Does Racial Discrimination Exist in the NBA?

Logan Kendall

University of Akron

Department of Economics

Senior Project

Spring 2020

Acknowledgement

I would like to thank Dr. Francesco Renna and Dr. Dinkar Kuchibhotla for their contributions to this project. I would also like to thank Dr. Michael Nelson, Dr. Elizabeth Erickson, and the rest of the Economics Department for everything they've done for me during my undergraduate career. Without the help from all these people this project would not have been possible.

Abstract

This paper aims to examines the affect of race on compensation for NBA players. Based on prior research, I pose the hypothesis that racial discrimination does exist in the NBA with there being a premium for white players. A second hypothesis I present based on the previous research is that the premium for white players exist for only some groups of players and for other groups there with not exist discrimination based on player race. By using career data from every player who played in the 2019-2020 NBA season and their 2019-2020 salaries, a few different models where ran. The OLS showed that as a whole, there is racial discrimination as black players made around 15% less than white players at a 0.1 significance level. The same model was ran, but as a fixed effect model accounting for team each player plays for. This model found that there was no evidence that discrimination exists. This showed that only some teams discriminate, and only certain players see discrimination. To look at my second hypothesis, three more fixed effect models were ran. One showing only rookie players and one showing low value adding or marginal black players found evidence that black players see discrimination in pay. The model showing high value adding black players found no evidence of discrimination, but possible reverse discrimination as there was a black premium of around 46%. These results show that there is evidence of racial discrimination in the NBA. It's likely from employer discrimination, but further studies would have to be done to determine if this is the most likely cause as they could look at ticket sales, city demographics, and other data surrounding the consumers of the NBA to see if there's evidence of consumer discrimination.

Table of Contents

I. Introduction

The NBA is a multi-billion-dollar industry. As of the 2019-2020 NBA season, the average NBA team was valued at $2.12 billion with the New York Knick leading the list with a net worth of $4.4 billion. These values are up 14% from last season (Forbes, 2020). The constant rise in value and popularity in the NBA can be attributed to many things such as ownership, management, globalization, the fans, and most importantly, the players. But do all the players that make the NBA reap the benefits of the multibillion-dollar business equally? What is the effect of race on an NBA player's compensation? This is a question that many people have looked to answer over the years but have come up with very different answers.

This paper aims to look at how the race of NBA players affects the salaries of players to see if discrimination exists in the NBA after controlling for the players statistics. The NBA has been around since the 1940s with their first season being in 1949. This was a time where many minority groups faced adversity and had to fight for basic rights daily. NBA players, specifically the black players began to feel uneasy regarding wages in the 1990s. Some players believed they were being discriminated against as they thought marginal white players (players that aren't viewed to be as good or important to a team or low valuing add) were receiving similar pay to the better black players. Some people even believe this to be one of the main reasons for the 1998 NBA lockout (Kahn, 2012). This time period is also where we can find the first studies on this issue [(Kahn et al, 1988) and (Dey, 1997)]. There are a few reasons why this is an interesting topic. The first is that there have been many studies and multiple different findings. Some of the early studies found evidence that there is a premium for white players, but then other studies slightly tweak the methodology and found that there are no pay differences. As the years go on some studies confirm previous findings, but others find the exact opposite. More recent studies

have even found that there is reverse discrimination as there is a premium for black players. So, there's not a clear answer for this question.

The other factor that makes this topic still relevant today is related to some recent restructuring within the NBA. The most recent study on this topic (Groothius et al, 2018), used data from the 2012-2013 season. However, in 2014 two significant things happened. The first was current NBA commissioner Adam Silver took over for David Stern. The other was on incident involving the L.A Clipper's owner at the time, Donald Sterling. Phone conversations were released where he was using many racial slurs and he was also known for making racial comments in public. When this happened, Adam Silver acted fast. He fined Sterling $2.5 million, forced him to sell the team, and banned him permanently from the NBA. As Adam Silver took over and handled this act of racism as well as one could, he earned a lot of respect from the players as the commissioner and as a leader. By evaluating current players data, we can see if how Silver acted in 2014 changed the way in which ownership compensates players (Zillgitt, 2014).

Today's NBA is a great place for black players to thrive. A study by Kahn (2005) showed that 80% of the league was black and 37 out of the 42 contracts over $10 million belonged to black players. This is still true today as 80% of the league is black, and the black athletes are dominating the game. For example, in the upcoming 2020 NBA All-Star Game, only 3 of the 25 players selected were white ("NBA All-Star 2020: Roster.").

As described in Becker's theory, there are 3 types of labor discrimination, employer (Owners, GMs, etc.), employee (teammates), and customers (fans). By looking at player salaries, I am making the assumption that employer discrimination is the main force behind pay differential. Based on the action Silver took in 2014, one would expect all levels of employer

discrimination to have diminished over the years. But has it? I plan to further previous research to try to determine whether we can with confidence say that there is possible employer discrimination in the NBA by analyzing player data to see if there are significant pay differences among players of different races after controlling for performance measures.

II. Literature Review

The topic of discrimination in the NBA has been extensively investigated. One of the interesting things about this topic is that there have been some contradicting results. Most papers included players stats such as points per game, assists per game, rebounds per game, steals per game, etc. Some used single season stats while others used career stats. Another similar variable was year drafted. Player salary was the dependent variable for these studies. Most studies looked at the difference between black and white players' salaries [(Kahn et al, 2005), (Groothius et al, 2013), (Dey, 1997)]. Other studies looked at pay differences between foreign born and U.S born players [(Naito et al, 2017), (Groothius et al, 2018), (Raburn, 2016), (Hoffer et al, 2016).

One of the first studies found that overall, there was significant evidence that there is a premium for white players if you use career earnings. But if you divide the players into different skill tiers, the results are different. For rookies and free agents (marginal players) they found that there was a white premium, but for players that were at least 3 years into their careers, there was a black premium. Since black players have overall better stats than white players, an NBA GM or owner can't afford to discriminate against good black players because they are the ones that make the win. But since marginal players don't contribute that much, they are discriminated against. At the time, white players had a better alternative than black players to playing in the NBA. They could argue for better pay because they could possibly get paid more to do

something else. Where the black players would do this because they didn't have a better substitute. This is known as the Nash Bargaining Model (Kahn et al, 2005).

