CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Bachelor thesis

2019 **Dmitrii Ivantsev**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Dmitrii Ivantsev

Wage Discrimination in US Sports: Comparative Analysis

Bachelor thesis

Prague 2019

Author: Dmitrii Ivantsev

Supervisor: PhDr. Radek Janhuba, M.A., Ph.D.

Academic Year: 2018/2019

Bibliographic note

IVANTSEV, Dmitrii. *Wage Discrimination in US Sports: Comparative Analysis.* Bachelor thesis. Charles University, Faculty of Social Sciences, Institut of Economic Studies.
Supervisor PhDr. Radek Janhuba, M.A., Ph.D.

Abstract

The issue of salary discrimination in the US sport has been a controversional topic, largely dscussed in the context of racial discrimination in American society. Numerous papers written in this field in the previous years investigated each of the US Major Leagues separately. This thesis provides a comparative analysis of wage discrimination in the Nationall Basketball Association, Major League Baseball and National Football League. The data on players' contracts for all leagues were collected based on 2015-2018 unrestricted free agents' signings with a different sample of players for each specific year. We created five unique datasets for the comparative analysis, including one NBA, two NFL and two MLB related samples, where the MLB datasets represent batting and pitching players, while the NFL samples refer to the players on the offensive and defensive positions. For the estimation of players' productivity, we considered the average performance statistics in three years prior to the first contract year. By adding dummy variables reflecting players' race into the regressions and applying the Ordinary Least Squares method we could compare the size of the wage premium paid to the representatives of the specific race. The results of our regressions do not provide sufficient evidence of salary discrimination in any of the considered leagues.

Abstrakt

Otázka platové diskriminace v americkém sportu stále zůstává kontroverzním tématem široce diskutovaném v kontextu rasové disckriminace v americké společnosti. Četné vědecké práce, zabývající se problémem platové diskriminace, zkoumaly každou z hlavních amerických lig zvlášť. Tato studie je věnována komparativní analýze platové diskriminace v Major League Baseball, National Basketball Association a National Football League. V práci jsou použité kontraktní údaje neomezených volných agentů z let 2015-2018, kde každý rok je reprezentován jiným vzorkem hráčů. Za účelem komparativní analýzy bylo vytvořeno pět různých datových sad: jedna pro NBA a dvě pro každou ze zbývajících lig, kde MLB vzorky jsou reprezentovany pálkaři a nadhazovači, kdezto NFL vzorky jsou tvořeny ofenzivními a defenzivními hráči. Za účelem analýzy produktivity hráčů jsou použity průměrné herní statistiky ze tří let před podepsáním konkrétní smlouvy. Přidání kategorické proměnné odpovídající rase hráče a použití metody nejmenších čtverců nám umožňuje porovnat velikost prémia, které dostávají představitelé určité etnické příslušnosti. Výsledky našich regresí neposkytují postačující evidenci existence platové diskriminace v žádné ze zkoumaných lig.

Keywords

Discrimination, basketball, football, baseball, NBA, NFL, MLB, wage, racism.

JEL Classification: J15, J30, J31, J71, Z21, Z22

Klíčová slova

Diskriminace, basketbal, fotbal, baseball, NBA, NFL, MLB, plat, rasismus.

Klasifikace JEL: J15, J30, J31, J71, Z21, Z22

Declaration of Authorship

1. The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

2. The author hereby declares that all the sources and literature used have been properly cited.

3. The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague 6.05.2019

Acknowledgments

The author is grateful especially to the thesis supervisor PhDr. Radek Janhuba, M.A., Ph.D. for his help with the choice of topic, time, and useful advices.

Bachelor Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author's name and surname: Dmitrii Ivantsev

E-mail: idmitriy221@gmail.com

Phone: +420774812471

Supervisor's name: PhDr. Radek Janhuba M.A.

Supervisor's email: janhuba@gmail.com

Thesis language: English

Wage Discrimination in US Sports: Comparative Analysis

Research question and motivation

Is there a place for wage discrimination as an example of irrational economic behavior in some of the most financially successful leagues in the world? And is there a difference in wage discrimination between leagues with African American majority (NFL, NBA) and MLB in which most teams are represented by non-African American players? (The Institute for Diversity & Ethics in Sport, 2016). The issue of salary discrimination in American sport industry is considered to be immensely controversial topic as numerous studies conducted on the US top sports leagues provided us with fairly mixed results varying with respect to specific dataset and selected approach. The 25% premia for Hispanic and white players was revealed in bottom quintile of MLB players (Holmes, 2011). Gordon (2015) claims that white players receive 13% higher base salary than their African American colleagues, but at the same time face a larger wage penalty when providing sub-average results. However, Hill et al (2014) based on NFL data concludes that players' career earnings are determined solely by their performance and not the race. As the evidence is mixed, a proper analysis of the latest data could help us understand the phenomenon of salary discrimination in the US sport industry.

Contribution

Recent research on wage discrimination in American professional sports leagues analyzed each league specifically. Applying the same methodology over multiple leagues at the same time we will perform a comparative analysis and consolidate the previous studies on the topic. We will investigate the tendency of wage discrimination across several major US sports leagues and try to reveal the pattern in the employers' behavior.

Methodology

Our analysis will rely on the open data about MLB, NFL and NBA players' wages and crucial performance statistics from the last seasons. For the purposes of our study we will create specific datasets relating to each league. In a theoretical part of our research we will review the existing literature on the

topic in order to compare the previous results with the inference derived from our model. The empirical part of the research will be provided using econometric analysis applied on our datasets.

Outline

1. Introduction
2. Literature Review
3. Theoretical Background
4. Data
5. Methodology and Model
6. Results
7. Conclusion

Bibliography

1. Fryer, R. (2010). Racial Inequality in the 21st Century: The Declining Significance of Discrimination. Handbook of Labor Economics, 4.

2. Ducking, J., Groothuis, P., & Hill, J. (2014). Compensation discrimination in the NFL: An analysis of career earnings. Applied Economics Letters (Vol. 21).

3. 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

4. Holmes, P. (2011). New evidence of salary discrimination in major league baseball. Labour Economics, 18(3), 320–331.

5. Gordon, S. (2015). An Examination of Wage Discrimination in the NBA
Available at: http://thesportsquotient.com/nba/2015/5/19/an-examination-of-wage-discrimination-in-the-nba

6. The Institute for Diversity & Ethics in Sport, (2016).
Available at: http://www.tidesport.org/reports.html

7. Wooldridge, Jeffrey M., (2009): "Introductory econometrics: a modern approach." 4th ed. Mason: Thomson/South- Western

List of Acronyms

AFC American Football Conference

AL American League

BRI Basketball Related Income

CBA Collective Bargaining Agreement

MLB Major League Baseball

NBA National Basketball Association

NFC National Football Conference

NFL National Football League

NL National League

OLS Ordinary Least Squares

USA United States of America

VIF Variance Inflation Factor

Contens

Chapter 1

Introduction

Major American leagues are considered some of the most attractive competitions in the world industry of professional sport. Record TV audience and game attendance make the games some of the most favourite events for US population. Aside from being highly competitive, the leagues demonstrate high level of financial stability.

The notion of profit maximization is crucial in the context of successful financial operating, which is not compatible with any manifestation of non-rational behavior, such as discrimination. In the context of this theory, US professional teams can be considered profit maximizing organizations (Zimbalist, 2003; Palmer, 2006). According to Becker (1957), prejudiced behavior, such as discrimination, cannot persist at the competitive labor market for a long time. Since the discriminating team can only hurt its own performance in the highly competitive sphere of professional sport, there is a strong incentive to provide the players with fair compensations (Irwin, 2004).

The segregation on the grounds of race as well as any act of discrimination from employers or labor unions became outlaw with the introduction of the Civil Rights Act (1964) and creation of the Equal Employment Opportunity Commission (1965), however, wage discrimination is often difficult to investigate in the context of labor market due to the lack of available information about the employee's efficiency. The datasets on proxy variables used to estimate employee's performance, such as years of experience or level of education, may be unavailable or wrongly interpreted due to the measurement bias.

The industry of professional sport provides us with all the necessary data for the analysis of different forms of discrimination. Team payrolls and detailed players' performance statistics as well as general information about each player are widely available on the internet. Additionally, the fact that the share of minorities in professional team sports is larger than in the labor force on the whole demonstrates that sport can provide exceptional opportunities to the representatives of minority group (Kahn, 1991).

Another attractive feature of professional sport industry is its "symbolic importance" for many Americans (Palmer, 2006). Millions of American children grow up with a dream to build a career of a professional player. NFL Super Bowl, MLB World Series and NBA All-Star weekend are some of the most watched television programs in

the USA in almost every given year (Statista, 2018). For many years the game of baseball remains "the national pastime" and along with the the other major leagues is expected to represent the nation's ideals (Palmer, 2006).

In this thesis, we will address the question of existence of wage discrimination in baseball, basketball and football US leagues. The National Hockey League is not included in the analysis. Since most of the NHL players are white Americans, Canadians and Europeans with a very low share of Afro-American athletes, the NHL would represent the case of wage discrimination against the foreigners as compared to the Canadian and American players, while this thesis is mainly dedicated to the racial discrimination.

As the numerous studies on the topic of salary discrepancies between white and non-white athletes available for each single league demonstrate fairly mixed results, the main goal for us will be to analyze the existing wage gap in each league and compare the premiums paid to the players depending on their race. Since the Afro-American athletes stand for more than 70% of the NBA and the NFL racial breakdown, while the share of Afro-American athletes is less than 10% in the Major League of Baseball (ESPN.com), comparative analysis of the revealed wage premiums in the mentioned leagues will help us in detecting the pattern in the employers' behavior.

