University of New Mexico
UNM Digital Repository

Health, Exercise, and Sports Sciences ETDs   Education ETDs

Spring 4-3-2019

MODELS OF WAGE DETERMINATION OF OFFENSIVE LINEMEN IN THE NATIONAL FOOTBALL LEAGUE

Daniel S. Krywaruczenko

Follow this and additional works at: https://digitalrepository.unm.edu/educ_hess_etds

Part of the Health and Physical Education Commons

Recommended Citation
Krywaruczenko, Daniel S.. "MODELS OF WAGE DETERMINATION OF OFFENSIVE LINEMEN IN THE NATIONAL FOOTBALL LEAGUE." (2019). https://digitalrepository.unm.edu/educ_hess_etds/105

This Dissertation is brought to you for free and open access by the Education ETDs at UNM Digital Repository. It has been accepted for inclusion in Health, Exercise, and Sports Sciences ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact amywinter@unm.edu.

Daniel S. Krywaruczenko
Candidate

Health, Exercise and Sports Sciences (HESS)
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Alfredo Martinez, Chairperson

Dr. Lunhua Mao

Dr. Evan Frederick

Dr. David van der Goes

MODELS OF WAGE DETERMINATION OF OFFENSIVE LINEMEN IN THE NATIONAL FOOTBALL LEAGUE

by

Daniel S. Krywaruczenko

B.S., Mathematics (Minor in Business Administration). University of Tennessee, Magna Cum Laude, 2008
M.S., Statistics, Mississippi State University, 2010
M.A., Economics, University of New Mexico, 2015

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physical Education, Sports and Exercise Science

The University of New Mexico
Albuquerque, New Mexico

May, 2019

© 2019, Daniel S. Krywaruczenko

ALL RIGHTS RESERVED

MODELS OF WAGE DETERMINATION OF OFFENSIVE LINEMEN IN THE NATIONAL FOOTBALL LEAGUE

by

Daniel S. Krywaruczenko

B.S., Mathematics (Minor in Business Administration). University of Tennessee, Magna Cum Laude, 2008
M.S., Statistics, Mississippi State University, 2010
M.A., Economics, University of New Mexico, 2015
Ph.D., Physical Education, Sports and Exercise Science, University of New Mexico, 2019

ABSTRACT

This report reviews factors that can impact an offensive lineman's salary in the National Football League. For years, factors measured by performance statistics have been gathered in various sports as a method to estimate a player's production. Understanding performance statistics that impact a baseball player's salary was popularized by the movie *Moneyball*, but limited work in this regard has been done to measure effectiveness or efficiency of offensive linemen in American football. The lack of publicly available data and many interdependencies in football make it difficult to objectively understand how salaries can best be determined. This study uses a quantitative approach and a single-equation model with a unique data set to explore the determinants of wages of offensive linemen in an open market. The data set consists of explanatory variables measured by the previous season's individual statistics, team statistics, and statistics based on nonoffensive line positions along with characteristics of offensive linemen, such as the player's age. The study found that several variables impact the salary of an offensive lineman in free agency, such as the number of

games started. It is hoped that this study will provide a building block for additional research on wage determination in American football.

Table of Contents

List of Tables

Introduction

A play on offense in American football begins with the ball being snapped from the center to the quarterback. After the snap, if a quarterback moves backward on a pass play, the offensive line, consisting of five players: left tackle, left guard, center, right guard, and right tackle has a duty to stop the opponent's defensive players on the field from tackling or disrupting him (i.e., pass blocking) so that he can successfully throw the football to a wide receiver, tight end, or running back (Alamar & Weinstein-Gould, 2008). On a running play in which the quarterback typically hands off the ball to a running back, the offensive line's goal is to move or stop the opponent's defensive players on the field (i.e., run blocking) from impeding the player carrying the football.

The offensive line plays an important role on different running plays and passing plays that make up the offensive play selection for a football team. Using running plays and passing plays, the goal for the offense is to move the ball into the opponent's end zone without being stopped by the opposing defense. After each offensive play, decision-makers such as general managers question and evaluate how much each player's performance on the field contributed to the success of the play. For example, in a single football play in which an offensive running back runs for 20 yards, how much did his performance contribute to the success of the play? How much did the entire offensive line's performance contribute to the play's success? Or in a pass play of 50 yards to a wide receiver, how much did the performance of the quarterback, the wide receivers and offensive line contribute to the play's success? The answers to these questions are often subjective in that different decision-makers will value different players based on what they see just like scouts have different opinions on players based on what they individually see. In light of this dilemma, establishing the wages of players, specifically, that of an offensive lineman with only subjective measures may be problematic.

From a more objective perspective, the amount of success on the field produced by the labor of players is a factor that influences their respective wages. However, objectively measuring success and how much a player deserves for their labor on a team is still difficult. Labor, one of the four factors of production, is difficult to study within the sport of football and specifically with offensive linemen because of the interdependencies within the unit (e.g., teamwork between right guard and center on run block) and between units on offense (e.g., quarterback's decision-making skills when throwing ball and pass blocking by the offensive line) are difficult to measure and evaluate, and the availability of public metrics for measuring individual performance of offensive linemen is limited (Fizel, 2006). This paper is one of the first to analyze in a quantitative manner the interaction of performance between offensive linemen.

Quantitatively, a week-long annual showcase event for a select group of top professional football prospects known as the National Football League (NFL) Combine where, current metrics such as the 40-yard dash provide data to evaluate offensive linemen but do not measure the production of the offensive lineman in a game. Players go through physical tests, medical evaluations, interviews with NFL player personnel, and intelligence exams. Physical tests, such as the broad jump, help quantify the explosiveness of the athlete and provide data to talent evaluators to help project how well an ex-college football player such as an offensive lineman will perform in the NFL and thus be compensated (Parekh, 2017). On the field, coaches study game scenarios to evaluate how offensive lineman perform. For instance, a success/failure metric might be utilized by an offensive line coach to determine the proportion of play in which an offensive lineman succeeded in terms of his technique.

The method of grading offensive linemen such as the one mentioned above is unique and can differ from team to team. Additionally, many current metrics, which would be beneficial to evaluate offensive linemen and determine their wages, are not available to the public because coaches don't reveal grades on their players. Coaches do not want another team's coaches to know their evaluation of personnel because such knowledge could give the opposing team an advantage.

In addition to the status quo, history reflects the lack of public metrics as the compilation of statistics in American football began in the 1920s and did not differentiate between yardage earned via pass plays, running plays, or kick returns. The data lacked completeness because statistics such as kickoff returns and punt returns were not kept until 1941. Additionally, there was a lack of credibility with the statistics that were gathered because the game was faster paced than, for example, baseball. Thus, those in attendance had a difficult time compiling a statistic while watching the game live (first NFL televised game was in 1939). By 1935, more-complete and more accurate individual and team statistics were available, such as for rushing and passing. Still, the league lacked statistics on key situations, such as yardage gained by the opponent after an interception and after a blocked punt (Carroll, Palmer, & Thorn, 1989).

For the position of offensive lineman, statistics are scarce because their work is often not documented statistically to the public. NFL.com (the official site of the NFL) has a statistical board to enable fans to help gauge which players are performing their best at their respective position. However, the web site does not list statistics related to the performance of offensive linemen on its “League Leaders” board, while every other position, such as quarterbacks and wide receivers, are represented.

In addition to the lack of metrics and publicly available statistics, the joint-production nature of this team sport makes it difficult to measure what individual players contributed toward winning a football game and thus how they should be compensated (Fizel, 2006). Joint production is a distinct characteristic of the sport because football is a team sport of many interdependencies on every play. This is especially true within the offensive line, which functions as a unit of five players in run blocking and pass blocking. In total, 11 players work together to attempt to complete different tasks (e.g., blocking, throwing, catching, running) to move the ball forward in order to score a touchdown or field goal, while 11 players on the opposing team try to stop the other team from moving the ball down the field (Berri, Brook, & Schmidt, 2007). Even in situations where one offensive lineman makes a key block, the others need to also deliver effective blocks, or else the outcome of the play will be undesirable for the team on offense (Carroll et al., 1989). While the offensive line works as a single unit, NFL teams pay offensive linemen based on their respective evaluations of the individual player.

In a perfectly competitive market, many NFL teams would bid against each other for the services of an individual player, causing a player’s salary to approach the amount of revenue he produces for the team based on his individual productivity (Bradbury, 2008; Quirk & Fort, 1992). The economic metric for this is marginal revenue production (MRP), and much research on MRP has been done in other sports, such as baseball. In football, however, in light of the difficulty of connecting individual productivity to revenue based on the interdependency of the positions, one can look at production of other players at other positions as a factor in determining wages for offensive linemen. For instance, the financial compensation for a left tackle rose along with the compensation of quarterbacks (a position in which wage determination has been shown in past literature) because the left tackle was viewed as insurance on the owner's investment (Lewis, 2007).

The financial compensation for linemen such as the left tackle has grown dramatically as well. In 1957, the minimum salary for an NFL player was $5,000 per season. Many players during that and previous eras would hold offseason jobs to pay their bills (Michael, 2014). For offensive linemen, the top-paid players in the 1960s, such as Chicago Bears center Mike Pyle, were paid $14,000 per season. Long-time starting left tackle Billy Shields, who played in the 1970s and 1980s for a total of 11 seasons, was paid less than a total of $1 million over his entire career. According to a survey in 1981, the average salary of linemen was the lowest of all positions on offense ("Average N.F.L. Salary," 1982). The Cincinnati Bengals football organization told offensive lineman Anthony Munoz in the 1980s that no offensive lineman was worth $500,000 per season (Lewis, 2007). Literature in sport on racial discrimination brings into question whether racism on behalf of the Cincinnati Bengals was a determining factor in the devaluation of Munoz's worth. During his playing career, players won rights through strikes to earn a higher percentage of compensation from team revenue and to gain the ability to collectively bargain, which enabled average salaries to continuously grow to $414,900 per season in 1992 (average salary was $30,000 in 1977) to approximately $496,300 per season by the start of the 1993 season (Stellino, 1992).

With the introduction of free agency in 1993 came an increase in spending on the purchase of offensive linemen as free-agent pickups. During that year, lesser known offensive linemen such as Harry Galbreath signed a three-year contract worth $1.52 million per season. Top-rated free agent left tackle offensive lineman Will Wolford signed a unique contract in 1993 worth $7.65 million over three years with the Indianapolis Colts. The contract included a clause that guaranteed Wolford that he would be the highest paid offensive player annually on the team. Virtually no statistics were available at that time to measure his productivity as a player, yet he was compensated as if he was the most productive offensive player on the team (Lewis, 2007). With proper metrics, one can measure how much a player such as Wolford contributed to his team's success (i.e., marginal productivity) and then use this value to compensate him accordingly during free agency. By identifying and understanding wage determinants of offensive linemen through the present research, the general manager of a football team can gain a competitive advantage by better understanding the labor and nonlabor forces that impact salaries for offensive linemen.