One factor that leads to some studies finding career salary discrimination is career length. Once this is factored in, one can conclude that the white premium exists because white players have longer careers. This would look at exit discrimination, which would be customer discrimination (fans). One study found that there was no pay discrimination between white and black players because the difference could be explained by white players playing for a few extra seasons. This means that the per season salaries were the same, so there was no discrimination by the owners or GMs (Dey, 1997).

The same analyses were done between foreign born and non-foreign-born players. After controlling for player statistics, the foreign players and U.S born players did have a difference in salaries. But it was by such a low amount that it was considered economically insignificant (Raburn, 2016). Other studies have used slightly different variables and got different results. For example, Hoeffer et al, added a college variable to account for foreign-born players that came to the U.S for college. The reasoning is that a foreign-born player may be not be considered as good as a U.S player because some experts say that overseas competition isn't as competitive as NCAA teams. So, this accounts for players that are foreign-born, but have also experience the competitiveness of NCAA basketball. They also used single season salaries and not career earnings (Groothius et al, 2018).

This leaves an opening for new research. There have been studies aimed to see if there was significant evidence for racial discrimination after controlling for player statistics. There have been studies that look at the demographics of the area a team is in and if that effects pay differences between black and white players. There have also been studies looking to see if black

players of different groups (such as rookies and non-rookies). None have done all three. I will be doing these in this paper, plus redefining certain things from previous studies such as what a marginal player is. But first, is there significant evidence that racial discrimination exists in the NBA?

III. Theoretical Model

Discrimination is a topic that has been studied for a long time. One of the first to develop a model for discrimination in the workplace was Gary Becker in *The Economics' of Discrimination* (1957). An important point he makes is that in perfect competition, discrimination can't exist or at least will start to go away. The reasoning behind this is that competitive firms simply can't afford to. If a company has a taste for discrimination or doesn't like to hire people of a certain race, they must pay for it. This would be called a premium. A firm with a taste for discrimination will hire white workers at a higher wage and not hire any of the black workers at lower wages. Not only are they not hiring enough workers, but they are also pay more for the same output. Since white and black workers are substitutes, if a firm pays one more than the other (white firms) this alone means they aren't maximizing profits because if they didn't discriminate, they could get the same output for cheaper. Then there are also firms that will only hire black workers. They will also not be maximizing profit because they also don't have enough workers. Then there are "color blind" firms. They will hire any employee regardless of race. These companies in the long run will be the only firms that will make a profit as they are the only ones hiring the correct amount of workers for the correct salaries. So, if this were to happen, a color blind firm would eventually be able to buyout all of the competition. This is known as employer discrimination.

For the NBA, there are two arguments that it may not be perfectly competitive. The first is from the standpoint of team competitiveness. If you take out last season, the NBA Finals consisted of only two different teams for 5 years (2014/15 – 2018/2019). The other argument would be from team values. The top few teams are worth way more than other teams. For example, the New York Knicks, as mentioned before, are worth $4.6 billion. And only the top 11 teams are worth more than $2 billion.

The author argues that there are three ways in which a person can see discrimination in the labor market. The first which has already been covered is employer discrimination. This would be discrimination by one's boss, manager, or any higher up. The second type of discrimination would be employee discrimination. This would come from co-workers. This is seen in the form of segregation. The final form is customer discrimination. This is when customers have a taste for discrimination. This is when a customer will refuse to buy a product from or order a service from a certain group of people. When this happens, it leads to employers not hiring that certain group of people. This is why it is sometimes hard to determine if discrimination is caused by employers or customers as customer discrimination causes employers to act.

Discrimination isn't hard to define, but what some economists have found is that it isn't the easiest to measure in the labor market. It's easy to determine the race, gender, age, etc. of a worker, but how can you control for worker productivity? You can control for position, education, and experience but this still does not perfectly catch laborers productivity. If two people have the same position, same education, and same experience but belong to different races and differ in productivity that doesn't equate to discrimination if the wages are different. This is true as long as the productivity gap isn't smaller than the salary gap. This just means one

worker outperforms the other. This is why we look to sports. It is much simpler to determine productivity when dealing with sports. All players statistics are recorded and publicized which makes it easy to compare players. The stats cannot be biasedly reported, so by using NBA players it would be more straightforward which players outperform others. This way it would be a clear answer whether or not a player with the same production output are compensated equally regardless of race.

When looking at the NBA, employee discrimination would be found in the compensation of the players. If certain racial groups were discriminated against by their teammates, you might find that member of that group would be underreported at a certain position. If it were customer-based discrimination, you would look at ticket sales, attendance, television views, merchandise sales, a player's career length, etc. This would be found by looking at data regarding the fans and people in the area.

For this project, we will focus on the first type of discrimination. This would be employer discrimination. In the case of the NBA this would be between owner/GM and players. If discrimination were to exist, the observation would be a certain group of people receive less in wages than another group even after controlling for the players stats. So, in other words, if two players with similar statistics but different races have unequal pay that could be due to employer discrimination or customer discrimination. For this paper, we will not be looking at customer data, so it will be assumed to be employee discrimination. Further studies could prove that the racial discrimination is customer based. But if this paper shows there is evidence of discrimination, there will only be evidence of employee discrimination. I will be estimating the effect of a players race on their salaries. The first models will be a normal OLS and a Fixed Effect Model that accounts for the different teams. Previous studies found that only marginal

players, or players of average statistics, were the only ones discriminated against. This was because a team can't be competitive without paying the best players what they are worth. And as can be seen by All-Star rosters and league statistic leaders, many of the top players happen to be black. Becker's theory aligns with this as well. He said in a competitive market, a firm cannot discriminate because they will make no money in the long run. So, by considering the team each player plays for, we can see if this holds true.

IV. Data

The data on player statistics was gathered from Basketball Reference. Basketball Reference is a website that collects past and present data on NBA players. They have customizable data tables that can be downloaded as Excel Spreadsheets. There was one element that you could not download which was race. I'm an avid NBA fan, so most of the time I knew the players that were black or white. When I wasn't sure, I visited their Basketball Reference webpage and looked at their picture.