The structure of the thesis is following. In the second chapter, we will review the existing literature discussing the problem of wage discrimination in each league specifically and the problem of discrimination in general. In the third chapter, we will provide a short introduction into the history of the Major American leagues and briefly discuss the fundamentals of each game. Chapter 4 provides the data presentation. We will introduce the criteria used to collect and process the corresponding datasets. Chapter 5 is dedicated to our models and methodology and provides the explanation of the variables selected for our regressions. Results of the regressions will be analyzed in Chapter 6. Chapter 7 concludes.

Chapter 2

Literature Review

This chapter provides a general discussion on the topic of discrimination and reviews the related empirical studies on the Major League Baseball, National Basketball Association and National Football League datasets conducted previously.

2.1 Theories of Discrimination

2.1.1 Taste-Based Discrimination

The structured analysis of wage discrimination in the context of the labor market started with Becker's "The Economics of Discrimination" (1957). Becker's analysis focused on the correlation between prejudice among white people and discriminating behavior against minorities. Becker introduced a formal definition of racial preferences as a distaste to cross-racial interaction. Becker refers to three types of taste-based discrimination: customer, employer and employee with respect to the corresponding source of discriminating behavior.

In the reality of professional sport, customer discrimination is reflected in the fans' attitude towards the players of different race in their team. The research conducted by Bodvarsson and Pettman (2002) for MLB and Brown et al (1991) for NBA found that the racial composition of the metropolitan area, where the team is based, is correlated with the salary discrimination against minorities. Burdekhin and Idson (1988) used two different datasets to detect the correlation between the racial composition of the team and the area and found that the game attendance is positively correlated with the number of players of the same race in the team.

Co-worker or employee discrimination appears in sports when players of a majority race refuse to work with their teammates belonging to the other group. The most famous example of co-worker discrimination is the story of Jackie Robinson described by Jules Tygiel (1997) in the book "Baseball's Great Experiment: Jackie Robinson and His Legacy". After signing a contract with Brooklyn Dodgers in 1947 he became the first

Afro-American baseball player to join the league and faced the cases of discriminating behavior from his teammates' side.

Finally, employer discrimination is a phenomenon that appears when the employee is treated differently from his colleagues based on race, religion, gender or national origin. This type of discrimination will be largely discussed in the studies present in the second part of this chapter.

2.1.2 Statistical Discrimination

Statistical discrimination results from the deficit of available information about people's performance. As a result, the specific employee is treated based on the information about the group that he belongs to. In this case, a rational decision-maker uses the aggregate characteristics of the group to evaluate the unobservable characteristics of the individual, consequently, inequality may exist between two groups even if the economic agents are rational and non-prejudiced (Moffatt, 2018).

While Becker's research was mainly dedicated to the investigation of the taste-based discrimination models, Phelps (1972) and Arrow (1973) introduced the statistical discrimination models in the papers "The Statistical Theory of Racism and Sexism" and "The Theory of Discrimination" respectively. Unlike Becker's theory, which explained discrimination as a consequence of the racial taste factor, the theory of statistical discrimination refers to the employers and workers who have no distaste to hire or work with the members of the other racial group.

Players' productivity today can be perfectly monitored thanks to the technological development. Advanced software used in professional sport provides team management with complete analysis of every possible individual characteristic. For this reason, we consider the case of taste-based, but not the statistical discrimination in this thesis.

2.2 Related Emprirical Studies

In order to summarize the existing state of knowledge on wage discrimination in the sphere of professional US sport we will go through the available empirical results for each league specifically.

2.2.1 National Basketball Association

The number of papers dedicated to the racial discrimination in the National Basketball Association is fairly extensive. The first research conducted in this field relates to the 1970s. The earlier articles of Rockwood and Asher (1976) and Mogull (1977,1981) found no evidence of discriminative behavior against non-white athletes, since they were paid as much as their white colleagues demonstrating the same performance. Rockwood and Asher reviewed the earlier Mogull's "Wage Discrimination in Professional Basketball" (1974) and re-ran the model to test his conclusions. Considering the sample of 28 players with equal number of white and Afro-American players from the seasons of 1970-71, Mogull found that an Afro-American player has to perform better in order to be paid equally as his white teammate. Rockwood and Asher, however, concluded that players are rewarded solely in accordance with their performance, but not their race. Scott, Long, and Somppi (1985) found that both white and Afro-American players provided similar performance and were paid equally, however, the investigated sample consisted of 26 players only and was not large enough to reflect the state of affairs in the league on the whole.

Later papers on the topic of wage discrimination worked with bigger datasets and could provide more accurate results, since they could capture the overall trends rather than characteristic of specific samples (Koch and Vander Hill, 1988; Kahn, 1991). These studies succeeded to provide the evidence of wage discrimination based on the salary differences between white and non-white players of equal performance. Wallace (1988) used a dataset of 229 players form the 1984-85 season and ran an OLS regression with various performance variables and race dummy variable included in the model and found 16.8 percent salary gap between the players from white and Afro-American groups. Koch and Vander Hill (1988), Kahn and Sherer (1988), and Brown, Spiro, and Keenan (1991) applied the same econometric methods on the datasets of approximately similar size, collected from the seasons 1984 – 1986, and revealed the salary gap between white and African-American players of 11 percent, 20 percent, and 14 percent respectively.

As these studies considered a set of wages and statistic performance from a single season, the long-term contracts signed before the measured season could not accurately reflect current player's value, which led to the systematic measurement errors. Jenkins (1996) introduced a new approach to the data measurement. In the "Reexamination of Salary Discrimination in Professional Basketball", the author collected the data on the

newly signed contracts from 1983 to 1994 and excluded the players under their first contract, since they had no previous NBA experience. The regression and a Chow test proved the insignificance of the race variable and showed the same return on increase in performance for both races.

Other literature of 1990s provided fairly mixed results concerning racial discrimination in the NBA. Dey (1997), Gius and Johnson (1998) and Bodvarsson and Brastow (1999) failed to find any evidence of salary discrimination in the end of 1980s and beginning of 1990s. At the same time Hamilton (1997) showed the differences in the upper part of wage distribution from the 1994-1995 season. Dey (1997) reapplied the model implemented by Brown, Spiro, and Keenan (1991) on each of the seasons between 1987-1993 and found salary discrepancies between Afro-American and white athletes in two of the mentioned seasons as well as in the pooled sample which, however, disappeared after adding the dummy variable for centers. Hence, the initial wage premium indicated by Brown et al (1991) could be in reality a premium for centers, since this position was mainly represented by white players. This conclusion was later confirmed by Hill (2004). The author found only a slight evidence of salary discrimination in a sample from 1990-2000 period and referred to the correlation between race and height. Wage premium for centers is also revealed in the more recent paper by Pacák (2016).

More recent studies on the topic of racial salary discrimination were generally not able to support the evidence of unequal compensation to different race groups. Groothuis and Hill (2011) failed to find the evidence of wage discrimination considering 1990-2008 NBA data and applying a modified Heckman procedure while controlling for survival bias. Pacák (2016) used both all-career performance as well as players' statistics in the last three seasons before the contract signing and considered the sample of 230 NBA players with the current contracts. He applied OLS regression with a race dummy variable and found no evidence of salary discrimination, but detected a 50% premium given to the centers.

Naito and Takagi (2016) created an unbalanced panel dataset including data on players' salaries from the 1985-1986 season up to the 2015-2016 season. Contrary to the previous studies, the authors found the emerging salary gap in the beginning of 2000s increasing up to 20% by 2010. Johnson and Minuci (2018) used Oaxaca-Blinder decomposition and weighted linear regression models with the sample of 800 free agents from the seasons 2011-2017 and concluded that 72.7% of the wage gap between white and

Afro-American athletes was explained by race. Non-white athletes were found to receive 20,5% less than their white colleagues keeping other factors equal.

2.2.2 Major League Baseball

Salary discrimination in the Major League Baseball emerged as an object of active discussion in the 1970s. Rapping and Pascal (1971) found no evidence of discriminating behavior in the baseball labor market, but revealed the racial difference in performance. In the other paper dedicated to the MLB, G.Scully (1974) found the evidence of salary discrimination against Afro-American non-pitchers in the season 1970, but no evidence concerning non-white pitchers. He concluded that white hitters received higher wages than their non-white colleagues demonstrating similar performance in the season.

However, later studies on the topic of wage discrimination in the baseball labor market conducted at 1970-1985 period found only slight or zero discrimination against Afro- American players. According to Medoff (1975), the results provided by both Scully and Rapping were not accurate as the theory of marginal productivity was not considered in their papers, which also missed the significant variables in the corresponding models.

The issue of racial discrimination was not actively discussed in the scientific literature since the middle of the 1980s as the problem seemed to disappear. Kahn (1991,2000) provided a review of the existing literature and stated in his paper of 2000 that "…regression analyses of salaries in baseball and football have not found much evidence of racial salary discrimination against minorities."

In the article by Bodvarsson and Pettman (2002) "Racial Wage Discrimination in Major League Baseball: Do Free Agency and League Size Matter?", the authors used a logarithmic-linear model to detect the salary discrimination in the labor market for pitchers. Including various performance statistics and other related variables, such as population of the team metropolitan area and racial composition of the team metropolitan area, in their model, they came to the conclusion that league expansion in 1993 actually helped to eliminate the salary discrimination as with increased competition in the league the owners tried to provide fair salaries to get the best players in their team. Will Irwin (2004) in his investigation of wage discrimination among outfielders detected that variable indicating the race of the specific player is not significant.

