural log in order to avoid issues with heteroscedasticity due to the high range of values.

To control for potential factors that might influence player salary, we included a number of variables in the analysis. These include player age, age squared, team, position, height, weight, games started, and minutes played. We also created several dummy variables to identify whether a player is white, American, or from another country. These dummy variables were used to create three interaction terms: American nonwhite, other country white, and other country nonwhite. The reference category for these interaction terms is American white.

Initially, we calculated the mean salary for white and non-white players in order to assess whether there was any preliminary evidence of discrimination. We found a $400,000 difference in mean salary between the two groups. However, when we examined the average player stats for the two groups, we found that black athletes had higher scores in most categories, with the exception of total rebounds. Given that black athletes tend to have better performance statistics overall, it is not surprising that they would be paid more. Therefore, we cannot definitively conclude that there is discrimination in the NBA based on these findings alone, and further regression analysis is needed.

Table 1. Independent variable value list.

Variables	Count	Mean	Std	Min	Max
GS	408	31.669	30.672	0	154
PER	408	10.091	6.0785	0	36.094
GDPpop	408	599.96	196.84	3.1029	934.329
AW	396	0.1085	0.3115	0	1
AN	396	0.6439	0.4794	0	1
OW	396	0.1111	0.3146	0	1
ON	396	0.1363	0.3436	0	1
Gift	405	0.0958	0.0688	0.003	0.29241
Age	408	25.713	4.3005	19	41
Age2	408	679.62	233.27	361	1281

In total, after we drop the nan value in dataset our regression included 392 observations, which are summarized in the Table 1 "Independent variable value Lists" and described in detail in Table 2 "Data Description."

4 Econometric Models and Testing

4.1 Econometric Models

To study whether there is discrimination against race and country in NBA salary, we introduced three dummy variables in the regression model. We used the interaction term to represent athletes of different races and countries. Considering that the dependent variable is salary, we use the Mincer earnings function as the basic model and further optimize the NBA salary problem on its model. The following are the equations used in the regression model [9]:

$$lnsalary = \beta_0 + \beta_1 GS + \beta_2 PER + \beta_3 GDP_{pop} + \beta_4 AN + \beta_5 OW + \beta_6 ON + \beta_7 Gift + beta_8 Age + \beta_9 Age^2 u \tag{1}$$

The dependent variable is lnsalary, which is the natural log of the players' 2022–2023 salaries. Calculation by natural logarithm, which can reduce the range of salary and make it closer to normal distribution.

The first independent variable GS is the number of starts for the player in the 2021–2022 season. Usually, we think that a player's appearance in the starting lineup reflects their stronger ability and better skills. We generally think that starting players will have higher wages, which can also be understood as a proxy variable of the star effect.

The second independent variable is PER, which is the player's efficiency rating in the 2021–2022 season. Calculated as follows:

$$PER = \frac{(PTS + TRB + AST + STL + BLK) - (2PA + 3PA - 2P - 3P) - (FTA - FT) - TOV - PF}{G} \tag{2}$$

For the PER which is a factor developed by ESPN's John Hollinger to measure a player's per-minute production in the NBA. PER includes all good performances and bad performances of players in the game, which are summed and averaged for each game. We believe that PER is the embodiment of a player's overall strength.

The third variable is GDP Pop, interpreted as the GDP per capita of the athlete's home country. Calculated as follows:

$$GDP_{pop} = \frac{GDP}{Population} \tag{3}$$

It is generally believed that countries with higher per capita GDP tend to have more resources and wealth, which can translate into higher salaries for athletes, including basketball players. This is because these countries may have more disposable income and a greater willingness to spend money on sports, including supporting their national teams and investing in the development of athletic talent. Additionally, higher per capita GDP may also be indicative of a larger market for basketball and other sports, which can lead to higher revenue streams for teams and leagues and ultimately result in higher salaries for players. However, it is worth noting that there are many other factors that can influence an athlete's salary, such as the individual's skill level, popularity, and negotiating power.

Table 2. Data description.

Data name	Description	Unit
Pos	The position of the player on the field	non
Age	Player's age on Feb 1st of the season	years
G	Games	times
GS	Games Started	times
MP	Minutes Played	minutes
FG	Filed Goals	points
FGA	Filed Goal Attempts	times
3P	3-pointed filed goal	points
3PA	3-Pointed Field goal Attempts	times
2P	2-Pointed Field goal	points
2PA	2-Pointed Field goal Attempts	times
FT	Free Throw	times
FTA	Free Throw Attempts	times
ORB	Offensive Rebounds	times
DRB	Defensive Rebounds	times
TRB	Total Rebounds	times
AST	Assists	times
STL	Steals	times
BLK	Blocks	times
TOV	Turnovers	times
PF	Personal Fouls	times
PTS	Points	points
white	= 1, if the player is white	non
country	= 1, if the player is American	non
Salary	Salary of player in 2021–2022	dollars
lnsalary	= ln(Salary)	non
Position	the position that the players were selected	number
Height	the height of player	cm
weight	the weight of player	kg
Age2	square of Age	non
AW	American white	non
AN	American nonwhite	non