Additionally, to this author's knowledge, this report is the first study of wage determination of its kind for offensive linemen. The response variable consists solely of salaries of offensive linemen who are unrestricted free agents. This models an open, competitive market. A unique mix of explanatory variables from different levels, such as the offensive unit, the offensive line unit (e.g., interaction component), and a unit for an individual player enables one to gain a thorough understanding of labor and nonlabor forces (such as play calling) that impact wages of offensive linemen. Future research and development of new individual performance statistics, coupled with a revenue function, might enable one to evaluate MRP of offensive linemen.

Research Topic: Wage Determination

A key situational component in determining wages is a person's experience in their labor market (Mincer, 1974). Dr. Jacob Mincer, credited as a pioneer for his single-equation model, used experience (variable T) as an explanatory variable in his earning's function:

(A1)

$$Ln(SAL) = \mathrm{Ln}(\mathrm{SAL_0}) + \alpha_1 S + \alpha_2 T + \alpha_3 T^2 + u_i$$

Where SAL = earnings, SAL_0 = earnings of someone with no education and no experience, S = years of schooling, and T = years of potential labor market experience (Mincer, 1974).

Log of salaries was taken because earnings are naturally positively skewed. T acts as an inverted "U" in the life cycle of earnings because it acts as a method of formal training and also brings about a deterioration of production (Mincer, 1958).

A person's earnings should approach their MRP, which is defined as the additional revenue credited to an additional worker in a perfectly competitive market (Bradbury, 2008; Quirk & Fort, 1992). Individual player productivity in sport can be measured by marginal production (MP), which is defined as the number of wins an individual generates for their respective team (Santo & Mildner, 2010). A common method used to determine production is linear weights (using ordinary least squares). Of key interest in the process of calculating MP is the team percent win function for the sport of baseball that was credited by Scully (1974a) as follows:

(A2)

$$PCTWIN_t = \mathrm{x_0} + \mathrm{x_1} TSA_t + \mathrm{x_2} TSW_t - \mathrm{x_3} NL + x_4 CONT_t - x_5 OUT_t$$

Where PCTWIN = percentage of wins, TSA = team slugging average, TSW = team strikeout-to-walk ratio, NL = one for playing in the National League (dummy variable), CONT = one for pennant or divisional winners (dummy variable), and OUT = one for teams that are 20 or more games out of first place in their division.

A key determinant in the wage model is individual performance. In that sense, TSA and TSW are our variables of interest. These performance variables have been attributed to individual performance of baseball players (Scully, 1974a). Improvement in TSA and TSW will impact PCTWIN (what's that?) in a positive manner, and an increase in PCTWIN will increase the revenue of a baseball team (Scully, 1974a). Thus, an individual's batting average and a pitcher's strikeout-to-walk ratio are individual statistics that act as wage determinants for a major league baseball player on a respective team. Once a player's MP is determined, a revenue equation, which is a function of MP, can determine the player's MRP, which is what a player produces in revenue for a particular team. In a perfectly competitive market, a player's salary would approach their MRP (Scully, 1974a).

Aside from individual performance, individual characteristic variables of players, such as race, are considered as a possible determining factor of wages. Racial discrimination occurs when people show bias against certain racial groups (Anderson & La Croix, 1991). The study of racial discrimination is of interest to many labor economists in sport because of readily available data on wage and performance statistics that are not common outside of the field of professional sport (Hakes & Sauer, 2006). Following Becker (1971), a standard approach to determine racial wage discrimination in sport is to estimate the following model:

(A3)

$$Y = \beta_0 + \beta_1 P + \beta_2 R + e_i$$

where Y = salary (or other decision variable); P = worker productivity, player characteristics, and market variables; R = race (dummy variable); and e_i = error term.

A significant value for β_2 lends evidence of discrimination, although this interpretation proves to be difficult, unless all factors in P are accounted for (Berri et al., 2007). The variables representing P and R act as wage determinants (Becker, 1971).

Research Problem: Wage Determination of NFL Offensive Linemen

In the NFL, offensive linemen play a key role in maximizing production, which general managers encourage through high salaries (Lewis, 2007). Yet, traditional game statistics, such as rushing yards and receiving touchdowns, are not available for this particular position that would enable the researcher to better understand wage determinants. However, this paper will contribute an additional avenue to evaluate offensive linemen performance by using a unique set of variables (individual and team) for measuring the wage determinants of offensive linemen in the NFL.

Additionally, an offensive line works as a single unit in itself (Carroll, Palmer, Thorn, & Pietrusza, 1998). This presents the difficulty of measuring other factors that influence individual productivity such as the general environment, in which there are several interdependencies occurring during each play (Bradbury, 2008). Further, Idson and Kahane (2000) found complementary effects (interaction) exist between inputs of labor (players), individual productivity could be poorly measured as one attempts to separate the individual from the team. The individual productivity of a given player might differ from team to team, depending on the amount of assistance a player receives in each environment. Such complementary effects in human-capital inputs create a team dynamic that impacts individual effects and the player’s respective compensation (Idson & Kahane, 2000). This paper will investigate possible complementary effects of offensive linemen.

Research Purpose: Consistency of measures and explanatory power of outcomes

The purpose of this study is to investigate wage determinants of NFL offensive linemen. This paper will utilize statistics from STATS LLC, Pro Football Reference, NFL.com, and Spotrac. The STATS LLC statistics, to my knowledge, have not been used in the literature.

Research Questions

1.) **How are the wages of different positions in the NFL offensive line affected by individual player performance?**

    1a.) Does the previous regular season proportion of knockdowns on quarterbacks by individual offensive linemen affect wages?

    1b.) Does the number of previous regular season games started by individual offensive linemen affect wages?

    1c.) Does the previous regular season proportion of penalties committed by individual offensive linemen affect wages?

2.) **How are the wages of different positions in the NFL offensive line affected by team performance?**

    2a.) Does the number of previous regular season rushing yards of the offensive lineman's team affect wages?

    2b.) Does the number of previous regular season passing yards of the offensive lineman's team affect wages?

    2c.) Does the number of previous regular season total points scored by the offensive lineman's team affect wages?

3.) How does relative performance of NFL offensive linemen affect their wages?

3a.) Do the previous regular season difference between a player's proportion of knockdowns allowed on quarterbacks and his offensive line teammates' have an impact on the player's wages?

3b.) Do the previous regular season difference between the player's proportion of knockdowns allowed on quarterbacks and his new team's proportion of knockdowns allowed on quarterbacks have an impact on the player's wages?

4.) How does the individual performance of nonoffensive line positions affect the pay of different positions in the NFL offensive line?

4a.) Do the previous regular season adjusted line yards of the offensive lineman's team affect wages?

4b.) Do the previous regular season running backs' yards before contact of the offensive lineman's team affect wages?

4c.) Do the previous regular season running backs' average yards per carry of the offensive lineman's team affect wages?

4d.) Does the previous regular season power percentage of the offensive lineman's team affect wages?

4e.) Does the previous regular season stuff percentage of the offensive lineman's team affect wages?

5.) How does the play selection of a team affect the pay of different positions in the NFL offensive line?

5a.) Does the previous regular season location run proportion of the offensive lineman's team affect wages?

5b.) Does the previous regular season run attempts to pass attempts ratio of the offensive lineman's team affect wages?

Delimitations

- Study focuses only on offensive linemen in the NFL.
- Uses publicly available data through STATS LLC, Spotrac, Pro Football Reference, and NFL.com.
- Data comes from NFL regular seasons of 2010-2016.

Limitations

- Data set from STATS LLC, Spotrac, Pro Football Reference, and NFL.com does not include elements such as current score, time remaining in game, down, and distance of play, which would enable one to measure a player's "clutch" ability and could be used for other wage determination metrics, such as expected points and win probability.
- No data is available on individual offensive lineman run-blocking performance; thus, run-blocking ability is measured through other means, such as team performance statistics.
- No data is available logging in number of offensive snaps per individual offensive linemen for 2010-2012 seasons; thus, offense snaps logged during these seasons are estimated based on games the player started.
- No data is available with regards to specific injuries of players, which can impact player performance on field.

Definition of Key Terms

Conceptual Definitions of Key Terms

Accrued season: a season that counts toward NFL free agency that one earns by being on a team's roster for six or more regular season games at full-pay status.

Defensive back: one of two defensive backs positioned in the secondary between the linebackers and safeties, responsible for covering the outside areas near the sidelines against end runs and pass plays.

Defensive end: a defensive player positioned across from an offensive tackle.

Defensive tackle: a defensive player positioned across from an offensive guard.

False start: an illegal movement by an offensive participant before the snap of the ball.

Fullback: a particular type of running back who is typically larger and more adept at blocking.

Halfback: a term that is synonymous with running back.

Holding: a person illegally restrains a person who is not holding the ball.

Linebacker: a player on defense who takes a position close behind the linemen.

Line of scrimmage: an imaginary line parallel to the goal lines that passes from one sideline to the other through the point of the football closest to the goal line of each team.

Marginal production: the number of wins a player contributes to a team over a season.

Marginal revenue production: the revenue produced for a team by the individual production of a player.

Offensive center: a specific position on the offensive line who is responsible for snapping the ball to the quarterback.

Offensive guard: a specific position on the offensive line that is lined up between the center and offensive tackle.

Offensive lineman: position on the offense whose primary job is to run-block on run plays and pass-block on pass plays.

Offensive tackle: a specific position on the offensive line that is lined up outside the offensive guard.

Quarterback: position on the offense that receives the ball at the start of the play from the center and is primarily responsible for communicating and orchestrating the play.

Running back: position on offense that traditionally lines up behind the offensive linemen.

Safety: a player on the defense who lines up farthest behind the line of scrimmage.

Tight end: an offensive player positioned at one extremity of the line directly beside a tackle, used as both a blocker and a pass receiver.

Unrestricted free agent: any player with at least four accrued seasons at the time his contract expires.

Wide receiver: an offensive player positioned wide of the formation, as a split end, used primarily as a pass receiver.