Traditional quantitative methods used in the models consider the effect of race for an average player. Even though the discrimination might not be evident on a large scale and might not appear in the middle and high salary groups, non-white players with below-average performance might still be discriminated. Palmer (2006) in his paper "Has Salary Discrimination Really Disappeared from Major League Baseball?" states that minorities in the lower salary group are not compensated for the improved performance as much as white players. His conclusion is that white fans might still prefer to see the players of the same race in their team, but as long as it does not interfere with the team results. Palmer's results are supported by Holmes (2011). Considering the contracts signed between 1998 - 2006 and applying the quantile regression on the corresponding data he found up to 25% premium for Hispanic and white players in the lower quantile.

In the context of recent research, no sufficient evidence of wage discrimination has been found. Arias's (2016) analysis of 100 players from the season of 2015 does not detect any wage differences caused by the race variable. However, the number of stolen bases was found to be undervalued by the team owners and could result in the compensation gap for non-white players, since they performed more steals per season.

Masiella (2017) used data from 2012-2015 and applied simple OLS along with a two-stage least squares regression, but failed to reveal any sign of employers' discriminative behavior towards minorities.

2.2.3 National Football League

The problem of wage discrimination in American Football received relatively less attention compared to the National Basketball Association and the Major League Baseball. The first research on the NFL data was provided by Mogull (1973,1981) and Kahn (1992) with the use of 1970 and 1989 salaries and performance statistics accordingly. In the first of his two studies, Mogull used the data of 96 NFL athletes and applied the descriptive statistics to the sample. In the later research, the sample of 64 players was used to perform a Chow test. Whereas datasets used by Mogull were relatively small, Kahn applied the regression model on a sample of 1363 players, which stood for about 94% of all the players in the league. No empirical evidence of essential wage premium for whites was found in either of Mogull's research, however Kahn found that player's salary depends on the racial composition of the metropolitan area where the

team is located and found the maximum of 4% wage premium for whites. According to Kahn's results, white and non-white players were more popular among fans of the same race and were better paid by the team owners. In his research, Kahn also introduced three possible reasons for wage discrimination in American football based on the Nash's theory. He indicates owners' prejudice, customers' preferences and lower social and financial status of non-whites outside of the professional sport field as the factors affecting the wage gap between whites and non-whites.

In the later research on the wage discrimination, Guis and Johnson (2000) obtained the empirical evidence of the reverse discrimination. Applying linear regression and a Chow test on a dataset of 938 players from the season of 1996, they concluded that white players earned 10% less than their Afro-American colleagues. The authors introduced no performance explanatory variables in their model and argued that no performance statistics can be applied to the whole NFL dataset due to the different players' positions.

Even though the position of quarterback is typically occupied by white players, Afro-American quarterbacks are not uncommon in today's NFL. Berri & Simmons (2007) investigated the case of wage discrimination among quarterbacks. They revealed the style differences in the play manner of white and Afro-American quarterbacks. Using the quantile regression, they found no compensation for additional rush yards achieved. Since non-white players tend to have more rush yards than their white colleagues, they receive no additional reward for extra passing in some range of salary distribution.

More recent studies provided us with fairly mixed results. Keefer (2013) revealed the wage premium for white linebackers by using a quantile regression with a dummy variable technique. Following the same approach Ducking et al (2014) found the reverse discrimination in the defensive lineman market and failed to detect the labor market discrimination in the defensive market as a whole. At the same time, Burnett and Van Scyoc (2015) followed the same approach, however, they applied it on the rookies rather than all current players and limited for the tight ends and found no evidence of racial discrimination.

Chapter 3

Theoretical Background

This chapter provides the general information about each league along with the discussion on the labor agreements between the corresponding league and the players' labor unions. Collective Bargaining Agreement is an agreement between the players' association and the league, which determines the rules of the employment and the financial structure of the game. Since the details of this agreement play an important role in determining player's compensation, it should be discussed before we proceed to the introduction of our models.

Several studies conducted on the topic of wage discrimination investigated the impact of the CBA negotiations on the players' earnings. Bodvarsson and Brastow (1998,1999) came to the conclusion that the new NBA Collective Bargaining Agreement signed in between 1985-86 and 1990-91 seasons diminished the monopsony power of teams over their players and led to disappearance of the employers' discrimination in the end of 1980s, while Larsen et al (2006) using the NFL data from 1970-2002 seasons found that the introduction of salary cap and free agency led to the increased competitiveness among the NFL teams.

Free agency and salary cap are two crucial CBA rules for our research. The status of a free agent gives his owner a possibility to negotiate with other teams about his future career and select the best option, while salary cap directly determines the limit for the team spending on players' contracts. The salary cap is supposed to balance the teams in terms of contract spending and stop the wealthier teams from outspending the poorer ones in the negotiations with free agents.

3.1 National Basketball Association

The game of basketball was invented by Canadian professor of physical education James Naismith. For many years the game process and rules have been vastly modified. Basketball is a team sport involving 5 players on the court from each side competing against each other. Today's NBA game consists of four 12-minutes quarters and additional overtimes if none of the teams defeats the opponent in the previous 48 minutes. The goal

of the game is to outscore the opposing team by shooting more points in the opponents' ring, while defending its own ring.

The National Basketball Association was founded in 1946 as the Basketball Association of America and included 11 teams divided into two divisions. The league received its current name after the merger with the National Basketball league in 1949. Today the most popular basketball league in the world consists of 29 American and one Canadian team competing in the Western and Eastern conferences. The regular season consists of 82 games which are followed by play-off with eight best teams from each conference and continues until one team from each conference proceeds to the final series, in which the winner is determined.

3.1.1 Free Agency and Salary Cap

The first attempts to organize players' labor union were taken by Bob Cousy in 1954 as many players competing in the NBA were not satisfied with the financial conditions defined by the league. Later, in 1957 the NBA Board of Governors recognized the modifications of several financial aspects including retired players' pensions, payment of back salaries and other payments. The first CBA agreement was signed in 1970 for three years providing an increase in minimum salaries, playoff pool and per diem allowance.

The 1976 agreement introduced the limited free agency which gave the players more bargaining power in the contract negotiations as before the free agency rule players were prohibited to negotiate with the other teams about their future contracts and were obliged to stay in their current team after the termination of their contract unless the team decided to trade them to another team. These modifications let the players negotiate the contract conditions with several teams interested in them and choose the best option.

Even though the salary cap existed in the league in its first season after the merger in 1946- 1947, the modern salary cap was introduced in 1984-1985 season at $3.6 million. The latest CBA signed in 2011 for the period of ten years set the salary cap at 51.2% of the basketball related income with a range of 49-51% in the following years of the agreement. The salary cap introduced by this CBA represents an example of a soft cap system. It means that teams can spend on salaries above the introduced salary cap by using different exceptions without paying a penalty up to the certain point defined by the

threshold of the luxury tax. For the latest season considered in this thesis the salary cap was set at $99 million, while the threshold for the luxury tax was set at $119 million.

3.2 Major League Baseball

Official history of baseball starts from the 40s of the 19th century. The rules of the game, developed in 1845 by Alexander Cartwright, one of the founders of the New York Knickerbockers baseball team, formed a base of the modern baseball rules.

Contrary to the other games considered in this thesis, a game of baseball has no time restrictions. A single baseball match consists of nine innings. In each of the nine innings, the opposing teams play in the offense and defense, taking turn batting and fielding. The runs are gained in the offensive part of the game and a team with a higher number of runs wins the game. If both teams have the same number of runs after nine innings the extra innings are played until the winner is known.

Baseball is considered to have been the first professional sport in the US with the National Association of Baseball Players and the National League founded in 1850 and 1876 respectively. The Major League Baseball, the oldest US major league, was founded in 1903 after the merger of the National League and the American Leagues. Modern MLB consists of 29 American and one Canadian team with 15 teams playing in each of the leagues. The regular season of 162 games is followed by the postseason playoffs with 5 best teams from each league. The winner is determined in World Series, the final series between the NL and AL champions.

3.2.1 Free Agency and Salary Cap

The first MLB collective bargaining agreement between the Major League Baseball Players Association and the Major League Baseball signed in 1968 was the first CBA in professional sports history and the first agreement guaranteeing a minimum salary. The agreement served as an important step in limiting the power of team owners, requiring most of the changes in the league governing to be negotiated in the collective bargaining between players and owners.

The CBA signed in 1976 played a crucial role in the regulation of free agent's status in the league. The agreement prescribed a player to become a free agent after six

years of service in the Major League Baseball in case the current contract is not renewed by the player. It also introduced a compensation for the club losing a free agent in the form of the draft choice and limited the clubs in the number of free agents they could sign based on the size of the free agent pool.

The Major League Baseball is the only American major league with no fixed salary cap introduced. However, the clubs are restricted in terms of contract spending by the luxury or competitive balance tax which is paid in case the team exceeds the predetermined threshold reviewed annually. The current CBA ratified in the end of 2016 for 5 years introduced the luxury tax threshold at the level of $206 million in the season of 2019 with further increase in the seasons 2020 and 2021.

3.3 National Football League

The game of American football was introduced in the third quarter of the 19th century as a combination of traditional soccer and rugby. The first official game took place in 1869 between two college teams representing Princeton and Rutgers universities.

American football match lasts for 60 minutes divided into 15-minute quarters and followed by the overtime in case of a tie score in the regular time. It includes two teams with 11 players on the field from each side. The purpose of the game is to move the ball to the opponent's end zone. The points are achieved by crossing the opposition's goal line with the ball, kicking the ball through the post or uprights or tackling the opponent with the ball in their end zone.

The National Football League was founded in 1920 as the American Professional Association with 11 teams competing under the name of the league in its first season and received its current name in 1922. After the merger with the American Football league in 1966 the newly created NFL expanded to 24 teams.

Nowadays NFL has the highest attendance among the professional sport leagues (Statista, 2018), while american football is considered the most popular game in the US (GALLUP, 2018). The league consists of 32 teams equally divided between the American and National Football conferences with the regular season of 16 games scheduled from the beginning of September until the end of December. The best six teams from each conference by the results of a regular season advance to the playoffs. The champion is

determined in the final playoff game between the NFC and AFC champions called Super Bowl.