The variables refer to the 2019-2020 season. The dependent variable is salary for the season, which was converted to its natural log. This was done to avoid any issues that might arise for having such a high distribution, such as heteroskedasticity. The control variables are; points per game, win shares (advanced statistic that evaluates a number of wins the player can be held responsible for), rebounds per game, experience in the NBA, a created variable for experience squared because the experience variable will not have a linear relationship, steals and blocks per game, player position, and games played.

I also created a few dummy variables. There was a dummy variable created for each team so that a fixed effect model could be ran to account for all of the different NBA teams. I made one to identify if a player was a first-round pick (first_rd), one to identify second round picks (second_rd), and one to identify undrafted players. There are dummy variables for both white and black players as well. Another group of dummy variables I created was for position. I made a dummy variable for guard, forward, and center. The last set of dummy variables was created to split up the player types. Previous studies have defined a marginal player as one that is currently a free agent and not on a team. But for this study, a marginal player is defined as a player with more than 4 years of experience in the NBA and averages less than 10 points per game. This will represent the group of players that aren't very essential for a winning team. I chose to define them this way because a player needs a few years to decide whether or not they would be considered a marginal player. For example, reigning MVP Giannis Antetokounmpo averaged at least 27 the last three seasons. But in his first season, he averaged only 6.8 (Basketball-Reference.com). Since we don't want to consider a player that may not have adjusted to the NBA yet, I chose to make the cut off over 4 seasons. Then as far as 10 points per game, most basketball fans would agree that a player that doesn't even score more than 10 would be considered a role/bench player. These players are very interchangeable and aren't as important as the starters or stars of the team. The final dummy variable is to account for players that add the most value to their teams. These are defined as players that average more than 18 point per game. I chose to define them as players of more than 18 points per game because again, as a fan, most other fans would agree the best players on a team or players with max contracts will average at least 18 points per game.

There were also two extra data sets that were created. The first is players on their rookie contracts. This is because there is a rookie pay scale. This means that for the first 3-4 seasons (the 4th year is a team option to keep the player under contract) salary is solely dictated by their draft position and is not at all effected by their NBA stats. The other data set is the rest of the NBA.

To first see if there is preliminary evidence of discrimination, I calculated the mean salary gap between white and black players. There was a $400,000 difference. After using a paired two-sample t test, I fail to reject the hypothesis that there may not be significant evidence to say that these are different. But I would still like to further investigate this as a paired t-test doesn't take into account the difference in player stats. If the output of two players is different, and the salaries are also different this doesn't conclude that discrimination exists. For example, if a black player doesn't play as well statistically as a white player and makes less, this isn't due to discrimination. The salary differences are from one player being better than the other. But looking at the averages in player stats, the black athletes on average have higher player stats in all areas except in total rebounds. They also on average make less. By just looking at this, it appears as though discrimination exists in the NBA. By using the 2019-2020 salaries converted to their natural log for the dependent variable and career stat data, position data, experience data, and race data I will further investigate whether there is significant evidence that employer discrimination still exists in the NBA.

V. Empirical Model

To examine whether there is evidence that discrimination exists in the NBA an OLS regression was ran. The equation was a Mincer Equation. Jacob Mincer came with a function today known as the Mincer Equation. He theorized that a person's wage is related to their education, experience, the square of their experience (experience is a non-linear function), and output or other performance variables. In the case of output, for the NBA that would be a player's statistics. The following is the equation used in the regression model.

Equation: *lnsalary = β_0 + β_1 (second_rd) + β_2 (undrafted) + β_3 (black) + β_4 (Forward) + β_5 (Center) + β_6 (PTS) + β_7 (TRB) + β_8 (AST) + β_9 (STL) + β_{10} (BLK) + β_{11} (MP) + β_{12} (WS) + β_{13} (Experience) + β_{14} (exp2) + ε*

The dependent variable is lnsalary. This is the natural log of the players 2019-2020 salary. The salary value was converted to natural log because this condenses the distribution of salaries. This is a way to avoid heteroskedasticity. This transformation leads to the coefficients for all the variables to represent a percentage and not an actual value.

The next variable is second_rd. This is a binary dummy variable for players that were selected in the second round of the NBA draft. This variable is expected to be a negative value. The baseline is players drafted in the first round, a player drafted in the second round would make less.

Undrafted is similar to second_rd, but it represents players that were not drafted in the NBA. These would be players that were not picked in the drafted, but later signed as free agents. Since teams didn't view these players as good enough to be drafted, the expected coefficient for this variable is negative to a degree even more than that of second_rd.

The next variable is black. This is a binary variable representing black players. This is the main variable in this study. If it comes out as a negative value, this could mean that there is racial discrimination in the NBA. If it isn't negative, then it could mean that there doesn't exist racial discrimination in the NBA. After the research that has been done on this topic, the expected result would be for the value to be negative, meaning there does exist racial discrimination in the NBA.

The next variable is Forward and Center. These are both dummy variables. The baseline represents players that play the guard position. Previous studies found that the guard position is viewed as a leadership role. Based on that it might be expected that both variables be negative. But it wouldn't be surprising to see these variables be positive since this was only a theory and nothing that has been proven.

The next variable is PTS. This is points per game. This is expected to be a positive value. Since the more points per game a player scores, the better that player is. I anticipate that this will be one of the more significant variables. Points per game is one of the main stats many people look at when comparing players.

Then next variables are TRB (total rebounds per game), AST (assists per game), STL (steals per game) + BLK (blocks per game). These are the rest of the player statistic variables. All of these variables are expected to be positive as the higher any of them are, the better the player is expected to be. This would result in higher pay.

The next variable is MP. This is minutes per game. This variable is expected to be positive as well. Better players tend to play more minutes. The more minutes a player plays, the more they are expected to make.

Then next variable is WS. This stands for win shares. This is another variable that is expected to be positive. Win shares are the number of wins a player can individually be responsible for. It is calculated using multiple player statistics. So better players have higher win shares. Players with higher win shares would be expected to get paid more annually.