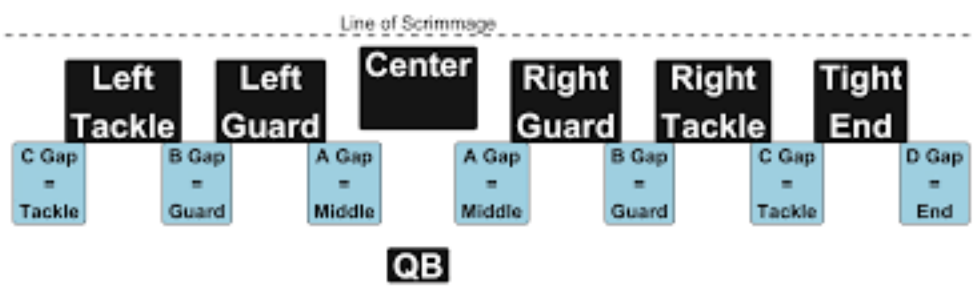
Operational Definitions of Key Terms

Adjusted line yards: metric created by Football Outsiders (football analytics company) based on regression analysis; the adjusted line yards formula takes all running back carries and assigns responsibility to the offensive line based on the following percentages:

- Losses: 120% value.
- 0-4 yards: 100% value.
- 5-10 yards: 50% value.
- 11+ yards: 0% value.

Knockdown: a statistic in football credited to a defensive player or players when physical contact with a quarterback causes him to fall down behind the line of scrimmage during a passing play.

Location run proportion: Proportion of runs where a running back runs to the left, right, or through a gap of an offensive lineman's position.



(Yurko, Ventura, & Horowitz, 2018).

Power percentage: "Percentage of runs on third or fourth down, two yards or less to go, that achieved a first down or touchdown. Also includes runs on first-and-goal or second-and-goal from the two-yard line or closer" (includes quarterback runs, football outsiders.com).

Sack: a statistic in football credited to a defensive player or players when physical contact with a quarterback causes him to be ruled down by contact behind the line of scrimmage, prior to an attempt to throw a forward pass or if the play is determined to resemble a passing play.

Stuff percentage: "Percentage of runs where the running back is tackled at or behind the line of scrimmage" (footballoutsiders.com).

Yards before contact: yards gained by a player rushing on offense prior to being contacted by a defensive player.

Review of Literature

Wage Determination in Sport

The proposed model includes four distinct arenas that previous literature suggests can lead to determining a fair-market assessment for employee salaries in sports- marginal revenue production (MRP), race discrimination (non-MRP, emphasis on wage discrimination), performance statistics, and the National Football League (NFL) Combine metrics. First, MRP relates to the proposed wage-determinants model due to the usage of performance variables in the production function. Further, in an open, competitive market, an offensive lineman's salary will approach their MRP (Kahn, 1992). Race discrimination (non-MRP, emphasis on wage discrimination) literature often encompasses the explanatory variable of race (to measure possible discrimination), with performance variables acting as controls, both of which are of interest to the researcher. The performance statistics section enhances the reader's understanding of metrics used in terms of evaluating a player's production, while NFL Combine metrics looks at common measurements of physical ability used in the league that act as another form of evaluation and prediction.

a.) Marginal Revenue Production (marginal production included)

No research in MRP in football has been investigated specific to the sport, to this author's knowledge. Other research in sports, such as baseball, will be reviewed because those academic works shed detailed light on important concepts covered in this dissertation. Research related to MRP has been done in basketball and hockey, posing more difficulties in estimating values of their respective professional athletes.

i.) Baseball

Scully (1974a) was the first to analyze MRP in sports, in baseball. He estimated the amount of economic loss for baseball players thought to be caused by the reserve clause, which restricted labor movement in Major League Baseball. In a perfectly competitive market, a player's salary approaches his MRP (Scully, 1974a).

To find player MRP values, Scully (1974a) used a two-equation model beginning with a production function (to estimate individual player production and introduced in the research topic) and then a team-revenue function. The production function consisted of explanatory variables in the form of performance statistics, such as slugging average and strike-to-walk ratio (discussed more in detail in the topic section of this research). The criterion variable was win percentage (Scully, 1974a).

The team-revenue function consisted of explanatory variables, such as win percentage and market characteristics of the ball clubs, such as the Standard Metropolitan Statistical Area (SMSA) and intensity of fan interest (denoted as MARGA). MARGA was estimated by calculating the relationship between team percent win and attendance. Scully found that most players in Major League Baseball were exploited (player salary was less than their MRP). The exception was mediocre players whose salaries were greater than the net MRP, which takes into account costs other than player labor of running a business (1974a). A key limitation of his work was that the production function omitted explanatory variables of interest, such as stadium investment and managerial quality of the staff in the dugout who make decision during the game such as changing pitchers.

Medoff (1976) built on the work of Scully (1974a) by changing variables to better capture a player's productivity, using more-recent data, and using two-stage least squares regression (instead of ordinary least squares) to address an endogeneity issue. He also incorporated race as an explanatory variable in the revenue function. To improve on capturing player productivity, Medoff (1976) used the runs scored metric for hitters because it incorporated various offensive contributions, such as singles, doubles, triples, and stolen bases, with their respective weights. Percentage wins (PCTWIN) was regressed onto the runs scored metric (mentioned above) for the entire team (TRUNS), earned run average of the team's pitching staff (TERA), and whether a team played in the NL or not (dummy variable) (Medoff, 1976). Medoff's (1976) revenue function consisted of many of the same variables as Scully's revenue function, such as percentage wins (PCTWIN) and SMSA. Key differences were the addition of new explanatory variables, such as percentage of non-White players on a team (BBPCT) and age of stadium (dummy variable denoted as STD). All of the explanatory variables in the revenue function were found to be significant at the significance level of 0.10, with the exception of BBPCT.

Two-Stage least squares regression was utilized for estimating the revenue function due to the possible correlation between the error term and PCTWIN in the revenue function, thought to be plausibly based on error terms of both equations being correlated (simultaneous relationship). Using ordinary least squares would have caused an upward bias of the coefficient PCTWIN. Ultimately, Medoff found, similar to Scully's findings, that players faced monopsonistic exploitation because they received approximately 11% of their economic worth (Medoff, 1976).

Continuing with the theme of monopsonistic exploitation by Scully (1974a) and Medoff (1976), Haupert (2009) obtained a unique data set of financial records for the New York Yankees for 1915-1937. This data set provided him with actual team revenue data that was previously unavailable and typically was proxied by variables such as attendance and television revenue. He utilized this unique data set from the pre-free agency era to better understand the exploitation of baseball players in the past compared to those who later earned the right to bargain for their wages in the 1970s via free agency. He found the degree of exploitation occurred more in situations where workers were less able to bargain for their compensation. Although prior to 1968, teams were not required to publicize their finances, the calculation of MRP for a baseball player followed Scully's three-equation approach. The Yankee's exploitation rates over the seasons followed a parabolic path when compared to player experience because the exploitation rate increased significantly between 1914 and 1917 but then decreased steadily through 1925 (Haupert, 2009).

Again, similar to Scully's work on MRP estimations, Sommers and Quinton (1982) incorporated expansion teams into the MRP analysis. Another key addition was the utilization of interaction between SMSA and win percentage of that particular team with SMSA. This was done to demonstrate how winning is more important for teams in larger markets. The researchers' two samples consisted of the first batch of free agents and nonfree agents. The researchers estimated the cost of the player by taking the total value of the contract and dividing its length in years. By following Scully's two-equation method, the researchers found that salaries for free agents approached their respective MRP values (Sommers & Quinton, 1982).

Blass (1992) shifted from the previous philosophy of measuring rates of exploitation and focused on the human capital model of investment. He sought to find out whether more-experienced players were paid more solely because of their production on the field (according to the human capital model of investment). To answer this, he utilized a new three-part approach to measure MRP of baseball players. The first part related offensive statistics to runs (runs scored function). The second part related runs to winning percentage (a production function), and the third part determined how win percentage impacted attendance and broadcasting revenue (two separate revenue equations). The third part used lagged variables, such as the previous season's team win percentage of the player (Blass, 1992).

The offensive statistics utilized in the first function were singles, doubles, triples, home runs, nonintentional walks, hit batsmen, stolen bases, grounding into double plays, caught stealing, sacrifice flies, and outs. For the second function, the author estimated win percentage in logit form (logarithm of the odds function), because it is difficult to transform win percentage into a useful form when utilizing a negative binomial distribution for runs scored. Blass then determined the necessary number of runs needed over the season to contribute to an additional win for the specific team. For the third part, he created two separate equations that incorporated how winning percentage uniquely impacted attendance revenue and local broadcasting revenue for individual MLB teams (Blass, 1992).

For his data set, Blass (1992) used a cross-sectional panel data set of players with at least 10 years of Major League Baseball experience to determine if experience and productivity impacted salaries. Regressing deviations in real salary from a nine-year mean onto variables such as experience and deviation from mean offensive productivity, he found that experience played a major role in determining the dependent variable while the introduction of the offensive variable of productivity had little impact on the dependent variable. He found that workers initially were underpaid but were overpaid later in their careers relative to their productivity (Blass, 1992).

ii.) Basketball

Some 11 years after the first research began on MRP in baseball, Scott, Long, and Somppi (1983) investigated the relationship between MRP and salary in the National Basketball Association. The authors found that a key advantage of investigating professional basketball players' MRP values versus that of Major League Baseball players was a lack of team investment in player development and training in the National Basketball Association because college basketball teams played that role. The researchers noted that prior to a 1976 federal court ruling, players in the National Basketball Association faced restraints in selling their services to the highest bidder. From 1976 to 1980, players gained more mobility to sell their services to other teams. After 1980, National Basketball Association players were granted nearly unrestricted mobility. Thus, one would expect that the level of restraint in the market would have a negative correlation in terms of players being paid less than their MRP. Also of interest to the authors was the possibility of racial discrimination. This could be investigated by evaluating whether Black players were paid less than their White counterparts of equivalent performance (Scott et al., 1983).

To address these hypotheses, the authors created two functions. The Cobb-Douglas production function used a team's winning percentage as its criterion variable. This was regressed onto the following performance statistics: team field-goal percentage, opponent field-goal percentage, team free-throw percentage, opponent free-throw percentage, team rebounds, opponent rebounds, team assists, opponent assists, team fouls, and opponent fouls. For a revenue function similar to Scully's (1974a) method, team revenue was regressed onto win percentage, population (MSA), arena capacity, Black percentage of city's population, per-capita income, years the team had been in the city, number of superstars on the team, playoff contender (dummy variable), and the percent of team players who were Black (Scott et al., 1983).

To estimate a player's defensive ability, variable points allowed was divided by 5 and then multiplied by the decimal value of minutes played in a game. This assumed that each player bore the same responsibility for points allowed based on their minutes played. The authors acknowledged that the team defense assumption understated the value of good defensive players and overstated the value of bad defensive players. It was found that player salaries approached the MRP as the freedom to negotiate their contracts increased. The researchers found no evidence of customer discrimination against Blacks as a significant positive coefficient was detected for the variable percentage of Black players on a team in the revenue function at the 0.01 significance level. The authors also found no evidence of racial discrimination at the management level because race was not a significant factor in determining a player's salary (salary regressed onto MRP and race) at the 0.01 significance level (Scott et al., 1983).