3.3.1 Free Agency and Salary Cap

The National Football League Players Association was recognized by the league in 1968 as a response to the players' decision to go on a strike. After a brief stoppage, a collective bargaining agreement was signed in the same year, setting the regulations on the minimum salary and contributions to the players' pension.

The institute of the unrestricted free agency did not exist in the National Football League until the introduction of the new collective agreement in 1993. Previously, players' movement was restricted by numerous regulations which did not let the players freely change the current team after the termination of the contract. According to the newly signed CBA, any NFL athlete with at least four years of NFL experience could become an unrestricted free agent.

Additionally, the CBA of 1993 set a salary cap as a response to the introduction of the unrestricted free agency. For its initial season the salary cap was set at the level of $34.6 million. Unlike the NBA salary system, the NFL salary cap is hard, which means that teams are not allowed to exceed it. The latest CBA was signed in 2011 with the expiration after the season of 2020 and introduced an initial salary cap of $120 million in the first season after the CBA. The official salary cap for the season of 2019 is set at $188.2.

3.4 Comparison of Salary Cap and Free Agency

In the previous part of the chapter, we described the regulations concerning salary cap and free agency separately with respect to each league. This section provides a comparison of the existing financial rules in the major leagues in order to understand how these institutes affect players' earnings in the mentioned competitions.

Since we consider players' contracts signed in the period of 2015-2018, the size of the salary cap set by the corresponding league should have a direct impact on the players' earnings. In order to estimate a possible effect of this regulation we proceed to the specific salary caps set by each league in the covered seasons. Table 3.1 demonstrates the

size of the salary cap in the major leagues set before the start of the corresponding season. Since there is no official salary cap in the Major League Baseball, the threshold for the luxury tax is mentioned in the table.

Table 3.1 Salary Caps by Season					
Season	**NBA**	**Season**	**MLB***	**Season**	**NFL**
2015-2016	$70,000,000	2015	$189,000,000	2015	$143,280,000
2016-2017	$94,143,000	2016	$189,000,000	2016	$155,270,000
2017-2018	$99,093,000	2017	$195,000,000	2017	$168,000,000
2018-2019	$101,869,000	2018	$197,000,000	2018	$177,200,000

*threshold for the luxury tax

As we can see, the salary cap in the NBA and NFL was consequently increased during four seasons, while the threshold for the competitive tax in MLB was kept the same in the first two of the considered seasons and increased in the last two. Due to the permanent increase of these limits in each of the leagues, dummy variables indicating the year of the contract signature are expected to have a negative sign as compared to the base year 2018.

The magnitudes of the thresholds are sufficiently larger for the MLB and NFL, as the roster for the season can contain up to 40 and 53 men respectively, while an NBA team can have only 15 players on its active roaster during the season. For this reason, single NBA contracts are expected to be larger on average, while the NFL contracts should possess the lowest salary. Additionally, the existence of the formal salary cap in basketball and football might involve higher risks for players to be discriminated in the corresponding leagues, since the teams are limited in terms of contract spending and might prefer to offer higher wages to the players of the preferred ethnicity when considering the

athletes with similar productivity, while the owners of the MLB teams are less bounded in terms of the budget allocation.

Since there are certain differences in functioning of the institutes of free agency in the three leagues, the limitations for our datasets are set specifically based on the features of free agency in each league. As we consider only unrestricted free agents' signings in this thesis, the rules of the unrestricted free agency are to be discussed in this section. Contrary to the MLB and NFL, where a player can become an unrestricted free agent after 6 and 4 years of the league service respectively, in the National Basketball Association most of the agents are unrestricted with the exception of rookies after the expiration of their frist contract signed next to the draft procedure and players on the two-way contracts. These differences in the free agency system could possibly lead to the difference in the experience of players included in the estimated samples. Since NBA rookie contracts are generally signed for the period from 2 to 4 years, where first two years are guaranteed, while the last two years can be exercised by the team decision, the NBA sample could appear to be younger and less experienced compared to the NFL and MLB samples.

Chapter 4

Data

This chapter provides the overview of the approach selected for the creation of the datasets for each league.

4.1 General Data characteristics

Three types of data are used for each league in our analysis. First, we need the data on the average players' earnings. We will consider the average annual contract value including a guaranteed part calculated based on the unrestricted free agents' signings available on the www.spotrac.com for the period of 2015-2018.

The selected range of four seasons represents comparatively recent contracts, while providing a bigger number of observations for our analysis compared to a one-year set of observations considered in some of the papers mentioned in the previous parts of the thesis.

The second type of data is performance statistics. They are obtained based on the game specifics for each league. We use the data on the crucial players' performance indicators available online.

Opposite to the numerous papers on the topic of wage discrimination written before and considering a single year salary with respect to the current athlete's performance, we rely on the performance statistics collected strictly for three seasons preceding the first year of the current contract. This approach helps us to estimate the performance which has a direct impact on players' earnings. In case of multi-year contracts, performance demonstrated in the previous or current season cannot have any influence on the already signed contract (Masiella, 2017). All-career statistics are not considered, since player's efficiency can vary vastly during the professional career and can be misleading in the context of our analysis. For similar reason we do not select a larger timeframe for our research, since teams focus mainly on the performance demonstrated in the latest seasons before a possible signing.

The time frame of three seasons is selected for several reasons.

Firstly, the data collected based on this time period can be considered a more reliable indicator of players' performance rather than simple one-year estimation, since it indicates an ability to demonstrate a steady level of game productivity for several years.

Secondly, this approach helps us eliminate the players inclinable to injuries from our dataset, since they played too few games to provide enough input for analysis or even missed some of the seasons.

Finally, considering the performance demonstrated in the major leagues we exclude players with no previous experience and players for whom the current contract is the first professional agreement due to the lack of reliable information. Performance demonstrated in the league other than a US major league, does not provide sufficient input for our analysis due to the differences in the level of competition and game rules.

Apart from the lack of professional experience, rookies' contracts are excluded from the sample since their value is restricted by the league regulations and freshmen demonstrating an outstanding performance in their first years in the league are not able to apply for a higher than prescribed salary.

Statistics collected for each league relate exclusively to the performance demonstrated in the regular season. Adding the data from play-off to our datasets would give an advantage to a certain group of players in our sample, since the teams participated in a different number of play-off games in the considered periods. Moreover, players' individual results in play-offs can differ from the regular season, as the teams might prefer to follow less risky strategies.

Finally, the third type of data includes players' personal information as well as other factors which might have an impact on their earnings. The data concerning the athletes' race are obtained from the www.espn.com based on the player's profile pictures and name characteristics.

4.2 Data Specification

Variables related to players' experience, length of the current contract and the year of contract signature will be present in all the models appearing in this thesis based on the logic described below.

The first factor apperaring in all the considered models is player's experience. We expect it to be positively correlated with the salary, however, player's physical conditions tend to decrease with any additional year after a certain moment, they become more inclinable to injuries and need more time to recover after the games and trainings which seriously reduces their usefulness for the team. For this reason, the square of experience is also added into the model.

Two other variables common for all our datasets refer to the date of contract signature and the contract length. Since the size of the salary cap as well as the luxury tax threshold are reviewed every year, we add the dummies for the year when the corresponding contract was signed in order to count for possible trend. The second variable is added in order to investigate if there is a discrimination with respect to the long-term contracts.

Variables specific to individual leagues are discussed separately.

4.2.1 National Basketball Association

4.2.1.1 Restrictions and Dataset Specification

Our dataset includes contracts signed in the seasons 2015-2018 and represent the most recent time frame among the research performed in this field. Since we consider multi-year contracts, the average contract value is analyzed. The average annual salary is obtained by dividing the actual contract value with a guaranteed part included by the length of the contract.

For the estimation of players' efficiency, we use the average performance indicators calculated based on the annual performance data available on www.nba.com. The data are collected for each of three seasons prior to the contract signature year and then the corresponding averages are used in the analysis.

Players participated in less than 20 games per season on average are not included in our dataset as their performance cannot be reliably estimated based on this small number of played games.

For our sample we consider all the non-Afro-American athletes to be white, including Hispanic players. Arabic and Asian players are excluded from the dataset due

to the low share in the league racial composition. In case of mixed parents, players with at least one black parent are considered black.

4.2.1.2 Performance Statistics

For the National Basketball Association, performance statistics are collected for the following variables.

Player's importance for the team is represented by the number of games played per season. This variable demonstrates teammates' and coaches' trust towards the player as well as player's ability to maintain the necessary physical conditions and avoid serious injuries during the whole season and is expected to be positively correlated with player's earnings.

The number of points per game is one of the crucial indicators in our model. Since the winner of the game is determined according to the scored points, talented scorers are highly appreciated in almost every NBA team nowadays.

The next performance variable indicates the number of rebounds made per game. Gaining a control of the ball after a missed shot or a free throw which bounces from the board or rim is referred to as a rebound. Rebounds are defined as offensive and defensive with respect to the board at which it was performed. For our model, the total number of rebounds is considered. This statistic is an inherent attribute to taller players such as centers and forwards. It has a favorable effect on the team's results and is expected to be positively correlated with the monetary compensation, since it is supposed to decrease the number of points missed, while offensive rebounds provide the team with a second chance to attack.

Turnovers per game are another variable in our regression. Since it represents the loss of possession for the team, it is expected to be negatively correlated with the player's wage.

The last performance variable considered in the context of NBA is the number of committed fouls. The impact of this variable on the players' wage is fairly ambiguous and depends on the game episode. Fouls committed by mistake or due to the lack of defensive skills have a negative impact on team results, however, some players are appreciated for their aggressive defensive manner and might commit more fouls on

average, furthermore, depending on the situation, preventive foul can be the wisest possible solution.