(*continued*)

Table 2. (*continued*)

Data name	Description	Unit
OW	Other country white	non
ON	Other country nonwhite	non
Gift	position/(height + weight)	number
GDP	The GDP of one country in 2021	$1 billion
Population	The population of one country in 2021	10 million
GDP Pop	GDP/Population	$10
PER	[(PTS + TRB + AST + STL + BLK)-2PA + 3PA-2P-3P-(FTA-FT)-TOV-PF]/G	points

The fourth to sixth independent variables are dummy variables. AN represents non-white Americans, OW represents non-white Americans, and ON represents non-white non-Americans, so the base of the model is black Americans. We use these three variables to divide the athletes in the NBA into four categories. If a certain category is met, the variable value is 1, and the other dummy variable values are 0.

The seventh independent variable is Gift, which is used to represent the player's physical talent. The lower the Gift value, the better the physical talent we think. The calculation formula is as follows:

$$Gift = \frac{Position}{Height + Weight} \tag{4}$$

The eighth independent variable and the ninth independent variable are age and the square of age respectively. According to the thinking of the Mincer earnings function, we believe that wages will show a trend of rising first and then falling with age. The same is true in terms of subjective understanding. Athletes will reach a peak period as they grow older, and their performance will decline with age, and their salary will also decline.

To summarize, our research aims to investigate whether there is salary discrimination in the NBA based on race and nationality. We use the Mincer equation as a foundation and incorporate various independent variables that we believe may impact salaries. Our results will provide insight into the prevalence of discrimination in the NBA and may help to inform policies and practices to address such discrimination in the future.

4.2 Testing

We conducted several statistical tests to assess the validity and reliability of our multiple linear regression (MLR) model. First, we checked for collinearity among the independent variables by performing a correlation test. The results which are shown in the following Table 3.

Next, we conducted a RESET test to check for misspecification in the model. The RESET test compares the fit of a full model, which includes all the independent variables,

Table 3. Correlation.

Variables	lnsalary	GS	GDPpop	PER	AW	AN	OW	ON	Gift	Age	Age ^2
lnsalary	1	0.61	0.08	0.74	0.07	0.08	0.01	0.06	0.43	0.31	0.29
GS	0.61	1	0.09	0.07	0.11	0.11	0.01	0.03	0.27	0.15	0.13
GDPpop	0.08	0.09	1	0.03	0.17	0.64	0.53	0.56	0.04	0.02	0.02
PER	0.74	0.7	0.03	1	0.11	0.08	0.02	0.03	0.42	0.17	0.16
AW	0.07	0.11	0.17	0.11	1	0.47	0.12	0.14	0.07	0.03	0.01
AN	0.08	0.11	0.64	0.08	0.47	1	0.48	0.53	0.01	0.02	0.01
OW	0.01	0.01	0.53	0.02	0.12	0.48	1	0.14	0.02	0.06	0.06
ON	0.06	0.03	0.56	0.03	0.14	0.53	0.14	1	0.03	0.05	0.05
Gift	0.43	0.27	0.04	0.42	0.07	0.01	0.02	0.03	1	0.03	0.02
Age	0.31	0.15	0.02	0.17	0.03	0.02	0.06	0.05	0.03	1	1
Age2	0.29	0.13	0.02	0.16	0.01	0.01	0.06	0.05	0.02	1	1

with a reduced model that omits certain non-significant variables. The p-value of the RESET test was 0.0664, which passed the test at a confidence level of 0.05. This suggests that the model is reasonable and does not suffer from misspecification.

Finally, we performed the White test and the BP test to check for heteroscedasticity in the model. Heteroscedasticity occurs when the variance of the error term is not constant across the range of the independent variables and can lead to biased and inconsistent estimates of the regression coefficients. The p-values of both the White test = 0.1951 and the BP test = 0.3588 were above the 0.05 confidence level, indicating that there is no heteroscedasticity in the model.

Overall, the results of these tests suggest that the MLR model is valid and reliable for making predictions or inferences about the relationship between the independent variables and the dependent variable.

5 Result

Table 4 shows the OLS model estimates for the 2021–2022 NBA season. As can be seen from Table 4, the model is significant, with a F-statistic of 77.24. And the model with basic variable AW equation is:

$$lnsalary = 9.4929 + 0.0070GS + 0.0004GDP_{pop} + 0.0001AN + 0.06980OW + 0.0286ON - 3.1944Gift + 0.3326Age - 0.0052Age2 \quad (5)$$

The coefficient for games started (GS) was positive and statistically significant, indicating that players with a higher value in this variable may receive higher salaries. This could be due to the"star" effect, where the starting lineup of each team is perceived as having a stable level of productivity and receives a lot of attention from the public. As a result, teams may be willing to spend more to retain these players in order to improve their chances of winning and attract more fans. Player efficiency rating (PER)

Table 4. Regression results.