In 1999, Berri looked into measuring the production of National Basketball Association players in hopes of determining who was the most valuable player. A key improvement from previous literature was his ability to address statistics tied to an opponent's rebounds in a manner that would better evaluate individual player's production as opposed to Somppi's (1985) assumptions that each player was equally responsible defensively for all of an opponent's production. Note that Berri still adjusted for team statistics, such as opponent's 3-point field goal percentage, by distributing it to players on the team based directly on the percentage of minutes played divided by 5, which was the number of players on each team who were on the court at any one time (1999).

Berri (1999) also noted that the Cobb-Douglass production function he used failed to isolate production of individual players from their respective teams. For example, a team that had difficulties rebounding the ball also faced the issue of a diminished impact of scoring on winning. Thus, identical player performances on different teams would be of different value based on the performance of the player's teammates. Berri (1999) also weighted each player's production by the team's tempo to reduce the bias in which a team that played at a higher tempo would naturally score more than a team that played at a slower tempo ceteris paribus (Berri, 1999).

Berri (1999) utilized a three-stage least squares method. He began with a model that demonstrated how wins was a function of points scored and points allowed. Using data from the 1994-1995 and 1997-1998 seasons, he found that approximately 95% of the variation of team wins was explained in this model. Points scored and points allowed were then regressed onto explanatory variables, such as assist-to-turnover ratio, in a points-scored equation and a defensive assist-to-turnover ratio in points-allowed equation. Finally, the marginal value of each statistic was calculated by merging the two-equation system into a single-equation model that had win percentage as the response variable and individual variables that determine points scored and points allowed. Per-minute production was calculated by taking the summation of the (marginal value of the statistic)*(accumulation of the statistic) divided by the total minutes played of the individual player (Berri, 1999).

Building upon his previous work by adding performance statistics of assists and team rebounds, Berri (2008) sought to determine the production per 48 minutes played (P48) of all National Basketball Association players. Blocked shots and assists were determined not to fit in defensive or offensive efficiency due to the nature of both statistics, in which there would be a transfer of production from one player to another. With that in mind, two key adjustments were made involving the league average of blocked shots per team per 48 minutes being subtracted from a particular team's average blocked shots per 48 minutes. Likewise, the league average team assists per 48 minutes was subtracted from the team's average assists per 48 minutes. The two resulting numbers from these calculations were subtracted to calculate each player's 48-minute production. To account for the difference in positions that were deemed to be complements of each other in the National Basketball Association, each P48 performance was adjusted by subtracting the P48 league average of that player's specific position (e.g., center) (Berri, 2008).

iii.) Hockey

Some 5 years after the first research began on MRP in basketball, Jones and Walsh (1987) estimated the MRPs of National Hockey League player participants from the 1976-1978 seasons. The production function regressed percentage of maximum team points onto goals for (offensive performance), goals against (defensive performance), and team quality (proxying managerial and coaching effectiveness). The revenue function regressed total revenue onto percentage of maximum team points, stadium capacity, average household income and population of the team's city, Canada (dummy variable), and number of competitive league sports in the city. Combining both functions, using weights to account for expected defensive and offensive contribution of different positions in the sport, and adjusting for other costs such as travel (nonplayer cost), they calculated the MRPs. The authors found no evidence of exploitation of player's salaries and attributed this to the presence of the World Hockey Association, which was a competitor to the National Hockey League during this time period (Jones & Walsh, 1987).

Similarly, Bent and Sommers (2014) estimated the MRPs of National Hockey League player participants in the league's 2012 all-star game. They utilized a two-way recursive function. The first function regressed points onto scoring, goals allowed, and out, which is dummy variable that indicates whether a team is 20 or more points out of the eighth and final playoff spot. The explanatory variable of scoring was calculated as goals plus assists divided by the average of goals plus assists of every team in the league. Similarly, goals allowed was calculated by the number of goals allowed by the team divided by the average goals allowed per team in the league (Bent & Sommers, 2014).

The second function regressed revenue onto points, points*population, point*population^2, new (dummy variable indicating a new venue built within five years), Canada (dummy variable), and two or more teams in MSA (dummy variable). The authors evaluated the MRP of each player by first plugging the scoring into the points function to obtain the points estimate, which they in turn plugged into the revenue function along with the other explanatory variables in the revenue function, such as population (Bent & Sommers, 2014). Limitations of this study included how the effects of coaching and teammates were not accounted for in the production function.

Summary

The human capital model of investment and the concept of MRP play a role in my research. Sommers and Quinton's (1982) finding that free agent baseball salaries approach MRP supports labor economics theory that peoples' salaries approach MRP in an open, competitive market. Thus, one could make a case that research involving unrestricted free agent offensive linemen as the sample is giving an estimated MRP value of an NFL offensive lineman. Further, the variable experience, which is proxied by age in my models, will be used as a determining factor for wage determination of NFL offensive linemen and is representative of a trait found in the human capital model of investment.

b.) Race Discrimination (non-MRP, emphasis on wage discrimination)

While MRP estimates the true value of a player to a particular team, the topic of wage discrimination involves finding the factors associated with what the player is paid. Unlike studies on MRP, research on wage discrimination has been done in football. Many sports, such as soccer, baseball, basketball, and hockey, provided opportunities for researchers to analyze whether wage discrimination existed within the respective sport. Studies such as that done by Bellemore (2001), which looked at possible promotion discrimination, are included in this section because they provide similar methodologies and controls in terms of performance statistics that are present in wage discrimination literature.

i.) Football

Mogull (1973) was the first researcher to this author's knowledge to have investigated wage discrimination in football. At the time of his research, data of player salaries was not readily available, which means he resorted to sending surveys to all NFL players on the 1970-1971 rosters of teams. From the sample of 96 players who returned the questionnaires, subgroups consisting of rookie and nonrookies and whether the player played for a southern region team or non-southern region team were evaluated (Mogull, 1973).

For first-year Black players and White players, salary or bonuses were not significantly different (at a 0.20 significance level) between Whites or Blacks in either the South or non-South region. Similarly, White and Black players with one or more years of experience did not have significant differences between salary or bonuses. When experience was used as a proxy for playing ability, there appeared to be possible wage discrimination because Blacks had more playing experience in general but lower salaries than White players. However, the effect was not found significant (Mogull, 1973)

Mogull (1981) re-evaluated the same data from his previous work using an alternative methodology. His alternative methodology consisted of using the Chow test to study whether the coefficients of linear regression on two separate sets of data were equal. The F stat in the Chow test of 1.225 was not significant, indicating that football team management did not discriminate in wages based on race (Mogull, 1981).

Following Mogull (1981), Kahn (1992) utilized data on NFL players from the 1989 season to examine the issue of wage discrimination. The log of salary, where salary was defined as average annual compensation, was regressed onto performance variables, such as games played and games started, and onto race variables. The race variables were of main interest and were as follows: White (dichotomous variable), percentage of non-White residents in a metropolitan area of the team and the interaction term of both. The variable White was included to test for customer discrimination. Percentage non-White was assumed to be positive because non-White players should do better in home team areas that have higher percentage of non-White players. White players were assumed to earn more money in areas that have a higher percentage of White residents; thus, the sum of the coefficients for the interaction term and percentage non-White would be negative (Kahn, 1992).

Kahn ran separate regression on three different position categories, including the offensive line, and also ran a regression that included all positions. No variables of interest were found significant in the offensive line model. A key finding for all positions was that players made more money playing for teams that had a higher representation of their race, as demonstrated by three different models having significant coefficients that followed the signs of what the author expected (Kahn, 1992).

Gius and Johnson (2000) followed their study in 1998 about possible racial discrimination in the National Basketball Association by taking a similar methodological approach with the NFL. They utilized a log-linear wage equation and Chow test. Log of salary was regressed onto variables, such as race, experience, experience^2, and position variables, such as defensive tackle and offensive tackle. They found reverse discrimination in White players being paid 10% less (-0.10 coefficient) than Black players of similar ability with a significant Chow test (Gius & Johnson, 2000). In a similar study with a larger sample size and using several seasons of data, Doran and Doran (2004) found similar results, with the exception of there being a salary premium for White quarterbacks.

Transitioning to interest in racial discrimination for a particular position, in Berri and Simmons' study of "Race and the Evaluation of Signal Callers in the National Football League," the authors sought to determine whether Black quarterbacks faced performance-related discrimination (2009). Quantile regression was utilized in the salary model. Log salary was regressed onto experience, experience^2, draft round 1 (dummy variable), draft round 2 (dummy variable), veteran (dummy variable), change team (dummy variable), Pro Bowl participation, offense salary, log population of local MSA, career pass attempts, passing yards, Black (dummy variable), Black*(passing yards), touchdowns per attempt, completions per attempt, interceptions per attempt, rushing yards, and Black*(rushing yards). The singular performance measure significant across all percentiles was passing yards. The performance statistic of rushing yards had no significant impact on compensation throughout all quantiles, regardless of race (Berri & Simmons, 2009).

Thus, the skill of running as a quarterback was not compensated within the market. Black quarterbacks (who tend to run the ball well) faced performance-related discrimination (Berri & Simmons, 2009). A key limitation was the proxy used for evaluating the quarterback's teammates' abilities, a factor known as offense salary. It consisted of the summation of salaries for running backs, wide receivers, and tight ends.

Switching from the quarterback position to the linebacking core, Keefer (2013) sought to identify whether racial discrimination existed for this particular position. He built upon previous literature by focusing solely on a specific position on the defensive unit that had not been explored, to this author's knowledge. Quantile linear regression was utilized as a means to analyze where discrimination (if it did exist) occurred along the continuum of salaries for NFL linebackers. The explanatory variables of the model consisted of individual performance statistics such as tackles and sacks. Team statistics, such as points per game and yards allowed, were included in the model to control for the problem that poor defense correlated with more opportunities for making tackles. Both individual and team statistics were lagged 1 year because player contracts were to be impacted by the previous year's performance (Keefer, 2013). This author found that a non-Black player is compensated significantly more than a Black linebacker ceteris paribus at the 10^{th} and 25^{th} percentiles. Using an ordinary least squares model, the author found that non-Black players make on average 10% more than Black players ceteris paribus (Keefer, 2013).

Following Keefer's (2013) work, Burnett and Van Scyoc (2015) sought to determine whether wage discrimination involving rookie linebackers and rookie offensive linemen could be detected. These two positions were selected because they were relatively diverse with a number of Whites and Blacks playing at those respective positions. Using a methodology similar to that used by Keefer, while using only explanatory variables such as race and draft position, they found no significant evidence of racial discrimination (Burnett & Van Scyoc, 2015).

ii.) Baseball

In studies of baseball, much of the wage determination is found in the MRP section of this paper because MRP studies incorporate a wage determination component. However, topics such as customer discrimination based on fan attendance are of interest to the author and will be covered.