4.2.1.3 Other Related Variables

Several dummy variables invented to catch the effect of other factors than player's performance are introduced in our dataset.

The key variable in our research is race. For our NBA dataset, we introduce a variable *white* equal to 1 in case the considered player is a representative of white ethnicity and 0 otherwise. This variable is supposed to demonstrate the existence of wage premium for white athletes and determine the conclusion of this research.

Another dummy variable that appears in our model relates to a star status of the player. The star status is expected to increase the player's earnings, since the popularity of the player is determined by fans and athletes with such status are considered key players for their team. Since this term is relatively subjective and is not included in the official NBA statistics, we need to introduce our own measure to determine the star status of the player. For our dataset, the player is treated as a star in case he participated in at least one NBA All-Star game in some of the three seasons before the contract signature.

Since centers were found to have a wage premium in some of the previous studies on the NBA datasets (Hill, 2004; Pacak, 2016), the dummy variable *center* is introduced to check if there is a wage premium for the representatives of this position. It equals to 1 if the player's position is center and 0 otherwise. According to Hill (2004), taller players received an additional premium and since most of the centers present in the league were white, it resulted in the higher earnings for white players in general. Additionally, we add the dummy variable *guard* for both point and shooting guards in order to investigate if there is a difference between earnings with respect to a specific position. If neither of these variables is equal to 1, the player is considered a power or shooting forward.

4.2.2 Major League Baseball

4.2.2.1 Restrictions and Dataset Specification

For the analysis of the Major League Baseball, we consider two different datasets. The first is dedicated to the pitchers' performance, while the other includes batting players.

Since the MLB regular season starts in spring and continues until autumn, the free agents' contracts signed between the seasons 2014 - 2018 are added in our dataset.

Average players' performance during three years before the contract signature is calculated based on the statistics available on www.baseball-reference.com. Players with missing seasons are excluded from the datasets. Batters dataset consists of players with a minimum of 30 games, while the lowest required number of games for pitcher is 15.

Since the player is considered an unrestricted free agent after 6 years spent in the league, players with less than six years of experience at the moment of signing the contract are not included in our dataset.

Variables reflecting other players' qualities are similar for both MLB datasets.

Asian players are not included in our datasets due to the small share in the racial composition.

4.2.2.2 Performance Statistics

Since in each inning of a baseball game both teams play in defense and offense, two types of descriptive statistics are used: batting and pitching.

In order to better estimate players' batting performance, the following statistics are added into our model: batting average, on base percentage, home runs and stolen bases.

Batting Average is calculated by dividing the number of hits, when the batter reaches base after hitting a ball to a fair territory, by the number of at bats. For a long time, batting average has been considered a primary indicator of players' performance since it directly impacts team's ability to earn points in the offensive part of the inning. On Base Percentage indicates how often a batter reaches a base and is approximately calculated as the number of times the player reaches the base divided by the times he

appears on the plate. Home run is probably the most spectacular element of the baseball game known not only to baseball fans. Home run appears when the batter touches all four bases without any fielding error form the opponent side. Finally, stolen base occurs when the runner reaches the base that he is not entitled to. Since all these variables have a positive impact on the team's ability to score, they are expected to be positively correlated with players' earnings.

To estimate the pitching performance, variables earned run average, strike outs and win/loss ratio are added into our second regression.

Earned run average is considered a primary measure of pitcher's performance. This statistic is expressed as the runs earned by the opponent team times the number of innings in the game and divided by the innings pitched. This variable is expected to have a negative correlation with a player's salary, since a large number of earned runs by the opponent results in more resultative points for the opposing team. Strike outs are recorded when a pitcher throws any combination of three strikes to a batter and successfully defends against the opponent's offense. Lastly, win/loss ratio indicates the number of games where the pitcher's team takes a lead and wins the game divided by the number of games where the opposing team wins a game after taking a lead while the pitcher is in the game. Both strike outs and win/loss ratio are expected to be positively correlated with the player's compensation, since they indicate the efficiency of the team's defense.

4.2.2.3 Other Related Variables

Variables reflecting other players' qualities are similar for both MLB datasets, except the dummy variable for the starting pitchers in our defensive sample. Since this sample consists of starting and relief pitchers, a dummy variable equal to 1 if the corresponding player tends to start the games as a team pitcher and 0 otherwise is present in our regression.

Due to a large number of Hispanic players in the Major league Baseball two dummy variables related to player's race are added into our regressions. Variables *black* and *hispanic* indicate all the non-white athletes in our datasets, they are equal to 1 when the considered player is Afro-American or Hispanic respectively and 0 in case of a white player.

As for the other two leagues present in this paper, the dummy variable indicating a superstar status of the athlete is included in our dataset. For our MLB dataset, we consider a player to be the star if he was among the participants of at least one MLB All-Star Game during three years before signing the current contract, in this case variable *star* is equal to 1.

4.2.3 National Football League

4.2.3.1 Restrictions and Dataset Specification

For the NFL it is difficult to rate the quality of the player's performance based on the statistics for most positions. In order to provide a better estimation of players' efficiency we will create two different samples: one of them will include defensive positions, while the other will consist of the offensive positions. The first sample includes the data on the statistical performance of defensive linemen, defensive backs and linebackers. Wide receivers and tight ends are investigated in the second sample.

For each of the samples we use different models and run two separate linear regressions, since the corresponding performance statistics are different for both groups.

The NFL unrestricted free agents' contracts signed in the period of 2015-2018 are considered in our dataset. The corresponding performance statistics for three seasons before the signature of the contract are obtained from www.pro-football-reference.com and used to calculate the average performance.

Since only the players with four or more years of the NFL experience are able to become the unrestricted free agents after the termination of their contract, athletes with less than 4 years of experience at the moment of signing a current contract are not included in our dataset.

4.2.3.2 Performance Statistics

Since offensive and defensive samples are represented in our NFL dataset, two groups of performance variables are considered in this section.

The set of offensive variables is supposed to estimate the performance of wide receivers and tight ends and includes the average number of receiving yards, yards per reception and touchdowns.

Receiving yards are defined as the number of yards achieved on a passing play. This indicator includes the yards gained by the pass and additional yards gained after the reception, while yards per reception count for the number of yards gained for every reception made by the athlete.

The touchdown occurs when the player crosses the opponent's goal line with the ball. Touchdowns are the most resultative offensive actions in the game, since the player performed a successful touchdown brings six points to the team.

The group of the defensive variables consists of total tackles, sacks, interceptions, defended passes and forced fumbles.

The tackle occurs when the defensive player tackles the opponent with the ball to the ground or forces him out of bounds without interfering the rules. Total tackles include both solo tackles and tackle assists, which happen when the opponent is tackled by two defensive players. Sacks are referred to as tackles appearing when the defensive player tackles the opposition's passing player behind the scrimmage line before he can throw a pass.

Passes defended count for the number of passes stopped by the defensive player who blocks the ball, while the interception occurs, when the pass is caught by the defensive player, which results in the change of possession.

Fumble is a statistic indicating how many times a player possessing the ball accidentally loses it during the play. Considered variable *forced fumble* indicates how many times on average the defensive player forced the opponent to fumble the ball.

All the performance variables described above reflect offensive and defensive skills of the corresponding player and are expected to be positively correlated with the players' earnings as well as the number of games played per season which is present in both NFL related regressions.

4.2.3.3 Other Related Variables

Variables indicating factors which are not directly related to the players' performance are the same for both NFL samples. To catch the effect of other factors, variables reflecting body mass index and superstar status are included into our NFL regression in addition to the experience, race and contract related variables which are common for each league dataset.

Since the racial composition in the National Football League is represented by white and Afro-American athletes, variable *white*, equal to 1 in case of a white athlete is added into the regression.

To emphasize the superstar status of the NFL athlete, we implement the approach similar to the one used in the previous sections. The NFL athlete is considered a star, in case he attended at least one Pro Bowl, the NFL analogue of the All-Star game, during three seasons prior to the first contract year.

Finally, the body mass index variable is considered in our regression since it serves as a measure of player's physical development and is used in the context of the athlete's physical characteristics such as strength and speed.

4.3 Datasets Characteristics

This section reviews the general characteristics of the created datasets.

4.3.1 National Basketball Association

The dataset related to the National Basketball league consists of 236 players selected based on the above-mentioned criteria. Afro-American part of the sample is represented by 174 players and stands for more than 73% of the overall sample. The sample contains 22 athletes with the star status.

Table 4.1. National Basketball Association

Sample Means			
Variable	**Afro-Americans**	**Whites**	**Overall**
Age	29.94	30.00	29.98
Contract length	2.24	2.14	2.17
Experience	7.95	8.44	8.31
Games	64.20	63.93	64.00
Total rebounds	4.37	4.06	4.14
Turnovers	1.18	1.30	1.27
Fouls	1.88	1.88	1.88
Points	9.28	9.86	9.71
Average Salary ($)	7 907 950	8 157 411	8 091 875

4.3.2 Major League Baseball

Since each of the National Football League and Major League Baseball datasets consist of two different samples, each of them is described separately.

The sample containing batting players is formed by 140 athletes of Afro-American, White and Hispanic ethnicity. Majority of the sample is represented by white athletes, which contribute to approximately 51% of the sample, while the other half of the sample is evenly distributed between Afro-American and Hispanic athletes with a slight preponderance of the latter. The star status is assigned to 33 athletes.