The last two variables are experience and exp2. These stand for the number of seasons a player has played in the NBA and that same value squared. The experience must be squared because Mincer theorized that experience is a key factor for determining wages and it is a non-linear function. So, for accurate results, the experience variable must be square. The expected value for experience is positive because the more seasons a player plays the more money they tend to make as they get better. But experience is a non-linear function. This means that the function is a parabola with the curve being positive until it peaks at the center and then goes negative. This is because, in terms of basketball, years of experience can only add value to a player until a certain point. At this point the player usually peaks statistically. Then they begin to decline. In other words, a player doesn't keep getting better year after year. At a certain point, the player begins to digress. That is why the expected value for exp2 is negative, it is used as a correction for the experience variable.

A Fixed Effect Model was also ran. This model is the same as the initial model but has 29 team dummy variables. This is used to account for the team each player plays for as some teams may be color-blind while others aren't. Some teams may not care about staying competitive and discriminate against black players or some teams could have a higher value therefore can afford to discriminate in the short run. A fixed effect model can account for there being different teams, some being color-blind and some having a taste for discrimination.

VI. Empirical Results

The results of the initial regression can be seen in figure 2 of the Appendix. The regression equation is the following:

lnsalary = 13.27 -0.09 (second_rd) -0.35 (undrafted) -0.15 (black) + 0.06(Forward) + 0.26(Center) + 0.03(PTS) -0.04(TRB) + 0.03(AST) + 0.12(STL) + 0.12(BLK) + 0.04(MP) + 0.01(WS) + 0.30(Experience) -0.02(exp2) + ε

The number of observations is 455, 364 black players and 91 white players. The F test shows that the model as a whole is significant as the F value was high. The p value is less than .0001. This means that the hypothesis that the model as a whole is insignificant is rejected, or in other words that the model as a whole is significant. The root mean square error is 0.67. That means that the average error for the predicted dependent variable was incorrect by 0.67. Relative to the intercept of 13.27, this is good value for the root mean square error. The R^2 and adjust R^2 were 0.67 and 0.66 respectively. This means that the model can account for $^2/_3$ of the model's variation.

The variables that were statistically significant at a 0.01 significance level were undrafted players, minutes per game, win shares, experience and experience squared. Black players variable and center variable were significant at a .1 significance level. One reason that the player statistics could have been found to not be statistically significant is because of multicollinearity. All of the statistics were correlated to win shares and points per game (correlation coefficients can be seen in Table 3 of the Appendix) . This makes sense because if a player is good, they probably have high stats in points per game, win shares, and at least one other area. Minutes per game is the same, the more minutes the higher stats a player will have. The Forward and Center variable were also correlated to a few stats. For example, forward was correlated to assists. This

is because players of the guard and forward position are known to get assists as centers aren't. Centers are tall, so they get blocks and rebounds. Both of these were correlated to the center variable. I believe these variables are still important to the model and in previous studies these variables were still used.

To makes sure the correlation of these variables didn't affect the main variable which was the black player variable another regression model was ran omitting the correlated variables. When this was done, the sign of the black variable was still negative. This means that using these variables don't change the outcome of the results, therefore they should be kept in the model.

The expected signs of the coefficients were mostly correct as most of them were positive. The player statistics were almost all positive, which would be expected. The only ones that weren't were Forward, Center, and rebounds per game. As far as Forward and Center, I'm not surprised. I just hypothesized that they would be negative based on my knowledge of the NBA. It seems that most of the best players are point guards. If you look at the NBA All-Star rosters of recent years, there are almost no centers on them. The rebounds on the other hand was interesting that it wasn't positive. One reason I think this could be is because it was highly correlated to the center variable. So, the center variable could contain part of the rebound variable which would also explain why its coefficient was relatively high compared to the Forward variable.

The key variable for this study is the black player variable. The coefficient for this variable was -0.15. This is interesting because it is statistically significant at a 0.1 significance level. This can be interpreted as when controlling for player statistics, black players make 15% less. Recalling from earlier, the paired t-test showed that there wasn't a difference between white and black players' salaries. But a paired t-test doesn't account for players statistics. So just looking at the averages, black players only were making $400,000 less, which doesn't seem like

too much since the average salary for both is around $8,000,000. But since the regression model takes the stats into consideration, the results are different. For example, black players make slightly less, but they also have higher statistics in every statistic except for rebounds. So, based on this, black players have better stats but make less. This is why the regression model found there to be a significant difference in pay. An example would be a white player that makes $10,000,000 and a black player that makes $9,500,000. At a first glace it doesn't seem like the salary discrepancy is that bad. But what if they white player averages 12 points, and 5 assists while the black player averages 18 points and 8 assists. This is now clear discrimination towards the black player. Since his stats are better, that would mean he is considered a better player, but makes less.

Table 4 shows the results of the Fixed Effect Model compared to the initial model. They are similar. The only value that had an economically significant change would be the position variables. An explanation for this could be that certain teams have certain needs at different positions. The control team was the Atlanta Hawks (no real reason I chose them other than they came first alphabetically). So, the Center and Forward variable were both lower, so it could be that they don't have highly paid players at that position because they are a team that focus more on paying their guards rather than forwards or centers.

The big difference in this model was the black variable. In the Fixed Effect Model, this variable becomes insignificant. But this result makes sense. This shows that what Becker said was true. It appears that only non-competitive teams can afford to discriminate. Since the black variable was statistically insignificant, that means there isn't significant evidence that there is a pay difference between white and black players after controlling for player statistics. The OLS model shows as a whole, there is significant evidence that black players make less than white

players after controlling for statistics. Based on the OLS results, I am making the assumption that there are some teams in the NBA that discriminate against players. Based on previous studies, they found that only certain groups of black players are discriminated against. And Becker theories that firms could afford to discriminate if they wish to stay competitive in the long run. So, another assumption I made after these initial results is that there are certain groups of black players that see disclination.

To further investigate this to see if discrimination does exist for some groups of black players, the first thing I did was create two data sets. The first was for players only on their rookie contracts and players no longer on their rookie contracts. I also create two variables that I defined earlier, one to represent black marginal players or low value adding players (Marginal_black) and one to represent the better players in the league or high value adding players (Prime_black). A Fixed Effect Model was ran for the only rookie players and the primary variable being the black variable, one using non-rookies and the Prime_black variable as the main variable, and one that used all of the players and the Marginal_black variable as the main variable. The results can be seen in Table 5.