Prior to Scully's MRP analysis, he empirically analyzed wage discrimination between Blacks and Whites. He created racial salary functions of outfielders, infielders, and pitchers based on performance statistics, such as lifetime slugging average for the three outfielders position. The coefficients for performance used for infielders and outfielders were significantly different and favored White players at the infield and outfield positions, respectively. He argued that in order to evaluate wage discrimination, salary must be related to performance statistics (Scully, 1973).

Scully (1974b) also argued that fan attendance was impacted negatively by games pitched by Blacks. Using data from the 1967 baseball season, he regressed average fan attendance onto variables such as night, weekend, and race of the home team's starting pitcher. Race of the home team's starting pitcher was significant at the 0.05 level. On average, 1,969 fewer fans attended a game pitched by Blacks versus those pitched by Whites (Scully, 1974b).

Continuing with customer discrimination, Tainsky and Winfree (2010) investigated to see the change in effect of a foreign-born MLB player on ticket demand from 1985-2005.

Log of attendance was regressed onto variables, such as win percentage, lag-1 win percentage, number of foreign players, (number of foreign players*trend), (number of foreign players)*trend^2, and matching where trend was a component to measure the change over time. Matching was calculated by summing the product of each nationality proportion on an MLB team with the matching nationality in the SMSA of the team. They found that from 1985-1991, foreign-born players had a negative impact on ticket demand at a diminishing rate. In 1992, foreign-born players had a positive impact on ticket demand, and this positive effect increased until it peaked in 2000. From then until 2005, the positive effect slightly decreased. Matching had no significant impact (Tainsky & Winfree, 2010).

Shifting to promotion opportunities, Bellemore (2001) investigated whether discrimination in promotion from the minor leagues (AAA league) to the major league in baseball occurred during distinct time periods. Eleven years of data on batters from the late 1960s to the late 1990s were used. A probit equation was regressed onto statistics such as batting average, home runs, runs batted in, stolen bases and race. In the 1960s and 1970s, Blacks and Hispanics faced promotion discrimination. With the expansion of Major League Baseball in the mid-1990s, promotion discrimination subsided for Blacks and was not significant for Hispanics (Bellemore, 2001).

Using awards as a response variable, Depken and Ford (2006) sought to evaluate customer-based discrimination in the form of baseball's All-Star voting process from 1990-2000. The number of All-Star votes was regressed onto many variable types, such as race, player quality, and voters. Within the variable type of voters, dummy variables accounted for where a player's team was located. The researchers found no significant effect in terms of (‘in terms of’…) discrimination against minorities. Fans actually displayed a bias toward Blacks and Latinos in All-Star voting (Depken & Ford, 2006).

Jewell, Brown, and Miles (2002) also measured whether racial discrimination impacted one’s chances of being inducted into the National Baseball Hall of Fame. Race variables such as being Black; being born in Latin America and the interaction of those two; performance statistics and noteworthy accomplishments, such as World Series Championship appearances, were considered as explanatory variables. Percentage of votes received out of total votes was the dependent variable. The dependent variable was transformed by taking the log of percentage of votes received over percentage of votes not received because a logit chi-squared method was employed. Heckman’s correction was employed on the sample to account for no-vote players.

iii.) Soccer

Gandelman (2009) utilized data from a Uruguayan soccer league to address whether race plays a role in media player evaluation. The dependent variable, denoted performance, consisted of evaluations by journalists (average grades) and was regressed onto variables such as goals, race, and education. Race was a significant factor at the 0.10 significance level in four of the six models. He concluded that discrimination by Uruguayan media was present against non-White players in the national league (Gandelman, 2009).

In an Italian soccer league, Bryson, Rossi, and Simmons (2014) investigated player contributions to team wins and fan attendance and their relationship to team revenue. They built upon previous literature by including other performance statistics beyond goals and assists in their player productivity variable. For the salary model, log salary was regressed onto age, age^2, experience, player productivity, team effects such as team Serie A or B, season dummy, and nationality dummy variables (Bryson et al., 2014).

Player productivity was measured by the following performance statistics: career goals in Serie A, career goals in Serie B, appearances in Serie A, appearances in Serie B, minutes played, minutes $played^2$, lost balls, recovered balls, season goals in Serie A, goalkeeper saves, goal assist, shots on target, successful passes, tackles, fast breaks, footballer of the year award, World Cup selection, European Championship selection. The nationality dummy variables consisted of a dummy variable known as "local" to indicate whether the player plays close to his hometown and "Italian" to ultimately distinguish between local and nonlocal Italians. Using ordinary least squares and quantile regression, they found a growing wage penalty (starting at the 50th percentile) for Italian players versus non-Italian players as one earned higher wages (Bryson et al., 2014).

In order to explore the individual effects of change on team effects, which would impact attendance, the attendance model was created. Log team attendance (dependent variable) was regressed onto predicted salary (proxy for team quality), residual salary, team points, year, nationality, and team fixed effects. Using models such as the above, the authors determined that wage discrimination occurred in large part because fans preferred a player being a European Union (EU) migrant (Bryson et al., 2014).

iv.) Basketball

In the National Basketball Association, Kahn and Sherer (1988) sought to identify whether racial discrimination was present in the employment of its players. Their work built on previous literature on wage discrimination because they were better able to control for outside factors by utilizing data from means such as the Current Population Study. Salary was regressed on variables such as average minutes played, the selection number in which the player was chosen in the National Basketball Association draft, team win percentage, and race. The draft number of the player was used as a proxy for fan's perceived quality of a player. Two-Stage least squares method was employed to address a possible endogeneity issue of the variable draft number interacting with player quality and race. The draft number was regressed onto variables such as race and college field goal percentage. A key limitation is that the authors used only one year of data. They found racial discrimination exists and customer discrimination appears to be the driving factor (Kahn & Sherer, 1988).

Jenkins (1996) made a significant contribution to the wage discrimination literature in basketball by using a new sample selection. Only free-agent National Basketball Association players were selected for this study because Jenkins believed doing so gave a better fit between past performance and salary. He found no significant effect for race on wages in his model and no supporting evidence of race discrimination utilizing a Chow test (Jenkins, 1996).

Similarly, Hamilton (1997) utilized ordinary least squares and tobit (is 'tobit' correct?) regressions and found no difference between salaries of White players and Black players when controlling for player performance, such as seasons played, and team characteristics, such as a general manager being Black. However, censored quantile regression indicated that Blacks were discriminated against in the upper end of the salary distribution (90^{th} percentile) (Hamilton, 1997).

Continuing with possible wage discrimination using more-recent data, Gius and Johnson (1998) used a log-linear wage equation and also performed a Chow test. Log salary was regressed onto race and performance variables, such as assists per game and steals per game. Proxies for whether a player was a free agent and for whether the player stayed on the same team were also included as explanatory variables. Surprisingly (based on the theory that a player's salary approaches MRP in a free, open, competitive market), free agency had a negative effect on player salary. Gius and Johnson (1998) conjectured that this result occurred because of the salary cap rules in place in the National Basketball Association limited a team's expenditure on free agents. The Chow test was utilized to determine if the regression parameters differed for White athletes and Black athletes in wage determination. Both the model and Chow test found that wage discrimination was no longer prevalent in the National Basketball Association (Gius & Johnson, 1998).

Robst, VanGilder, Coates, and Berri (2011) looked into intraracial discrimination by introducing an objective measure of skin tone called the RGB score (used by Adobe Photoshop) in basketball. Using unrestricted free agent players in the National Basketball Association for a sample, a log of salary was regressed onto performance statistics, such as a player's points per game, racial composition of the city in which the team was based, and skin tone, among other variables. The researchers found minor support for customer discrimination in that lighter skinned Black players earned higher wages for teams that had a higher representation of Hispanic and Asians but that no significant evidence of employer discrimination was based on skin tone.

Going beyond just wage discrimination, Kahn and Shah (2005) investigated whether race impacted salary, total compensation, and contract duration for National Basketball Association players during the 2001-2002 season. Salary was defined as average annual compensation (including bonuses) over the duration of the contract. Compensation was defined as the total cash value of the entire contract. A key contribution to previous work performed on wage discrimination in the National Basketball Association was that they accounted for differing negotiation rights of players being first round picks, nonfirst round picks, and free agents. Using the negotiation rights, Kahn and Shah (2005) separated their players according to those three groups. The log of compensation and log of salary were regressed onto variables, such as rebounds, blocks, market variables, and race using tobit analysis because the league has a minimum wage of $332,817. Duration was analyzed as a dependent variable of the model using ordered probit because it is a discrete variable. The authors found significant discrimination for nonfirst round draft selections on total compensation, salary, and contract length by looking at the marginal effects of being White (Kahn & Shah, 2005).

Schollaert and Smith (1987) turned their attention to customer discrimination and sought to determine whether race played a role in attendance at National Basketball Association games. They created four models with different dependent variables to measure attendance as follows: total attendance, percentage of seats sold, number of unsold seats, and per-capita attendance. The independent variables were facility characteristics and market characteristics. Team racial composition had no significant effect on any of the attendance models (Schollaert & Smith, 1987).

Brown, Spiro, and Keenan (1991) sought further to determine if fan attendance impacted wage discrimination through their preference of watching White players on the court. Additionally, the researchers sought to determine whether a player's ability to be employed by a team is determined in part because of their race (entrance discrimination) existed. Prior to addressing the key points of interest, a salary determination equation was created with log of salary being regressed onto variables such as race and points per minute played over a lifetime. There was a significant negative coefficient on race of -0.156, indicating that Black players' salaries were 16% less than White players' salaries ceteris paribus.

To address possible customer discrimination as mentioned at the beginning of the above paragraph, attendance was regressed onto PCTMINB, which indicated the percentage of minutes played by Black players during the season and other variables, such as the number of years a team has played in a particular city. The variable PCTMINB was not significant, which indicated that no evidence was found that fan attendance was impacted by the playing time of Black players (Brown et al., 1991).

As for finding possible entrance discrimination, the authors sought to determine whether marginal White players were outperformed by marginally performing Black players, based on performance statistics such as points per game and assists per game. Players were grouped as "marginal" if they averaged 25 or fewer minutes per game. Racial performance differentials were investigated on performance statistics, such as field goal percentage. The researchers found weak evidence of entrance discrimination because only limited-performance variables (points per game and points per minute) were significant for marginal players (Brown et al., 1991).

From entrance discrimination to exit discrimination, Groothuis and Hill (2004) sought to determine if National Basketball Association players' careers were shortened because of their race. They used both nonparametric and semiparametric methods. The nonparametric method consisted of calculating hazard rates for Black and White players. Hazard rate calculated the percentage of players who left the National Basketball Association after a particular career length. Hazard rates for Black and White players were similar. A key limitation of this method was that it did not account for the productivity of the players. The semiparametric method consisted of a logit model, which consisted of a dichotomous dependent variable in which a player would receive a 0 for each year he did not exit the league and a 1 when he did exit the league (panel data set). The researchers' models consisted of variables such as steals, blocks, height, weight, tenure, tenure^2, tenure cubed, and tenure fourth power with the likelihood function statistic.