Table 4.2. Sample Means

MLB Batters				
Variable	**Afro-Americans**	**Whites**	**Hispanics**	**Overall**
Age	32.38	32.69	32.06	32.46
Contract Length	2.03	1.80	1.74	1.84
Experience	9.91	9.39	9.89	9.64
Batting Average	0.26	0.24	0.25	0.25
On Base (%)	32.38%	31.64%	32.90%	32.14%
Home Runs	10.76	12.25	14.01	12.33
Stolen Bases	11.48	3.15	4.74	5.57
Average salary ($)	6 603 039	5 293 639	6 230 952	5 845 964

The second MLB sample consists of 139 pitchers with the overwhelming majority of white players. While 68,5% of the pitchers' dataset is given to the white athletes, Hispanic and Afro-American athletes represent roughly 13% and 15% of the sample respectively. The sample includes 48 starting pitchers and 91 relief pitchers, 26 of which were selected for the All-Star Game in the previous seasons giving them a star status.

As we can see from the Tabel 4.3, Afro-American athletes left behind with a big handicap their white and Hispanic colleagues in terms of strikeouts and possessed better win-loss ratio as confirmed by the corresponding t-test.

Table 4.3. Sample Means				
MLB Pitchers				
Variable	**Afro-Americans**	**Whites**	**Hispanics**	**Overall**
Age	31.84	32.82	33.27	32.76
Contact Length	2.58	2.01	1.36	1.99
Experience	9.37	9.57	10.14	9.63
ERA	3.79	3.79	4.22	3.86
Strikeouts	105.33	79.63	79.59	83.13
Win/Loss Ratio	0.51	0.49	0.48	0.49
Average Salary ($)	9 154 825	6 643 146	4 434 091	6 636 835

4.3.3 National Football League

Our NFL dataset includes two samples: offensive and defensive.

The offensive sample consists of 130 players: 43 tight ends and 87 wide receivers with 14 athletes given a star status. For both of our samples the majority is represented by African-American athletes and in case of the offensive positions Afro-American players stand for almost 67% of the overall sample.

Table 4.4 provides an overview of the sample means. As shown below the Afro-American athletes demonstrated better performance in terms of the receiving yards and yards per receptions as reflected in the corresponding statistics, while the white players possessed on average a higher body mass index.

Table 4.4 Sample Means			
NFL Offense			
Variable	**Afro-Americans**	**Whites**	**Overall**
Age	28.36	28.96	28.48
Contract Length	1.88	1.88	1.88
Experience	6.20	6.65	6.29
BMI Index	27.90	28.99	28.11
Games	13.65	12.60	13.44
Receiving Yards	441.63	302.95	413.89
Yards per Reception	12.57	10.92	12.24
Touchdowns	2.84	2.50	2.77
Average Salary ($)	3 349 872	2 862 340	3 252 365

The defensive sample consists of 469 athletes: 179 defensive backs, 142 linebackers and 148 defensive linemen. The sample includes 419 Afro-American and 50 white athletes standing for 89.3 and 10.7 of the sample respectively as well as 39 Pro Bowl participants, referred to as stars.

The respective sample means are demonstrated in Table 4.5. As in the previously described sample, white players provided a poorer performance in most of the game aspects. According to the results of the t-test, they were worse in terms of interceptions, defended passes and forced fumbles.

Table 5. Sample Means			
NFL Defense			
Variable	**Afro-Americans**	**Whites**	**Overall**
Age	28.68	28.54	28.66
Contract Length	1.73	1.94	1.75
Experience	6.11	5.60	6.06
BMI Index	31.31	32.27	31.42
Games	13.03	13.57	13.09
Sacks	1.16	1.32	1.18
Tackles	36.80	31.53	36.24
Interceptions	0.57	0.21	0.53
Defended Passes	3.03	1.27	2.84
Forced Fumbles	0.49	0.35	0.47
Average Salary ($)	2 736 321	2 226 217	2 681 938

Chapter 5

Methodology and Model

This chapter provides an overview of the general approach and describes the models selected for our paper.

5.1 Methodology

The choice of a methodology for the research is dependent on the type of the collected data. Samples created for each of the league compared in this thesis represent the data on a certain number of players taken from four different seasons. Since we consider different time periods in our models and have a random sample of players for each specific season, our data are referred to as a combination of cross-sectional and time series type. Pooled cross-sections appear to be the most appropriate solution in this situation, since panel data approach requires the same sample in every investigated period, while each of the seasons in our datasets is represented by different group of players (Wooldridge, 2012).

Using cross-sectional data provides two advantages. Firstly, collecting data from several years instead of a single season enables us to include a larger number of observations in our samples. Secondly, we are able to include additional dummy variables identifying the year of the observation in order to control for possible effect of the inflation or time trend, since the data on players' earnings are provided in nominal, but not real dollars.

Most of the papers referred to in this thesis apply the Ordinary Least Squares (OLS) method for model estimation. The same approach is implemented in our research. Validity of the selected model requires OLS assumptions to hold with respect to our data. Since the investigated samples contain all the players satisfying the criteria set in advance and the relation between the explained and explanatory variables in all our models is assumed to be linear, the main threats for our models are represented by the possible presence of serial correlation, multicollinearity and heteroscedasticity. Since each sample related to the specific year in our datasets consists of different

players, no player appears in the corresponding dataset more than once, which rules out the autocorrelation. The assumptions of no perfect collinearity among the independent variables is verified using the Variance Inflation Factor (VIF) test. All variables exceeding the maximum acceptable level of 5 by the results of the test (Ringle et al., 2015) were excluded from the regressions. Finally, the assumption of homoscedasticity requires constant variance of the error terms in our models. In order to account for possible presence of heteroscedasticity, clustered standard errors at the team level are calcuated for each regression and presented along with OLS standard erros in the futher part of the thesis. The choice of this method is logically justified, since the expenditures on the players‘ contracts in the teams are affected by the internal regulations of each league as well as the size of the team budget and increase in one player's salary might result in the lower contracts offerred to his teammates, which means that standard errors can be correlated within the team.

5.2 Models

This part of the chapter introduces the models created for all the considered samples.

Before we move to the introduction of the selected models, the characteristics of the dependent variable should be explained. In all the models the natural logarithm of the players' salary is used. This transformation allows us to investigate the percentage change in salary resulted from the change in the explanatory variables rather than change in the physical amount of dollars paid to the athletes.

All our models share the same structure with small game-related adjustments. They include both numerical and dummy variables. Since we use the logarithmic transformation of the dependent variable, one unit change in the numerical variable represents $\beta_k * 100\%$ change in the explained variable, where k stands for k-th variable in the model. For example, every additional game in the player's statistics affects player's salary *by* β_1*100% in the NBA model. The nature of the dummy variables are different. They do not represent any numerical value and impact the dependent variable by β_{dummy}*100% in case the player posseses the estimated quality and have no effect on the explained variable otherwise.

Since each group of players is estimated based on the different aspects, regressions constructed for each of the leagues are introduced separately.

National Basketball Association

$$Log(Salary) \sim \beta_0 + \beta_1 games + \beta_2 points + \beta_3 rebounds + \beta_4 turnovers + \beta_5 fouls + \beta_6 experience + \beta_7 experiencesqr + \beta_8 star + \beta_9 center + \beta_{10} guard + \beta_{11} length + \beta_{12} white + \delta_0 y2015 + \delta_1 y2016 + \delta_2 y2017$$

Major League Baseball

Model for batters:

$$Log(Salary) \sim \beta_0 + \beta_1 homeruns + \beta_2 onbasepercentage + \beta_3 battingaverage + \beta_4 stolenbase + \beta_5 experience + \beta_6 experiencesqr + \beta_7 star + \beta_8 length + \beta_9 hispanic + \beta_{10} black + \delta_0 y2015 + \delta_1 y2016 + \delta_2 y2017$$

Model for pitchers:

$$Log(Salary) \sim \beta_0 + \beta_1 strikeouts + \beta_2 era + \beta_3 winnloss + \beta_4 experience + \beta_5 experiencesqr + \beta_6 star + \beta_7 length + \beta_8 startpitch + \beta_9 hispanic + \beta_{10} black + \delta_0 y2015 + \delta_1 y2016 + \delta_2 y2017$$

National Football League

Offensive model:

$$Log(Salary) \sim \beta_0 + \beta_1 games + \beta_2 bmi + \beta_3 receivingyards + \beta_4 yardsperreception + \beta_5 tds + \beta_6 experience + \beta_7 experiencesqr + \beta_8 star + \beta_9 length + \beta_{10} white + \delta_0 y2015 + \delta_1 y2016 + \delta_2 y2017$$

Defensive model:

$$Log(Salary) \sim \beta_0 + \beta_1 games + \beta_2 bmi + \beta_3 sacks + \beta_4 tackles + \beta_5 interceptions + \beta_6 forcedfumbles + \beta_7 passesdefended + \beta_8 experience + \beta_9 experiencesqr + \beta_{10} star + \beta_{11} length + \beta_{12} white + \delta_0 y2015 + \delta_1 y2016 + \delta_2 y2017$$

Chapter 6

Results

This chapter analyzes the results of the regression models introduced in the previous part.

6.1 National Basketball Association

Nine variables appear to be significant by the results of the NBA related regression. The value of the adjusted R-squared indicates that approximately 71% of the variation in the dependent variable is explained by the variation in the independent variables.

As expected, variable *points* is highly significant with approximately similar value of both OLS and clustered standard errors. For most of the players this variable is supposed to be crucial in the context of their overall performance estimation and in our case, each additional point in the player's average statistics leads to roughly 5 % salary increase. *Rebounds* and *games* are two other performance related variables demonstrating a high level of significance close to zero with respect to both types of the calculated errors. Both of them predictably possess positive coefficients, while each additional game played on average for the period of three seasons increased the salary by approximately 0.9% and every offensive or defensive rebound provided 8.9% increase in player's wage.