My assumptions were that with the rookie players there most likely isn't significant evidence for discrimination. But this model wouldn't be as important to the study as the rest since rookie contracts aren't based on the players stats, but their draft position. A rookie contract can last for either 3 or 4 seasons. The terms of this contract are based on what pick in the draft they were as there is a set range in which a team must pay their draft selection. This means that for possibly the first 4 years of a players career their salaries aren't based on performance but rather where they were drafted. For the high value adding black players, there shouldn't be evidence of discrimination. These are the players a team can't afford to discriminate against

because that player would just simply sign somewhere else. If a team were to discriminate against this group of players, they wouldn't ever sign a good player and therefore never be competitive. Finally, for the marginal players, based on the OLS and Fixed Effect models initially ran, I assume that these players will see discrimination. These players aren't as important as the high value adding players and are easily replaced. Therefore, a team can afford to discriminate against them.

My assumptions were correct. For the rookie model, there were 231 observations. The model accounted for around 50% of the variance of the predicted value as the adjusted R^2 was 0.501. The model as a whole was significant with an F-Value of 7.25. The main variable, black, was significant at a .005 significance level. It found black players on rookie contracts make 21.6% less than white players on their rookie contract. As far as this study, these results don't add that much information. This just shows that there is evidence that there could be racial discrimination in the drafting process.

Now on to the two more important models. The model of non-rookie players had 186 observations. The model accounted for around 52% of the predicted values variance as the adjusted R^2 was 0.52. The model as a whole was significant with an F-Value of 6.13. The main variable, Prime_black, was significant at a 0.01 significance level. The parameter estimate was a staggering 0.460. This means that high value adding black players make 46% more than the rest of the NBA. This actually shows there is evidence of reverse discrimination. I assumed this is the group of black players that would not see discrimination. This was true, but it went a little further as this group of players made more than the rest of the NBA after controlling for player statistics. These results make sense. Most of the top players in the NBA are black, so in order to sign these players teams must offer them the most money. It appears in most cases, that these players are

being over paid. This may not be attributed to reverse discrimination though. There's a theory

known as the Superstar Effect, which can be similar to the winner's curse that could explain why

these players are being paid significantly more. The Superstar Effect is the theory that

"superstars" are the individuals that are at the top of their respected field. This is a very small

group of people, but they are highly paid. This could be for a few different reasons. They could

add more value to the team than other plays based on their performance, or they could drive

ticket sales. The more people that attend the games, the more money the team makes. If a single

player can take credit for drawing this crowd, they may be put into this superstar group. With

them being paid so much more, this could show in the results as a significant change in salaries

when in reality it's only a few players making significantly more (Rosen, 1981).

Another effect of the superstar effect is similar to the winner's curse. The winner's curse

takes place in action type settings. It's when whatever is being bid on has a set value, but this

value isn't known by the purchaser, and the purchasers don't know what each other have set their

anticipated values at. For the NBA, this is similar to how free agency is. Not every team will

evaluate the player correctly. Some teams might not use the same analysis to determine player

values and no team could predict player health. Each team has a value for a player, but don't

know what other teams value the that same player at. But since a player doesn't have the same

value for each team, this is not exactly the winner's curse. The winner's curse requires there to

be a uniform value. So, it is considered part of the superstar effect. In order to sign the player to

stay competitive, a team may bid higher than they think the player is worth in order to sign them.

They may "win" the bid and sign the player, but if they did indeed win the bid it's most likely

because they paid too much. This is why you can't conclude that there is reverse discrimination

here because it appears that in this case that the over payment would have to do with teams overpaying the best players to ensure they sign them and stay competitive (Andreff, 2014).

For marginal players, the superstar effect doesn't apply. Teams won't have a bidding war for these types of players because they don't add as much value to a team and are replaceable. If they don't win the bid for this player, they can sign a different one. Meaning there would be no reason to overpay. This is the group that I assumed would see discrimination. This model had 455 observations and accounted for 67.7% of the variation of the predicted value for lnsalary as the adjusted R^2 was 0.677. The model as a whole was significant with an F-Value of 25.42. The main variable was Marginal_black. It was statistically significant at a 0.05 significance level. The parameter estimate was -0.157. So, this group of players makes around 15% less than the rest of the NBA after controlling for player stats and team played for.

VII. Conclusion

In this paper, a regression model was estimated to see if there was evidence for the existence of racial discrimination in the NBA. The model focused on position, stats, and skin color of an NBA player. These variables were used to predict the salary of the player. The results show that black players make 15% less than white players after controlling for player stats. The paired t-test showed there wasn't a difference in average salaries for white and black players, but since black players have all around better stats, they should be making slightly more than white players not slightly less. This model shows that there is evidence for there to be racial discrimination in the NBA. Since a firm will pay the same amount of white and black players, this means that these results show that firms or teams hire more white players than black players overall. But once the team that the player plays for is added to the model, there isn't significant

evidence for there being a differential in players' salaries after controlling for statistics. This follows what Becker theorized. A competitive firm can't afford to discriminate. In the NBA, it appears that a team can't stay competitive and discriminate against the good players.

Next set of results proved this. The high value adding players didn't see discrimination, but actually make significantly more than the rest of the NBA. This could be interpreted as reverse discrimination. This means that a team hires more black than white high value adding players. This makes sense as mentioned before, black players on average have better stats and therefore could be interpreted as better or higher value players than white players. But these results could also be due to what is known as the Superstar Effect.

The model used to examine the marginal players found that there is significant evidence that the black marginal players make 15% less than the rest of the NBA. The results show that not all of the NBA's black players are discriminated against. Based on this study, there is evidence that only non-competitive teams could discriminate and low value adding marginal players see discrimination.

The next steps to further this study would be to determine whether the possible discrimination is employer or customer a model that accounts for customer attendance and views could solve this issue. If data on each stadium and team television views were collected, a model could be used to rule out the discrimination possibly being due to the consumers. A limitation to this study would be that there may have been some variables that are undefinable. Some previous studies have stated that things such as personality and if a player is a good teammate could attribute to their salary.