Factors	OLS(AW)	OLS(AN)	OLS(OW)	OLS(ON)
GS	0.007(0.002)	0.007(0.002)	0.007(0.002)	0.007(0.002)
PER	0.0886(0.008)	0.0886(0.008)	0.0886(0.008)	0.0886(0.008)
GDPpop	0.0004(0.0003)	0.0004(0.0003)	0.0004(0.0003)	0.0004(0.0003)
AW	-	–0.0001(0.114)	–0.0698(0.193)	–0.0286(0.182)
AN	0.0001(0.114)	-	–0.0697(0.167)	–0.0285(0.155)
OW	0.0698(0.193)	0.0697(0.167)	-	0.0412(0.139)
ON	0.0286(0.182)	0.0285(0.155)	-0.0412(0.139)	-
Gift	–3.194(0.569)	–3.194(0.569)	−3.194(0.569)	–3.194(0.569)
Age	0.3326(0.091)	0.3326(0.091)	0.3326(0.091)	0.3326(0.091)
Age2	–0.0052(0.002)	–0.0052(0.002)	–0.0052(0.002)	–0.0052(0.002)
constant	9.4929(1.225)	9.4931(1.203)	9.5627(1.203)	9.5216(1.198)
n	392	392	392	392
R2	0.645	0.645	0.645	0.645
Adj. R2	0.637	0.645	0.645	0.645
F-statistic	77.24	77.24	77.24	77.24

were statistically significant and related to player performance which imply that that one player can receive a larger wage from his great performance.

However, the result of GDP Pop coefficient differs from the conclusion of Yang & Lin. We found that the GDP Pop has a positive impact on salary but it's not statistically significant. We also included a factor called"Gift" to measure player talent, including draft pick, height, and weight, with a lower Gift indicating greater talent. The coefficient for Gift was negative and statistically significant, which aligns with the definition of the variable. Based on the Mincer equation, we also included Age and Age2 in the model, both of which were statistically significant and had reasonable signs. The Age factors suggest that a player's salary may increase as they get older, but may decrease after reaching their peak physical fitness. It is worth noting that the OLS model only takes into account the variables included in the analysis and may not capture all factors that influence a player's salary. Further research could explore other variables that may affect salary in the NBA.

Moreover, the results of the OLS model estimate for the 2021–2022 NBA season suggest that there is no salary discrimination in the league. This conclusion is based on the coefficients for each dummy variable, which were all statistically insignificant. This finding differs from previous research, which may be due to the"winner's curse" disappearing in recent years. It is possible that teams have gained more experience in evaluating the talent of players, including international talent, and as a result, discrimination has decreased [10]. Additionally, society has become more focused on promoting

ethnic equality, and managers may be more aware of the consequences of discrimination. The banishment of Donald Sterling from the NBA serves as an example of the league's commitment to addressing discrimination and promoting inclusivity. Overall, it is important to continue efforts to eliminate discrimination and promote fairness in the NBA and other sports leagues.

6 Conclusion

Salaries discrimination in the National Basketball Association (NBA) has been a long-standing issue. In order to better understand the problem of salary discrimination in the NBA, we conducted research using data from the 2021–2022 season and the Mincer equation to build a model and perform OLS regression analysis. Our data include player productivity statistics as well as individual characteristics such as race and nationality. The result shows that there was no salary discrimination caused by race and nationality in the NBA based on the coefficients for the dummy variables in our OLS model, which were all statistically insignificant. While our research provides evidence that salary discrimination does not currently exist in the NBA, it is important to continue monitoring this issue and addressing any potential instances of discrimination that may arise.

References

1. Kahn, L.M., Sherer, P.D.: Racial differences in professional basketball players' compensation. J. Labor Econ. **6**(1), 40–61 (1988)
2. Dey, M.S.: Racial differences in national basketball association players' salaries: a new look. Am. Economist, **41**(2), 84–90 (1997)
3. Naito, H., Takagi, Y.: Is racial salary discrimination disappearing in the NBA? Evidence from data during 1985–2015. Int. Rev. Appl. Econ. **31**(5), 651–669 (2017)
4. Yang, C.H., Lin, H.Y.: Is there salary discrimination by nationality in the NBA? Foreign talent or foreign market. J. Sports Econ. **13**(1), 53–75 (2012)
5. Hoffer, A.J., Freidel, R.: Does salary discrimination persist for foreign athletes in the NBA?. Appl. Econ. Lett. **21**(1), 1–5 (2014)
6. Becker, G.S.: The Economics of Discrimination. University of Chicago Press, Chicago (1971)
7. Chiswick, B.R.: Jacob mincer, experience and the distribution of earnings. Rev. Econ. Household **1**, 343–361 (2003)
8. Heckman, J.J., Lochner, L.J., Todd, P.E.: Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In: Hanushek, E., Welch, F. (eds.) Handbook of the Economics of Education, vol. 1, pp. 307–458. Elsevier, Amsterdam (2006)
9. Chiswick, B.R., Mincer, J.: Experience and the Distribution of Earnings. Springer, New York (2006)
10. Eschker, E., Perez, S.J., Siegler, M.V.: The NBA and the influx of international basketball players. Appl. Econ. **36**(10), 1009–1020 (2004)