Kanazawa and Funk (2001) used Nielsen ratings as a dependent variable to determine whether customer discrimination existed in the National Basketball Association. Nielsen ratings were regressed onto variables such as home team win percentage, White fan percentage of SMSA, the number of players on the home team, the number of White players on the visiting team, and weekend as a dummy variable. Ultimately, six models were created, two of which were generalized least squares models, two of which were fixed-effects models, and two of which were random-effects models. The study found that viewer ratings increased based on greater participation of White players, even if they are on the bench (Kanazawa & Funk, 2001).

To better understand customer discrimination in the National Basketball Association, Burdekin, Hossfeld, and Smith (2005) measured the racial composition of an National Basketball Association team in three ways: percentage of members on the team who were White, percentage of starting players on a team who were White, and percentage of bench players on a team who were White. The three measures of racial composition were used as dependent variables for ordinary least squares regression and were regressed onto variables, such as the percentage of White residents in the SMSA population (denoted POPWHITE). It was found that the percentage of White starters on a team was significantly impacted by

POPWHITE, indicating that teams cater to customer discrimination. (be sure that 'cater to' fairly describes this)

To determine whether White players generated additional revenue in more White populated SMSA areas, home-game revenue was regressed onto important variables, such as the percentage of White players on a team (denoted TWHITE), the team's winning percentage, and the interaction term of TWHITE*POPWHITE. In the second model, the interaction term (what is 'interaction term'?) was found to be positive and significant, indicating White fans do prefer ('prefer' to do what?) White players (Burdekin et al., 2005).

Looking at performance-based rewards, Berri, Brook, and Schmidt (2007) turned their attention specifically to the variable of points scored and how points scored related to voting points for the National Basketball Association's all-rookie team. Further, wages for free agent basketball players, following wage determination, were calculated. For the all-rookie team, the natural log of voting points was regressed onto production variables, such as wins produced and points scored, and on other variables, such as draft position of the player. A tobit model was used to constrain the dependent variable of voting points earned by the rookie player. Looking at various tobit models, the researchers found points scored to be significant and to have the most significant impact on voting points (Berri et. al., 2007).

For the wage model, Berri et al. (2007) considered only free agents. This was done to focus on production variables at the time in which the contract would be determined. The dependent variable was the log of the average real salary over the duration of the contract. It was regressed onto variables such as points scored per game, rebounds per game, and wins produced per game. Similar to results found in studying different models on voting points, the various salary models demonstrated that average points scored per game had the largest significant impact on a player's salary (Berri et al., 2007).

v.) Hockey

In hockey, Jones and Walsh (1988) sought to determine if discrimination existed with French Canadians and their respective playing positions on the ice. Their work built upon previous research of racial discrimination by looking into the sport of hockey. The log of salary was regressed onto variables, such as height, weight, and the dummy variable of being born in Quebec. Skills such as average total goals and assists per game (denoted as points per game), which were of importance for forwards and defensemen, were found to be a major determinant of a player's salary. When looking at all positions, the researchers found no significant evidence of discrimination against French Canadian players in the National Hockey League. A key limitation of the study was that no variable was present to purely address a player's defensive capability (Jones & Walsh, 1988).

Similarly, Jones, Nadeau, and Walsh (1999) sought to determine if minority groups, such as Europeans or Americans, faced wage discrimination. The log of salary was regressed onto several variables, such as goals per game, assists per game, and ethnicity (dichotomous variables to account for minorities). Of key interest, the researchers found that salary was determined primarily by variables that measured productivity and skill of the player, while ethnicity played little to no role on wage determination (Jones et al., 1999).

Summary

Depending on the sport and the study, there are a variety of models and different results to determining whether racial discrimination in sport exists. Of key interest is evaluating the relationship between salaries and race, controlling for performance variables.

This is a common method utilized to determine whether wage discrimination occurs in all types of sport. Wage, race, age, age^2, position, and games started are variables used in previous research in this section that will be evaluated in my models.

c.) Performance Statistics

In the third subcategory of individual production, performance statistics in sports were analyzed. Key to the evaluation of performance statistics was determining how well a statistic correlated with the production of an individual (e.g., hitting doubles to winning in baseball). Comprehensive statistics are also of interest because these statistics can be used to objectively evaluate all players across different positions in a particular sport. Much of the early research occurred in baseball. Research of performance statistics in football became prevalent in the past decade.

i.) Baseball

Lindsey (1959) evaluated the usage of statistics for decision-making processes and player performance in baseball. Of particular interest, he investigated the shortcoming of batting average as a measure of batting effectiveness. Batting average failed to adequately account for productive hitting and circumstances such as a batter hitting into a double play. Slugging percentage accounted for this first issue, making it a better measurement of a batter's hitting ability (Lindsey, 1959). For runs batted in (RBI), Lindsey noted how that performance statistic was dependent on the performance of other players as well as on the individual who batted in a teammate to score (Lindsey, 1963). Using probabilities of situations and the expected runs in each particular situation, Lindsey created a linear-weights model to estimate expected values of a single, double, triple, and home run hit (Lindsey, 1963).

Cover and Keilers (1977) created a statistic known as offensive earned-run average (OERA). This statistic gives an estimation of the number of runs per game a team would score if its entire hitting lineup were composed of players identical to that single player. Five assumptions were made, such as runners not advancing on outs, singles advancing a player two bases, and a double advancing a base runner three bases. The assumptions were done so that calculations were independent of scorer judgment. While play-by-play computation of the OERA can be performed, the authors created a formula that they expected would give an appropriate approximation. The formula utilized the individual performance statistics of at-bats, singles, doubles, triples, home runs, and walks. After calculating the probability of each performance statistic and denoting the 24 different states in an inning (expected run matrix), a Markovian recurrence was established to calculate the expected runs per inning. Multiplying that by nine innings calculated the OERA (Cover & Keilers, 1977).

More currently, comprehensive statistics such as wins above replacement (WAR) have been employed as tools to measure overall player performance. These performance statistics are valuable, especially if they are highly correlated with wins. Issues with WAR include a lack of understanding of the uncertainty of the estimate and reproducibility, because there is no standard calculation of the statistic (Baumer, Jensen, & Matthews, 2015).

From sabermetrics to labor economics, the “Moneyball Hypothesis” stemming from Michael Lewis’ book *Moneyball* was based on the existence of a market inefficiency of how the skill of batting is priced in baseball. In particular, the inefficiency was an undervaluing of the performance of on-base percentage (OBP) versus an overvaluing of performance statistics such as making contact on a swing. Billy Beane’s exploitation of this market inefficiency as general manager of the Oakland Athletics helped his team compete with the likes of the New York Yankees, whose player payroll was almost three times that of the Athletics (Hakes & Sauer, 2006).

Hakes and Sauer (2006) sought to test to whether the Moneyball Hypothesis was indeed true. Win percentage was regressed onto the OBP of the team and the opponent's OBP. Some 82.5% of the variation of win percentage was explained by these performance statistics. Similarly, win percentage was regressed onto team slugging percentage and opponent's slugging percentage. Some 78.7% of the variation of win percentage was explained in this model. Comparing the respective coefficients of both models, the authors noticed the coefficients in the OBP model were significantly larger than the coefficients in the slugging percentage model (Hakes & Sauer, 2006).

Next, the authors sought to evaluate the market pricing of baseball players. The log of salary was regressed onto performance statistics such as slugging percentage and OBP. Dummy variables such as arbitration eligible and free agency were also incorporated because they impact the monopsonistic exploitation of a player. The coefficient of slugging percentage was found to be higher in the years of 2000-2003. Thus, there is an overvaluation of slugging percentage versus OBP, as indicated by Michael Lewis in his book. In 2004, the slugging percentage coefficient was larger than the OBP, indicating there was a market correction (Hakes & Sauer, 2006).

Following the Moneyball Hypothesis, Gerrard (2007) evaluated the applicability of this scenario in other sports. Prior to doing this, he used another method to evaluate whether the Moneyball Hypothesis was indeed valid. The win-ratio was regressed onto current payroll cost (Equation 1). Approximately 25% of the variation of the win-ratio was explained in that model. In another model, the win-ratio was regressed onto current payroll cost, lag 1 win-ratio, and a binary variable to account for teams that were new to the MLB (Equation 2). This model explained approximately 37% of the variation of win-ratio (Gerrard, 2007).

With these models, one can determine a benchmark for how many games a team should win, denoted as the benchmark, based on the team's payroll. A statistic, Wins Above Benchmark (WAB), was formulated. A negative WAB indicated that a team performed below its benchmark. Total WAB was broken down into Innovation Effects and Persistence Effects. Innovation effect was captured by the team's residual in equation 2. Persistence Effects, which included variables such as accrued experience captured the impact the previous season had on WAB. It was a difference of the residual in Equation 2 less the residual of Equation 1 (total WAB minus Innovation Effect). Over the period of 1998-2006, the Oakland Athletics led MLB in total WAB. That statistic lent support to the belief that Billy Beane's staff was innovative and held a competitive advantage (Gerrard, 2007).

ii.) Soccer

Focusing their research on soccer, Weimar and Wicker (2015) investigated the possibility of a market inefficiency in the Bundesliga. The researchers utilized probit models and found that effort measures of total distance run and number of intensive runs per player per match had a significant effect on the chances of winning in each separate model. The models consisted of other team performance statistics, such as the number of yellow cards and tackles won. In the full models (incorporating both effort measures), total distance run was significant and positive, and the number of intensive runs was significant and negative. The authors speculated that this was due to the high correlation between both measures (r^2 = 0.86). Based on the previous findings of Wicker, Prinz, Weimar, Deutscher, and Upmann (2013), in which effort did not have a significant impact on market value of a Bundesliga player, the authors speculated that the labor market undervalued the distance that a player ran during a soccer game (Weimar & Wicker, 2015).

Transitioning from the performance variable of effort to what is known as the footedness of soccer players, Bryson, Frick, and Simmons (2013) regressed the log of market valuation (a proxy for player salary) onto variables such as age, height, left foot, and two foot. The variable of interest, two foot, was a dichotomous variable indicating that the soccer player had equal power in both legs. To address a possible upward bias of this coefficient, the top and bottom 5% of market valuations were removed from the data set. Two foot was found significant at the 0.01 significance level with a coefficient of 0.171, indicating that players were paid a premium of 17.1% increase of their salary for this ability. Creating another model, total points was regressed onto relative average payroll, (relative average payroll)^2, and relative average share of two-footed appearances. Relative average share of two-footed appearances was not significant, leading to the conjecture of a market inefficiency (Bryson et al., 2013).