This research that as has been done in the past has been inconsistent as one person finds evidence of discrimination, and the next doesn't. But this paper has found there to be evidence of discrimination as black players are making 15% than white players after controlling for players statistics. Even when broking down the type of player, the black marginal players also see a 15% pay decrease, but reverse discrimination was found with the high value adding white players. Knowing different theories such as Becker's theory on discrimination or the Superstar effect are important for interpreting these results. Just by looking at the results, it appears that there is some sort of racial discrimination in the NBA, but it also seems as though it could be explained by other things. It appears that fewer black players are signed to contracts, except of the high value adding players which are the opposite. The results show that this is due to racial discrimination, but based on further analysis, could it be due to on average black players being better than white players? This could result in more black players being signed to higher contracts, which could lead to skewed results based on the Superstar Effect. This is why further studies in this area are important.

Over the years it seems that NBA has improved in many social issues. Adam Silver has made progress, but after this study there is still evidence that racial discrimination exists in the NBA. There is still work to be done to reach racial equality. Hopefully this research and further researched like this can be used to not only end racial discrimination in the workplace for the NBA, but for the rest of the world. Then not stop at just the workplace but go further and eliminate as much racial injustice as is possible.

Works Cited

Andreff, Wladimir. "The Winner's Curse in Sports Economics." *Contemporary Research in Sports Economics*, 1 Mar. 2014, doi:10.3726/978-3-653-04103-3/22.

Arrow, Kenneth J. "What Has Economics to Say About Racial Discrimination?" Journal of Economic Perspectives, vol. 12, no. 2, 1998, pp. 91–100., doi:10.1257/jep.12.2.91.

Borjas, George J. *Labor Economics*. McGraw-Hill/Irwin, 2020.

Dey, Matthew S. "Racial Differences in National Basketball Association Players Salaries: A New Look." The American Economist, vol. 41, no. 2, 1997, pp. 84–90., doi:10.1177/056943459704100211.

Groothuis, P. A., & Hill, J. R. (2013). Pay discrimination, exit discrimination or both? Another look at an old issue using NBA data. *Journal Of Sports Economics*, *14*(2), 171-185. ISSN: 1527-0025. doi: 10.1177/1527002511418515

Groothuis, Peter A., and James Richard Hill. "Career Duration in the NBA." Journal of Sports Economics, vol. 19, no. 6, 2017, pp. 873–883., doi:10.1177/1527002517690789.

Hoffer, Adam J., and Ryan Friedel. "Does Salary Discrimination Persist for Foreign Athletes in the NBA?" SSRN Electronic Journal, 2013, doi:10.2139/ssrn.2274725.

Kahn, Lawrence M. "The Economics of Discrimination." The Oxford Handbook of Sports Economics, Apr. 2012, pp. 20–38., doi:10.1093/oxfordhb/9780195387780.013.0002.

Kahn, Lawrence M., and Malav Shah. "Race, Compensation and Contract Length in the NBA:

2001-2002*." Industrial Relations, vol. 44, no. 3, 2005, pp. 444–462.,
doi:10.1111/j.0019-8676.2005.00394.x.

"NBA All-Star 2020: Roster." *NBA.com*, www.nba.com/allstar/2020/roster#/.

Rosen, Sherwin. "The Economics of Superstars." *The American Economic Review*, vol. 71, no.

5, Dec. 1981, pp. 845–858.,
http://home.uchicago.edu/~vlima/courses/econ201/Superstars.pdf.

"The Business of Basketball: Forbes Releases 22nd Annual NBA Team

Valuations." *Forbes.com*, 11 Feb. 2020, The Business of Basketball: Forbes Releases 22nd

Annual NBA Team Valuations.

Zillgitt, Jeff. "Adam Silver Gives Donald Sterling Lifetime Ban from NBA." *USA Today*,

Gannett Satellite Information Network, 30 Apr. 2014,

www.usatoday.com/story/news/usanow/2014/04/29/donald-sterling-fine-penalty-racism-

audio-commissioner-adam-silver-los-angeles-suspension/8460575/.

Appendix

Table 1: Description of Variables		
Variable Type	**Variable**	**Description**
Dependent	lnsalary	Natural Log of the 2019-2020 Salary
Independent	MP	Career Minutes Per Game
Independent	PTS	Career Pointes Per Game
Independent	TRB	Career Total Rebounds Per Game
Independent	AST	Career Assists Per Game
Independent	STL	Career Steals Per Game
Independent	BLK	Career Blocks Per Game
Independent	WS	Win Shares (Advanced Statistic)
Independent	Experience	Number of Seasons in the NBA
Independent	exp2	Number of Seasons in the NBA Squared
Binary	black	Dentotes if the Player is Black
Binary	first_rd	Denotes Players Drafted in the First Round
Binary	second_rd	Denotes Players Drafted in the Second Round
Binary	undrafted	Denotes Players that Went Undrafted
Binary	Guard	Denotes Players that Play the Guard Postion
Binary	Forward	Denotes Players that Play the Forward Postion
Binary	Center	Denotes Players that Play the Center Postion
Binary	Prime_Black	Denotes Players that are Considered High Value Adding to a Team
Binary	Marginal_Black	Denotes Players that are Considered Low Value Adding to a Team

All data was collected from Basketball-Reference.com

Table 2: Descriptive Statistics

Variable	White Players					Black Players				
	N	Mean	Std Dev	Min	Max	N	Mean	Std Dev	Min	Max
MP	91	20.48	7.27	3.00	33.40	364	21.66	8.36	1.30	38.50
PTS	91	8.85	4.35	1.40	24.00	364	9.43	5.52	0.00	27.10
TRB	91	4.08	2.30	0.50	11.20	364	3.84	2.28	0.00	13.90
AST	91	1.75	1.38	0.00	7.70	364	2.01	1.70	0.00	9.50
STL	91	0.56	0.32	0.00	1.90	364	0.72	0.40	0.00	2.20
BLK	91	0.44	0.40	0.00	2.00	364	0.46	0.43	0.00	2.40
WS	91	19.81	24.15	-0.10	144.10	364	19.55	29.36	-2.50	234.00
Experience	91	5.63	4.07	1.00	18.00	364	5.54	4.16	1.00	22.00
exp2	91	48.00	65.29	1.00	324.00	364	47.87	66.63	1.00	484.00
lnsalary	91	15.38	1.01	12.94	17.30	364	15.28	1.18	11.49	17.51
first_rd	91	0.52	0.50	0.00	1.00	364	0.61	0.49	0.00	1.00
second_rd	91	0.26	0.44	0.00	1.00	364	0.22	0.42	0.00	1.00
undrafted	91	0.22	0.42	0.00	1.00	364	0.16	0.37	0.00	1.00
Guard	91	0.35	0.48	0.00	1.00	364	0.50	0.50	0.00	1.00
Forward	91	0.36	0.48	0.00	1.00	364	0.40	0.49	0.00	1.00
Center	91	0.29	0.45	0.00	1.00	364	0.10	0.30	0.00	1.00