Variable *experience* is significant at 0.05 level for both errors calculation methods. Based on the regression results player's salary is increased every season by 7.6%. The significance of variable *length* can be explained by the fact that long-term contracts are usually offered to the most perspective and important players of the team, while short contracts might demonstrate a certain degree of risk aversion in case of players having poor career background or more inclinable to injuries than their colleagues. Due to the essential increase of the salary cap after the season 2015-16 high significance level of the dummy variable *y2015* is not surprising. Contrary to the results of some of the previous studies, the reverse discrimination against centers is detected by

our model. The variable coefficient indicates 27% smaller wages for the athletes representing this position.

Our main variable of interest *white* is far from being statistically significant with p-value of 0.834 and does provide sufficient evidence of racial wage discrimination in the NBA.

Table 6.1 National Basketball Association					
Variable	**Estimate**	**OLS Std.Errors**	**Sign.Level**	**Clustered Std.Errors**	**Sign.Level**
(Intercept)	12.8710	0.2450	***	0.2430	***
Star	0.0240	0.1496		0.1870	
Center	-0.2726	0.1214	*	0.1562	.
Guard	0.0892	0.1031		0.1328	
White	0.0181	0.0863		0.0674	
Length	0.4806	0.0348	***	0.0325	***
y2015	-0.5642	0.1115	***	0.1109	***
y2017	-0.0497	0.1058		0.1085	
y2016	-0.2151	0.1003	*	0.1090	*
Experience	0.0764	0.0384	*	0.0308	*
Expersqr	-0.0037	0.0019	.	0.0016	*
Games	0.0085	0.0032	**	0.0027	**
Rebounds	0.0887	0.0283	**	0.0341	**
Turnovers	-0.1065	0.1070		0.1114	
Fouls	0.0455	0.0969		0.0746	
Points	0.0528	0.0148	***	0.0142	***
Adjusted R-squared: 0.7106					
N=236					
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

6.2 Major League Baseball

The results of each MLB regression are discussed separately.

Batters

We start with the model constructed for the sample of batting players.

As demonstrated in the Table 6.2, four independent variables are significant. Independent variables selected for the model determine roughly 68% of the total variation in the players' salary as reflected by the corresponding R-squared.

Performance variables indicating the average number of home runs and batting average ratio are found to be significant at 0.001 level for both types of standard errors, while the average number of stolen bases possesses lower clustered errors compared to the OLS. Every additional home run gained on average for the period of three seasons increased the batter's salary by almost 4%, while increase in the batting average by one percent point resulted in 9% salary increase. Each stolen base tends to add a bonus of 1.6% to the player's wage. Notably, clustered standard errors in case of the variable indicating player's on base percentage are more than 1.5 times lower than OLS disturbances. According to the regression results, on base percentage improved by one percent point increased player's wage by approximately 2.6%.

Variable *length* is highly significant and demonstrates that athlete's average salary is increased by 31% with each additional year negotiated before the signing. The dummy variable *y2017* is significant and shows a 40% difference between the contracts signed in 2017 and 2018, which might seem counterintuitive since the luxury tax threshold in the Major League Baseball was higher in 2018, considered a base year for our models.

Variables *hispanic* and *black* possess negative signs detecting the lower wages for Hispanic and Afro-American players, however neither of these variables are significant, which does let us claim the act of discriminating behavior towards non-white batters.

Table 6.2 Major League of Baseball: Batters					
Variable	**Estimate**	**OLS Std.Error**	**Sign. Level**	**Clustered Std.Error**	**Sign.Level**
(Intercept)	11.1119	0.8594	***	0.7087	***
experience	-0.0619	0.1204		0.1259	
expersqr	0.0022	0.0054		0.0059	
y2015	0.0708	0.1590		0.2034	
y2016	-0.0390	0.1521		0.1726	
y2017	0.4055	0.1502	**	0.1768	*
black	-0.0949	0.1540		0.1230	
hispanic	-0.0392	0.1347		0.1742	
star	0.1094	0.1576		0.1615	
length	0.3137	0.0472	***	0.0517	***
hruns	0.0381	0.0076	***	0.0069	***
ob	2.6223	2.0471		1.2789	*
batavg	8.8948	2.7641	**	2.8962	**
sbase	0.0164	0.0089	.	0.0069	*
Adjusted R-squared: 0.6767					
N=140					
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Pitchers

The adjusted R-squared is slightly lower for our pitcher's model compared to the previous one since only 64% of the variation in the regressand is determined by the variation in the regressors.

Two out of three performance variables appear to be significant in this model. Variable *era* predictably has a negative sign with every additional one-point increase in the ratio reduces player's salary by almost 17%, while every additional strikeout leads to 0.5% increase in the corresponding pitcher's salary.

As in the previously discussed model, variables counting for contract length and a year dummy are found to be significant. However, in this case, the negative coefficient of variable *y2017* seems to be more logical.

Finally, neither of the race dummy variables appears to be significant.

Table 6.3 Major League of Baseball: Pitchers					
Variable	**Estimate**	**OLS Std. Error**	**Sign. Level**	**Clustered Std. Error**	**Sign. Level**
(Intercept)	14.1477	0.6715	***	0.9326	***
black	0.0056	0.1554		0.1029	
hisp	-0.0435	0.1446		0.1400	
star	0.1458	0.1545		0.1000	
experience	0.1113	0.1036		0.1417	
expersqr	-0.0047	0.0046		0.0065	
startpitch	0.1310	0.1495		0.2113	
length	0.3623	0.0544	***	0.0583	***
era	-0.1684	0.0672	*	0.0799	*
strikeouts	0.0054	0.0017	**	0.0017	**
winloss	0.0298	0.4058		0.4036	
y2015	-0.2055	0.1485		0.1993	
y2016	-0.2863	0.1311	*	0.1089	**
y2017	0.2431	0.1493		0.1506	
Adjusted R-squared: 0.6435					
N=139					
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

6.3 National Football League

Results of the NFL related models are introduced separately.

Offense

Table 6.4 provides the results of the regression related to the offensive NFL sample. Performance related variables indicating players body mass index and average number of receiving yards are significant at 0.001 and 0.01 levels respectively. Each additional BMI point provided the athlete with 8.4% salary increase, while each receiving yard gained on average in the last three seasons increased his salary by 0.09%. Variable *tds* appears to be slightly significant implying 8% bonus for every performed touchdown.

All the year dummies added into the regression are found to be significant and demonstrate the development of the trend with respect to the salary cap change, while variable *length* reflects almost 50% wage difference for the long-term contracts.

The sign of variable *white* might indicate the case of reverse discrimination among wide receivers and tight ends, however it does not possess sufficient significance to be a reliable indicator of the unequal treatment.

Table 6.4 National Football League: Offense					
Variable	**Estimate**	**OLS Std.Error**	**Sign. Level**	**Clustered Std.Error**	**Sign. Level**
(Intercept)	11.1967	0.9822	***	0.7722	***
white	-0.0230	0.1296		0.1216	
experience	-0.0975	0.0845		0.0811	
expersqr	0.0052	0.0051		0.0045	
star	-0.0019	0.1724		0.2072	
length	0.4788	0.0440	***	0.0352	***
bmi	0.0844	0.0302	**	0.0272	**
games	-0.0217	0.0250		0.0372	
recyds	0.0009	0.0004	*	0.0004	*
yrec	0.0320	0.0200		0.0213	
tds	0.0830	0.0487	.	0.0471	.
y2015	-0.2705	0.1393	.	0.1287	*
y2017	-0.2649	0.1236	*	0.1483	.
y2016	-0.5024	0.1403	***	0.1336	***
Adjusted R-squared: 0.6532					
N=130					
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Defense

The last regression in this thesis is conducted on the data on the NFL players representing defensive positions. This model demonstrates one of the highest goodness of fit among our models with adjusted R-squared equal to 0.73, outperforming the offensive model in terms of determination. Variables *passdef* and *sacks* are significant at the highest level among the performance variables and provide a wage increase of 4.5% and 12% for every gained statistical point respectively. The number of tackles and body mass index are less significant with the coefficients revealing 0.2% and 1.1% wage difference for every performed tackle and bmi point accordingly.

Dummy variables standing for the year of the contract signature are found to be significant at the highest possible level, while variable *length* reflects 50% bonus for every additional year of the signed contract similar to the results of the previously

considered NFL regression, confirming the overall NFL trend. Contrary to the results of the other regressions variable *star* appears to be highly significant representing almost 37% higher salary for the players possessing a star status probably due to the highest number of star players in the NFL defensive sample compared to the other leagues and subsets.

Finally, regression run on the sample composed of the NFL defensive players appears to be the only one, indicating the presence of discriminatory behavior towards the athletes. The negative sign detects the reverse discrimination with 10% wage premium paid to the Afro-American athletes, however it is found to be significant only at 0.1 level.

Table 6.5 National Football League: Defense

Variable	Estimate	OLS Std. Error	Sign.Level	Clustered Std. Error	Sign.Level
(Intercept)	12.7366	0.2389	***	0.2679	***
White	-0.1085	0.0664		0.0640	.
y2015	-0.3160	0.0568	***	0.0531	***
y2016	-0.2320	0.0566	***	0.0563	***
y2017	-0.2048	0.0538	***	0.0527	***
star	0.3673	0.0832	***	0.0920	***
experience	0.0831	0.0409	*	0.0336	*
expersqr	-0.0053	0.0027	*	0.0023	*
games	-0.0044	0.0088		0.0101	
length	0.5021	0.0200	***	0.0173	***
bmi	0.0112	0.0054	*	0.0048	*
sacks	0.1233	0.0159	***	0.0213	***
tackles	0.0027	0.0011	*	0.0010	**
interceptions	0.0033	0.0403		0.0443	
passdef	0.0449	0.0107	***	0.0101	***
forcedfumb	-0.0153	0.0431		0.0522	
Adjusted R-squared: 0.7371					
N=469					
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Chapter 7

Conclusion

This thesis investigates the case of wage discrimination in the US sport industry in the seasons of 2015 - 2018 concerning the athletes from the major American leagues. We have collected the data on players' performance from the years 2012-2017 in order to provide the estimation of players' earnings based on their efficiency and reveal a case of possible discrimination in the considered leagues. Since most of the recent papers written on the topic of wage discrimination in American sport demonstrated mixed results, the main goal of this research was to compare the leagues in terms of the unequal treatment towards the athletes and provide a conclusion on the overall US sport industry level, contrary to the previous research approaching each league separately.