Also looking at performance statistics and the relation to winning, Gerrard (2007) created a hierarchal model that consisted of team statistics and individual statistics. General play was broken down into the components of offense and defense. Offense consisted of score opportunities, conversion rate, and goals scored. Similarly, defense consisted of opponent's scoring opportunities, opponent's conversion rate, and goals allowed. Goals scored and goals allowed then contributed to match result (Gerrard, 2007).

Using benchmark analysis, six variables were measured as deviations from the benchmarks. The variables were (I will send via a separate memo rules for the use of 'include') general play, striking, goalkeeper, offense, defense, and result. General play consisted of performance statistics, such as the number of passes and the rate of completed passes. The first three variables were determined for individual play, while the other three variables were determined or team play. Breaking these down for team over the years gives one the ability to determine whether coaching (fitting into team effects) is of the utmost importance to a team's success. Gerrard found that player effectiveness was highly correlated to the player's wage, while team effectiveness was not. Thus, teams with limited resources can overcome this disadvantage with a knowledge-based strategy (Gerrard, 2007).

iii.) Football

Following Gerrard (2007) in relating performance statistics to winning, but transitioning to football, Berri (2007) found that 84% of wins in an NFL game could be explained by points scored and points surrendered. Points scored and points surrendered then could be regressed upon various independent variables, such as total team rushing yards gained, passing yards gained, and opponent's yards gained (Berri, 2007). Furthermore, the researcher then could attribute what a player produced by specific performance statistics, such as touchdowns thrown.

Utilizing performance statistics, such as total rushing yards of an individual player, Simmons and Berri (2009) investigated wage determinants based on specialization versus versatility of workers in the NFL. Their subjects were a team's running backs (RBs). The two components of total receiving and rushing yards were considered the focal points of performance measures for RBs. Similar to Berri's (2007) study, for determining which components contributed to offensive performance in the NFL (by their respective weights), the number of points a team scored was regressed onto several performance statistics, such as opponent's interceptions, opponent's fumbles lost, plays, passing yards, rushing yards, interceptions, and fumbles lost. For running backs, the authors made the case that the returns for receiving yards were more valuable than returns for rushing yards. By finding the average number of plays to reach 100 yards receiving and 100 yards rushing, respectively, and applying the costs associated with the number of plays it took reach those respective values, it was estimated that 100 rushing yards produced 3.13 points and 100 yards receiving generated 5.19 points (Simmons & Berri, 2009).

Next, the authors developed the wage determination model. Salary was measured by pro-rating a player's bonus over the life of the contract. Due to non-normality and excessive kurtosis for player salaries, the authors utilized a quantile regression method for estimation of the log salary. The explanatory variables were the player's experience, draft round 1 and draft round 2 (dummy variables), restricted free agent (dummy variable), veteran (dummy variable), stayer (dummy variable), change team (dummy variable), number of Pro Bowls in which a player was selected to, log salary of the offensive line (proxy for quality of offensive line), log of SMSA population (proxy for market size), fullback (dummy variable), career rushing yards, career passing receiving yards, rushing yards, passing yards, and (rushing yards)*(passing yards), which is a key variable to examine to determine whether specialization pays off. Specialization also paid running backs more than versatility, especially at the higher percentiles (Simmons & Berri, 2009). A key limitation of this study was that blocking ability of RBs was not considered.

Moving from performance statistics of a running back to a quarterback, passer rating was one of the most widely used statistics in football. The statistic acted as a measure of quarterback performance (von Dohlen, 2011). Stimel (2009) investigated the possibility of structural breaks in the four variables that compose passer rating—average yards per attempt, completion percentage, interception percentage, and touchdown percentage—in attempts to better understand a possible inflation of the statistic. He found structural breaks within the time frame of 1960-2007 and argued that rule changes involving passing plays and the formulation of the salary cap might have caused the overall change in passer rating for particular eras. Thus, comparisons between quarterbacks of different eras should not be done unless one accounts for these changes. Utilizing graph theory, he found evidence that the variables average yards per attempt and interception percentage affect completion percentage. Further studies are needed to support these interactions (Stimel, 2009).

Also focusing on passer rating, von Dohlen (2011) sought to evaluate and improve the effectiveness of the statistic. One of the criticisms of passer rating is that it increased steadily on average from 1940 to 2010, as cited by Stimel (2009). Dohlen (2011) proposed two modifications. The first involved utilizing the yards per completion instead of the yards per attempt within the model, because that would give a better balance between weighting production and throwing accuracy. The other modification adjusted the linear scale by taking the mean performance of the 10-year duration prior to the passer rating of that particular year, which would adjust for the inflation of the statistic (von Dohlen, 2011).

In a study of kicking metrics, Pasteur and Cunningham-Rhoads (2014) created a new expectation-based metric to evaluate place kickers. Instead of classic field goal percentage, their new metric accounted for important variables such as distance of the kick and weather conditions. Multivariate logistic regression was used to create an expected model of how a kicker would be expected to perform under these conditions. One could then compare all kickers' actual success versus what the model predicted their success would have been at the league average. Player salaries were weakly correlated ($R^2 = 0.12$) using the metric and Pro

Bowl, and all pro kickers were found to be mediocre, according to the metric. Some limitations of the study include not accounting for unusual wind patterns or psychological pressure on the kicker. Also, sample size could be enlarged by extending the sample to a decade worth of data (Pasteur & Cunningham-Rhoads, 2014).

In a study of offensive linemen, Alamar and Weinstein-Gould (2008) were the first known researchers to investigate the relationship between successful pass blocking and the completion rate of a quarterback. In their methodology, they utilized two logistic regression equations. The first related "time in pocket" (TIP) to an individual offensive lineman and the probability of successfully keeping the defensive lineman out of the pocket as the dependent variable. The other equation estimated the impact of a failure by an individual offensive lineman on the probability of the team completing the pass. The equation controlled for variables such as down, distance of the throw, and yards necessary to earn a first down. Plays that had penalties and designated quarterback runs were not included in the sample. Plays that resulted in a quarterback scramble were included because TIP was reported. A key limitation acknowledged by the authors is that they did not account for variation of the team defenses that offensive linemen played against (Alamar & Weinstein-Gould, 2008).

iv.) Hockey

As mentioned in the author's introduction, Idson and Kahane (2000) sought to determine if co-worker productivity impacted the salary of individual players. The log of salary was regressed onto individual player variables, coach variables, team variables, and interactions between team and individual variables, such as (team points)*(individual points). In terms of interactions, one conclusion was that a player with higher ability was rewarded more on teams that were more successful based on (team score)*(individual points) being positive and significant at the 0.10 level (Idson & Kahane, 2000).

Summary

The interaction effects I captured in my models were most similar to those presented in Idson and Kahane (2000), in which the authors looked into the interaction between team productivity and individual productivity. Simmons and Berri (2009) demonstrated the application of an interaction term with individual performance measures in a wage determination model in football. The development of performance statistics in research also enables us to further investigate possible market inefficiencies that are reflected directly in the research of Hakes and Sauer (2006); Gerrard (2007); Bryson, Frick, and Simmons (2013); and Weimar and Wicker (2015).

d.) NFL Combine Metrics

The literature on combine metrics at the NFL level has found mixed evidence of the relationship between physical ability and player performance, or the round of the annual draft in which the player is selected (Hartman (2011), Kuzmits & Adams (2008), Robbins (2010), McGee & Burkett (2003), Dhar (2011), Berri & Simmons (2011), and Wolfson, Addona, and Schmicker (2011). However, combine metrics do play a role as to player draft status (Sierer et al., 2008; McGee & Burkett, 2003) Thus, it is likely that combine skills are prerequisites to compete at elite levels (Robbins, 2010).

Prior to looking into previously cited studies, I will briefly explain the NFL Combine and discuss the medical tests and physical metrics used there. Each year, the NFL holds an event to evaluate the top prospects coming out of college, which is known as the NFL

Combine. In the combine, physical, psychological, and medical tests are performed on select players who accept the invitation (Brophy et al., 2008).

Medical tests consist of team physicians from each organization evaluating all major joints of the player, inquiring about missed games in college, and learning about the players' surgical history. In addition, a 10-year history of X-rays, CAT scans, and MRIs are reviewed. Using this information on each player, medical teams evaluate each player's career track (Futterman, 2017). Using medical evaluations of an anonymous NFL team from the years 1987-2000, Brophy et al. (2008) found that orthopedic grades (high, low, failure) correlate with the number of games played.

The most common physical tests that that players perform are a 40-yard sprint, vertical and broad jumps, 20-yard shuttle run, three-cone drill, and 225-pound bench press (Robbins, 2011) to measure acceleration and maximum speed, jumping and change of direction ability, and upper-body strength (Robbins, 2011). Coupling these physical tests with past performance and focusing on running backs, Hartman (2011) evaluated a possible relationship between players' physical ability and past production and their ranking in the annual draft. The only significant relationship found at the 0.05 level was between total yards during the player's final year of college and draft status (R = -0.66). Robbins (2011) used the same method of correlation but expanded to include all positions in his study. Combine data was examined in its original form and was scaled. Looking strictly at combine physical tests, Robbins (2011) found no significant results that would link performance at the combine to predicting future performance.

By examining only quarterbacks, running backs, and wide receivers, but looking at draft order and games played (measurement of performance), Kuzmits and Adams (2008) found that players' vertical jump and broad jump had significant relationships in terms of improving draft status for quarterbacks at the 0.05 significance level. For wide receiver draft status, 10-yard and 20-yard times were significant at the 0.05 significance level. For running back draft status, broad jump was the only significant correlation. However, at all position levels, the researchers found no consistent relationship between combine tests and future performance, with the exception of sprint tests for running backs (Kuzmits & Adams, 2008).

Focusing on quarterbacks and their NFL success as measured by games played and net points, Wolfson et al. (2011) found that college and combine statistics had little value on predicting NFL performance in their models. Their research differed from previous work because they based their analyses on all quarterbacks drafted by NFL teams (since 1997), not only those who have played in at least one NFL game. With similar findings, Berri and Simmons (2011) expanded the usage of performance measurements by using criteria such as net points, QB rating, and passing yards per attempt. The college performance metric of completion percentage was found to have had a positive impact on future performance as measured by NFL completion percentage at each level of experience. For a draft position model, height, speed, and whether a player played at the D-1A level (dummy variable) were significant. However, combine metrics and where the player was picked were found to have no significant effect on future performance (Berri & Simmons, 2011).

Similar to research by Berri and Simmons, Dhar (2011) also investigated draft order and performance models, but his focus was on the wide receiver position. To measure performance, he used WR score, which encompasses total NFL receiving yards, total NFL receiving touchdowns and years in the league. Using both ordinary least squares and a recursive partitioning tree (CART) for the draft order model, he found that the 40-yard dash time was an important factor in draft position. However, college performance was more of an overall determining factor in terms of where a player is drafted. For future performance, total college receiving touchdowns and final college receiving yard percentage were important factors, while combine metrics were of little importance (Dhar, 2011).