Table 3: Pearson Correlation Coefficients, N = 364

	MP	PTS	TRB	AST	STL	BLK	WS	Forward	Center
MP	1	0.8918	0.6273	0.66202	0.78504	0.32643	0.5871	-0.08625	0.05461
		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.1004	0.2987
PTS	0.89175	1	0.5785	0.6815	0.66962	0.29195	0.6172	-0.0697	0.04274
	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	0.1846	0.4162
TRB	0.62729	0.5785	1	0.20292	0.41721	0.75905	0.5202	0.15525	0.50857
	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	0.003	<.0001
AST	0.66202	0.6815	0.2029	1	0.7182	-0.04792	0.477	-0.30672	-0.13665
	<.0001	<.0001	<.0001		<.0001	0.3619	<.0001	<.0001	0.009
STL	0.78504	0.6696	0.4172	0.7182	1	0.18315	0.5333	-0.13649	-0.07527
	<.0001	<.0001	<.0001	<.0001		0.0004	<.0001	0.0091	0.1518
BLK	0.32643	0.292	0.7591	-0.04792	0.18315	1	0.3415	0.12716	0.58104
	<.0001	<.0001	<.0001	0.3619	0.0004		<.0001	0.0152	<.0001
WS	0.58714	0.6172	0.5202	0.47695	0.53328	0.34146	1	0.01077	0.14254
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		0.8377	0.0064
Forward	-0.08625	-0.0697	0.1553	-0.30672	-0.13649	0.12716	0.0108	1	-0.26845
	0.1004	0.1846	0.003	<.0001	0.0091	0.0152	0.8377		<.0001
Center	0.05461	0.0427	0.5086	-0.13665	-0.07527	0.58104	0.1425	-0.26845	1
	0.2987	0.4162	<.0001	0.009	0.1518	<.0001	0.0064	<.0001	

Table 4: OLS and Fixed Effect Models		
Variable	**OLS**	**FE**
Intercept	13.270*** (0.157)	13.233*** (0.227)
second_rd	-0.0894 (0.084)	-0.048 0.083
undrafted	-0.350*** (0.094)	-0.321*** (0.093)
black	-0.1491* (0.084)	-0.127 (0.084)
Forward	0.061 (0.087)	0.014 (0.086)
Center	0.25878* (0.15475)	0.165 (0.154)
PTS	0.02512 (0.016)	0.03172* (0.016)
TRB	-0.0411 (0.030)	-0.042 (0.030)
AST	0.02578 (0.035)	0.022 (0.035)
STL	0.11543 (0.157)	0.120 (0.158)
BLK	0.12442 (0.125)	0.172 (0.123)
MP	0.043*** (0.012)	0.037*** (0.012)
WS	0.010*** (0.003)	0.011*** (0.003)
Experience	0.298*** (0.027)	0.315*** (0.027)
exp2	-0.018*** (0.002)	-0.019*** (0.002)
N	455	455
Adj. R-Squared	0.6585	0.6834
F-Value	63.53	23.79

Standard Errors in parentheses.

*, **, and *** represent significance at the 90 percent, 95 percent, and 99 percent level, respectively.

Table 5: Rookie Contract Players, Non-Rookie Contract Players, and Marginal Players

Variable	Rookies	Non-Rookies	Marginal
Intercept	13.618*** (0.251)	13.188*** (0.571)	13.175*** (0.217)
Prime_Black	- -	0.460*** (0.192)	- -
Marginal_Black	- -	- -	-0.157** (0.078)
black	-0.216** (0.103)	- -	- -
PTS	0.024 (0.024)	-0.013 (0.028)	0.031* (0.016)
TRB	0.021 (0.044)	-0.045 (0.040)	-0.025 (0.026)
AST	0.055 (0.052)	-0.049 (0.048)	0.027 (0.033)
STL	-0.100 (0.234)	0.233 (0.229)	0.060 (0.154)
BLK	0.194 (0.173)	0.187 (0.183)	0.223 (0.117)
MP	0.044*** (0.015)	0.073** (0.025)	0.034*** (0.012)
WS	0.003 (0.013)	0.014*** (0.003)	0.011*** (0.003)
Experience	- -	0.282** (0.092)	0.318*** (0.027)
exp2	- -	-0.019*** (0.004)	-0.019*** (0.002)
N	231	186	455
Adj. R-Squared	0.501	0.520	0.677
F-Value	7.25	6.13	25.42

Standard Errors in parentheses.

*, **, and *** represent significance at the 90 percent, 95 percent, and 99 percent level, respectively.