The overall results of the thesis do not find large support for the claim of racial discrimination in the Major US Leagues in the latest seasons. Applying OLS approach to all our samples, we found no presence of racial discrimination in the investigated leagues.

The results of our study might be further deepened by increasing the investigated samples and applying different econometric methods on the corresponding datasets. Since players' earnings during four seasons only are considered in this paper, extending the time range of collected data could provide more reliable estimates and possibly affect the significance of the variables included into the models.

Applying the Ordinary Least Squares method in the thesis, we did not consider other econometric methods which could possibly bring additional value to our results. Based on the research by Holmes (2011), Afro-American athletes received 25% less than other MLB players in the lower quantile. Further comparative analysis could be specified by applying a quantile regression in order to investigate the case of salary discrimination in different salary groups, which could provide more space for comparison between the leagues.

Another possible modification concerns the interaction variables. Since the degree of racial discrimination in our models is measured solely by the coefficients at the dummy variables indicating players' race, adding the interaction of the race and performance variables could reveal potential discrimination in terms of compensation

for player's performance.

Finally, since this thesis focused only on the "Big Four" leagues with the National Hockey League excluded from the analysis for the reasons mentioned in the first part of this paper, including other competitions, such as Major League Soccer, in the comparative analysis could bring additional value to the research.

.

Bibliography

Arias J.(2016) The Labor Market of Major League Baseball. A Test For Salary Discrimination. Retrieved from: https://www.academia.edu/26079767/The_Labor_Market_of_Major_League_Baseball_A_Test_For_Salary_Discrimination.

Arrow, K. (1971). The Theory of Discrimination. Princeton University, *Department of Economics, Industrial Relations Section.*

Becker, G. (1971): The economics of discrimination 2nd ed. ,*University of Chicago Press, Chicago.*

Berri, D. & Simmons, R. (2007). Race and the evaluation of signal callers in the national football league, *Working Papers 591147, Lancaster University Management School, Economics Department.*

Bodvarsson, O.B., & Brastow, R. T. (1998). Do Employers Pay for Consistent Performance? Evidence from the NBA. *Economic Inquiry, 36(1)*, 145–160.

Bodvarsson, O.B., & Pettman, S. (2002). Racial Wage Discrimination in Major League Baseball: Do Free Agency and League Size Matter? *Applied Economics Letters (Vol. 9).*

Brown, E., Spiro, R., & Keenan, D. (1991). Wage and Nonwage Discrimination in Professional Basketball: Do Fans Affect It? *American Journal of Economics and Sociology, 50(3)*, 333–345.

Burdekin, R., & Idson, T. (1991). Customer Preferences, Attendance and the Racial Structure of Professional Basketball Teams. *Applied Economics (Vol. 23)*.

Burnett, N. J., & Van Scyoc, L. J. (2015). Compensation discrimination in the national football league. *Journal of Economics and Economic Education Research (Vol. 16)*, 60-71

Dey, M. S. (1997). Racial Differences in National Basketball Association Players' Salaries: Another Look. *The American Economist, 41(2)*, 84-90.

Ducking, J.& A. Groothuis, P.,A. & Hill, J.R. (2014). Compensation discrimination in the NFL: an analysis of career earnings. *Applied Economics Letters, Taylor & Francis Journals, vol. 21(10)*, 679-682.

ESPN (2018). Retrieved from: http://www.espn.com/mlb/story/_/id/23132253/major-league-baseball-gets-average-marks-hiring-practices

GALLUP (2018). Retrieved from: https://news.gallup.com/poll/224864/football-americans-favorite-sport-watch.aspx

Gius, M., and Johnson, D. (1998). An Empirical Investigation of Wage Discrimination in Professional Basketball. *Applied Economics Letters, 5(11)*: 703-705.

Gius, M. & Johnson, D. (2000). *Race and compensation in professional football, Applied Economics Letters, 7:2*, 73-75

Gordon, S. (2015). An Examination of Wage Discrimination in the NBA. Retrieved from: http://thesportsquotient.com/nba/2015/5/19/an-examination-of-wage-discrimination-in-the-nba

Groothuis, P., & Hill, J. (2011). Pay Discrimination, Exit Discrimination or Both? Another Look at an Old Issue Using NBA Data. *Journal of Sports Economics (Vol. 14)*.

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

Hill, J. (2004). Pay Discrimination in the NBA Revisited. *Quarterly Journal of Business and Economics, 43(1/2)*, 81-92.

Holmes, P. (2011). New evidence of salary discrimination in major league baseball. *Labour Economics, 18(3)*, 320–331.

Irwin, W. (2004) "Racial Discrimination in Major League Baseball: Can It Exist When Performance is Crystal Clear? *The Park Place Economist (Vol. 12)*.

Jenkins, J. A. (1996). A Reexamination of Salary Discrimination in Professional Basketball. *Social Science Quarterly (Vol. 77)*, 594-608

Johnson, C. and Minuci, E., Wage Discrimination in the NBA: Evidence Using Free Agent Signings (2018). Retrievde from: http://dx.doi.org/10.2139/ssrn.3276247

Kahn, L. (1991). Discrimination in Professional Sports: A Survey of the Literature. *Industrial and Labor Relations Review, 44(3)*, 395-418.

Kahn, L. (1992). The Effects of Race on Professional Football Players' Compensation. *Industrial and Labor Relations Review,45(2)*, 295-310.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *The Journal of Economic Perspectives, 14(3)*, 75-94.

Kahn, L., & Sherer, P. D. (1988). Racial Differences in Professional Basketball Players' Compensation. *Journal of Labor Economics, 6(1)*, 40–61.

Keefer, Q. A. W. (2013). Compensation Discrimination for Defensive Players: Applying Quantile Regression to the National Football League Market for Linebackers. *Journal of Sports Economics, 14(1)*, 23–44.

Koch, J. and Vander Hill W. (1988) Is there discrimination in the `black man's game? *Social Science Quarterly, 69(1)*, 83-93.

L. A. Rapping and A. H. Pascal (1971), Racial Discrimination in Organized Baseball. *RM-6227, The Rand Corporation*.

Larsen, A., Fenn, A. J., & Spenner, E. L. (2006). The Impact of Free Agency and the Salary Cap on Competitive Balance in the National Football League. *Journal of Sports Economics, 7(4)*, 374–390.

Masella, J. (2017). Racial Salary Discrimination in Major League Baseball : A Closer Look. Retrieved from: https://www.semanticscholar.org/paper/Racial-Salary-Discrimination-in-Major-League-%3A-A-Masella/ec10753000c0be75797ee8d0d24d4f0faa11bf49

Medoff, M. H. (1975). A Reappraisal of Racial Discrimination against Blacks in Professional Baseball. *The Review of Black Political Economy, 5(3)*, 259–268.

Moffatt, (2018). Retrieved from https://www.thoughtco.com/the-economics-of-discrimination-1147202

Mogull, R. G. (1973). Football Salaries and Race: Some Empirical Evidence. *A Journal of Economy and Society, 12(1)*, 109–112.

Mogull, R. G. (1974). Racial Discrimination in Professional Basketball. *The American Economist, 18(1)*, 11–15.

Mogull, R. G. (1977). Racial Discrimination in Professional Basketball: A Reply. *The American Economist, 21(1)*, 60–62.

Mogull, R. G.(1981). Salary discrimination in professional sports. *Atlantic Economic Journal (Vol. 9)*, 106-110

Naito, H., & Takagi, Y. (2017). Is racial salary discrimination disappearing in the NBA? evidence from data during 1985–2015. *International Review of Applied Economics, 31(5)*, 651–669.

PACÁK, V. (2016). The Salary Discrimination in NBA. *Charles University, Faculty of Social Sciences, Institut of Economic Studies.*

Palmer, M., & King, R. (2006). Has Salary Discrimination Really Disappeared from Major League Baseball? *Eastern Economic Journal, 32(2)*, 285-297.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review, 62(4)*, 659-661.

Ringle, C. M., Wende, S., & Will, A. (2015). SmartPLS 3. Retrieved from: http://www.smartpls.com

Rockwood, C. E., & Asher, E. (1976). Racial Discrimination in Professional Basketball Revisited. *The American Economist, 20(1)*, 59–64.

Scott, .A., Long, .E. & Somppi, K.(1985). *Atlantic Economic Journal, 13(3)*, 50-59.

G. W. (1974). Pay and Performance in Major League Baseball. *American Economic Review, 64(6)*, 915–930.

Statista (2018). Retrieved from: https://www.statista.com/topics/2113/sports-on-tv/

The Institute for Diversity & Ethics in Sport, (2016). Retrieved from: http://www.tidesport.org/reports.html

Tygiel, J. (1997). Baseball's great experiment: Jackie Robinson and his legacy. *Oxford University Press, USA.*

Wallace, M. (1988). Labor Market Structure and Salary Determination Among Professional Basketball Players. *Work and Occupations, 15(3)*, 294-312.

Wooldridge, Jeffrey M., (2012): "Introductory econometrics: a Modern Approach." 5th ed. *Mason: Thomson/South- Western*

Zimbalist, A. (2003): "Sport as business." *Oxford Review of Economic Policy 19(4)*, 503–511.