X. SAS Code

```
libname Sr_Prjct "D:\Senior Project";
DATA Sr_Prjct.NBA ;
  SET Data;

Experience = To - From  + 1;
exp2=Experience**2;
lnsalary = Log(_2019_2020_salary);

IF rd = 1 THEN first_rd = 1;
    ELSE first_rd = 0;
IF rd = 2 THEN second_rd = 1;
    ELSE second_rd = 0;
IF rd = "." THEN undrafted = 1;
    ELSE undrafted = 0;
IF race = "Black" THEN black = 1;
    ELSE black = 0;
IF race = "White" THEN white = 1;
    ELSE white = 0;
IF pos = "G" THEN Guard = 1;
    ELSE Guard = 0;
IF pos = "F" THEN Forward = 1;
    ELSE Forward = 0;
IF pos = "C" THEN Center = 1;
    ELSE Center = 0;
IF Current_PPG < 7 and Black = 1 THEN Black_Marginal = 1;
        ELSE Black_Marginal = 0;
IF Black = 1 and Current_PPG > 18 THEN Black_Prime = 1;
        ELSE Black_Prime = 0;
IF Team = "ATL" THEN ATL = 1;
    ELSE ATL = 0;
IF Team = "BOS" THEN BOS = 1;
    ELSE BOS = 0;
IF Team = "BRK" THEN BRK = 1;
    ELSE BRK = 0;
IF Team = "CHI" THEN CHI = 1;
    ELSE CHI = 0;
IF Team = "CHO" THEN CHO = 1;
    ELSE CHO = 0;
IF Team = "CLE" THEN CLE = 1;
    ELSE CLE = 0;
IF Team = "DAL" THEN DAL = 1;
    ELSE DAL = 0;
IF Team = "DEN" THEN DEN = 1;
    ELSE DEN = 0;
IF Team = "DET" THEN DET = 1;
    ELSE DET = 0;
IF Team = "GSW" THEN GSW = 1;
    ELSE GSW = 0;
IF Team = "HOU" THEN HOU = 1;
    ELSE HOU = 0;
```

```
IF Team = "IND" THEN IND = 1;
    ELSE IND = 0;
IF Team = "LAC" THEN LAC = 1;
    ELSE LAC = 0;
IF Team = "LAL" THEN LAL = 1;
    ELSE LAL = 0;
IF Team = "MEM" THEN MEM = 1;
    ELSE MEM = 0;
IF Team = "MIA" THEN MIA = 1;
    ELSE MIA = 0;
IF Team = "MIL" THEN MIL = 1;
    ELSE MIL = 0;
IF Team = "MIN" THEN MIN = 1;
    ELSE MIN = 0;
IF Team = "NOP" THEN NOP = 1;
    ELSE NOP = 0;
IF Team = "NYK" THEN NYK = 1;
    ELSE NYK = 0;
IF Team = "OKC" THEN OKC = 1;
    ELSE OKC = 0;
IF Team = "ORL" THEN ORL = 1;
    ELSE ORL = 0;
IF Team = "PHI" THEN PHI = 1;
    ELSE PHI = 0;
IF Team = "PHO" THEN PHO = 1;
    ELSE PHO = 0;
IF Team = "POR" THEN POR = 1;
    ELSE POR = 0;
IF Team = "SAC" THEN SAC = 1;
    ELSE SAC = 0;
IF Team = "SAS" THEN SAS = 1;
    ELSE SAS = 0;
IF Team = "TOR" THEN TOR = 1;
    ELSE TOR = 0;
IF Team = "UTA" THEN UTA = 1;
    ELSE UTA = 0;
IF Team = "WAS" THEN WAS = 1;
    ELSE WAS = 0;
LABEL Rd = "Round Drafted"
      Pk = "Pick Selected"
      To = "Most Resent Year Played"
      From = "First Year in NBA"
      G = "Career Games Played"
      MP = "Minutes Per Game"
    PTS = "Pointes Per Game"
    TRB ="Total Rebounds Per Game"
    AST ="Assists Per Game"
    STL = "Steals Per Game"
    BLK = "Blocks Per Game"
    FG_ = "Field Goal %"
    _2P_ = "2 Point Field Goal %"
    _3P_ = "3 Point Field Goal %"
    FT_ = "Free Throw %"
    WS = "Win Shares (Advanced Statistic)";
 RUN;

DATA NR_NBA ;
```

```sas
   SET Sr_Prjct.NBA;
   IF experience > 5;
   RUN;

 DATA R_NBA ;
   SET Sr_Prjct.NBA;
   IF experience < 5;
   RUN;

 DATA Black_NBA ;
   SET Sr_Prjct.NBA;
   IF race = "Black";
   RUN;

PROC MEANS DATA=Sr_Prjct.NBA;
   TITLE "Total Data Means";
   RUN;

PROC MEANS DATA=Sr_Prjct.nba;
   CLASS black;
   TITLE "Black and White Player Split Means";
RUN;

PROC TTEST DATA=Sr_Prjct.NBA;
CLASS black;
VAR _2019_2020_Salary ;
TITLE 'Pair t-test Black vs White';
RUN;

PROC CORR;
VAR MP PTS TRB AST STL BLK WS Forward Center;
TITLE 'Performance Variable Correlation Test';
RUN;

PROC REG DATA=Sr_Prjct.nba;
MODEL lnsalary = second_rd undrafted black Forward Center PTS TRB AST STL BLK
MP WS experience exp2;
TITLE 'All Player OLS';
RUN;

PROC REG DATA=Sr_Prjct.nba;
MODEL lnsalary = second_rd undrafted black Forward Center PTS TRB AST STL BLK
MP WS experience exp2 BOS     BRK    CHI    CHO    CLE    DAL    DEN    DET    GSW
       HOU    IND    LAC    LAL    MEM    MIA    MIL    MIN    NOP    NYK    OKC    ORL
       PHI    PHO    POR    SAC    SAS    TOR    UTA    WAS;
TITLE 'All Player Fixed Effect Model';
RUN;

PROC REG DATA=work.NR_NBA;
MODEL lnsalary = Black_Prime PTS TRB AST STL BLK experience exp2 MP WS BOS BRK
CHI CHO CLE DAL DEN DET GSW HOU IND LAC LAL MEM MIA MIL MIN NOP NYK OKC ORL
PHI PHO POR SAC SAS TOR UTA WAS;
TITLE 'Non-Rookie Regression Fixed Effect Model';
RUN;

PROC REG DATA=work.R_NBA;
```

```
MODEL lnsalary = black PTS TRB AST STL BLK MP WS BOS BRK CHI CHO CLE DAL DEN
DET GSW HOU IND LAC LAL MEM MIA MIL MIN NOP NYK OKC ORL PHI PHO POR SAC SAS
TOR UTA WAS;
TITLE 'Rookie Only Regression Fixed Effect Model';
RUN;

PROC REG DATA=Sr_Prjct.nba;
MODEL lnsalary = Black_Marginal PTS TRB AST STL BLK MP WS experience exp2 BOS
BRK CHI CHO CLE DAL DEN DET GSW HOU IND LAC LAL MEM MIA MIL MIN NOP NYK OKC
ORL PHI PHO POR SAC SAS TOR UTA WAS;
TITLE 'Marginal Player Regression Fixed Effect Model';
RUN;
```