Similarly, Mulholland and Jensen (2014) used both CART and ordinary least squares to create draft and performance models. They focused on the tight end position. Both draft order models using CART and linear regression, respectively, gave evidence that height, body mass index, 40-yard dash, bench press, and career college yards impact where a player is drafted. Performance models (using NFL games started, NFL career score, and NFL career score per game as different measures in respective models) using linear regression and CART weight consistently show the importance of the broad jump as a factor in a player's future performance. However, little evidence demonstrates that the variables that significantly impact where a player is drafted can predict performance (Mulholland & Jensen, 2014).

While most combine measurements offer little in terms of predicting player performance, support for vertical jump being worthwhile in the combine can be found in research at the college level. To measure football ability, Barker et al. (1993) had college coaches of their respective team evaluate players as starters or nonstarters (dichotomous variable), while Sawyer, Ostarello, Suess, and Dempsey (2002), and Daniel, Brown, and Gorman (1984) had player ability measured by coaches' ranking. Sawyer et al.'s (2002) study ranked all of the players on offense and defense separately, while Daniel et al. (1984) ranked player at the position level. A player ranking of 1 was considered the highest playing ability for the respective category; 2 was the second highest in playing ability; and so forth until every player was ranked. Using physical tests performed at the respective football programs, Barker et al. (1993) and Sawyer et al. (2002) found that football ability correlated with the test of vertical jump. Daniel et al. (1984) found that vertical jump was significant in terms of athletic ability of offensive linemen and defensive backs.

While the literature shows that combine measures were not sufficient predictors for player success, there is research that supports the idea that they act as signals of player ability (Robbins, 2010). Sierer et al. (2008) found that the probability of a player being drafted by a team was positively affected by his combine results. Furthermore, McGee and Burkett (2003) found that the combine measures not only determined whether a player would be drafted but also the round in which it occurred. McGee and Burkett (2003) successfully created many prediction models based on player position. The seven models were categorized for the positions of quarterback, wide receiver, running back, offensive line, defensive line, defensive backs, and linebackers with combine metrics and physical characteristics as explanatory variables. There was greater explanatory power for the wide receiver, running back, and defensive backs model. However, the linebacker model had only an R^2-value of 0.223 (McGee & Burkett, 2003).

Summary

Based on the lack of significant results in trying to find a relationship between many physical tests and future performance, there is skepticism about the value of the NFL Combine. Additional suspicion comes from public outsiders based on exceptional players such as Tom Brady ranking near the bottom of vertical jump for quarterbacks or wide receiver star Odell Beckham Jr. ranking near the bottom on the bench press test for wide receivers (Salfino, 2015). In the NFL, executives share similar concerns and express the need for other measures of player ability and performance. My research will look at NFL performance measures to better understand what determines player salary, which is an estimate of future player performance.

Methodology

This chapter describes the methodology used to conduct this study. A quantitative approach using a single-equation model was linked to relate individual statistics of offensive linemen, such as previous regular season proportion of knockdowns on quarterbacks by individual offensive linemen; individual characteristics of offensive linemen; such as race, team statistics such as total points scored, and nonoffensive line position statistics such as adjusted line yards, toward the determination of a National Football League (NFL) offensive lineman's salary. Ultimately, one is better able to identify the factors that determine the wages of any NFL offensive lineman in the open market. This chapter is divided into four sections: research design and procedure, population and sample, data source, and data analysis.

Research Design and Procedure

The model presented in the study provides a foundation of what determines an offensive lineman's salary and if individual offensive lineman performance statistics, such as regular season games started, impact his wages. Individual performance statistics might be nonsignificant in the wage determination model, but team performance statistics, such as total points scored, might be significant. This could indicate that current individual performance statistics of offensive linemen do not adequately measure what a player at this position contributes to winning. Thus, general managers would look at team performance statistics as better and easier to understand estimates of the true value of an offensive lineman. In this scenario, where there is a lack of current individual metrics, team managers lack complete information.

While evaluating individual offensive lineman performance of run blocking would be optimal using individual performance statistics, individual data related to run blocking is not available. Thus, statistics at the unit level of the entire offensive line are used, e.g., running backs' yards before contact (YBC), which are mentioned below in the model.

The base model most closely resembles the model in wage determination in research by Simmons and Berri (2009) but now focuses solely on unrestricted offensive linemen instead of on running backs. The interaction term of (proportion of knockdowns on quarterbacks by an individual offensive lineman)*(proportion of knockdowns on quarterbacks by other OL teammates) most closely resembles studies by Idson and Kahane (2000).

Model 1: Wage Determination Model-All OL Positions

(M1)

$$LOG(SAL_{it}) = a_0 + a_1 W_{i(t-1)} + a_2 X_{i(t-1)} + a_3 Y_{i(t-1)} + a_4 Z_{i(t-1)} + \psi_{i(t-1)}$$

Where i denotes an individual player, t denotes the year, SAL_{it} is the average salary, $W_{i(t-1)}$ is a vector of individual performance statistics, $X_{i(t-1)}$ is a vector of individual characteristics, $Y_{i(t-1)}$ is a vector of team performance statistics, $Z_{i(t-1)}$ is a vector of nonoffensive line position statistics, and $\psi_{i(t-1)}$ is the error term. Additionally, interaction terms and controls are used. The interaction terms of interest are (proportion of knockdowns on quarterbacks by individual offensive linemen)*(proportion of knockdowns on quarterbacks by other OL teammates) and (proportion of knockdowns on quarterbacks by individual offensive linemen)*(proportion of knockdowns on quarterbacks by OL of new team). The dichotomous variable of whether the player played in the same offensive scheme, pass rush and rush defense of opponents are used as controls.

Log of salary is used to account for salary being positively skewed in the NFL. Salary is defined as the annual average pay of a player's contract over the duration of a contract. Thus, if player A were to sign a five-year, $30 million contract, the average value of that contract would be $6 million per year. Salary is adjusted for inflation based on the hard salary cap for that particular season.

$W_{i(t-1)}$ consists of the following variables: number of regular season games started by an individual offensive lineman, proportion of penalties committed by an individual offensive lineman, and proportion of knockdowns on quarterbacks by an individual offensive lineman. To calculate the proportion of penalties committed by an individual offensive lineman, the total number of penalties on a particular offensive lineman are divided by the total number of offensive snaps in which the offensive lineman participated during the regular season. Similarly, to calculate proportion of knockdowns on quarterbacks by an individual offensive lineman, the regular season number of knockdowns on quarterbacks by an individual offensive lineman would be divided by the total number of offensive pass plays in which an individual lineman was involved during the regular season.

In terms of coefficients that are assumed to be negative--proportion of penalties committed, proportion of knockdowns on quarterbacks--are individual performance statistics related to not making mistakes, in which higher values reflect negatively on a player's ability. In particular, a lower proportion of knockdowns on quarterbacks by an individual offensive lineman would seem to indicate that the offensive lineman is doing a good job in pass protection because he is not responsible for allowing an opposing defender to hit the quarterback. In terms of proportion of penalties committed by an individual offensive lineman, penalties could be called during pre-snap (before a play begins), during a play, or a after a play. A typical pre-snap penalty might be false start, while holding might be a typical penalty during a play. Unnecessary roughness could be a penalty that occurs after the play ends. Whatever the scenario, penalties on players reflect errors on that particular individual. Thus, a higher proportion of penalties committed by an individual offensive lineman indicates that the player is more prone to mistakes and would be expected to be paid less because he is inhibiting the offensive unit from successfully moving the ball down the field.

The coefficient of the number of regular season games started by an individual offensive lineman is expected to be non-negative. The variable itself ranges in values from 0 (no games started) to 16, which would be the full regular season number of NFL games. The number of games started demonstrates player quality and availability, which is valued by an NFL team.

$X_{i(t-1)}$ consists of the following variables: age, race, overall pick, and position. Age acts as a proxy for experience, which can be a factor in improving one's skills but also could contribute to a deterioration in skill, because the offensive line position is a form of highly skilled manual labor. Age is expected to have a negative coefficient because each additional year decreases the player's salary from his peak production year, which could coincide with the year he becomes a free agent. Age squared is incorporated in the model and is used to determine if there is a nonlinear decrease in player salary. Overall pick, in which a player is selected in the NFL draft acts as a proxy for talent level. Overall pick is expected to be negative because a larger number indicates that a player was drafted in later rounds.

For position, centers are paid less and offensive tackles are paid the most with the position of guard being in the middle. This is conjectured based on the opinion that the most athletic offensive linemen play at the offensive tackle position. For race, non-White or

Hispanic (dichotomous variable) are included in the model because sport gives one the unique ability to investigate racial discrimination empirically. Race is not expected to play a role in the wage determination of an offensive lineman because an open, competitive market would bid away discrimination.

$Y_{i(t-1)}$ consists of the following variables of interest: rushing yards, passing yards, and total points scored. $Z_{i(t-1)}$ consists of the following variables of interest: adjusted line yards, running backs' yards before contact and average yards per carry, power percentage, location run proportion, and stuff percentage by the offensive lineman's team. A greater number of rushing yards, passing yards, and points scored indicate good performance of the team unit, which reflects positively on an offensive lineman's contribution. Similarly, a larger adjusted line yards, running backs' yards before contact and average yards per carry, and power percentage indicate a successful unit on offense. A higher location run proportion indicates a team's tendency to run behind or near a particular offensive lineman, which would reflect his strength in terms of being a run blocker relative to his fellow offensive linemen. Thus, the coefficients mentioned above in $Y_{i(t-1)}$ and $Z_{i(t-1)}$, with the exception stuff percentage are expected to be non-negative.

A higher value of stuff percentage, which indicates the percentage of runs in which a running back gains no yards or is tackled for a loss, implies a higher failure rate of the offensive line as a unit. Thus, it is expected to have a negative coefficient. For calculating running backs' yards before contact, a weighted average for running backs only is calculated for each particular team because rushes performed by wide receivers tend to occur rarely and as a form of deception (e.g., wide receiver reverse), and rushes by a quarterback often are not the design of a play (e.g., quarterback scrambles) or are not of interest to the researcher (e.g., quarterback sneaks) as an evaluation tool for OL.

For the pass control variable, a higher value indicates a higher level of difficulty of opponent pass rush ability over the course of the season, which would reflect positively on an offensive lineman's ability to pass block. Pass control will be calculated as follows:

Step 1: Find all knockdowns on the quarterback of a particular team for an entire season.

Step 2: Estimate total pass plays (pass attempts + sacks) made against that team for the entire season.

Step 3: Take total knockdowns on quarterback of a particular team and divide by total pass plays.

Step 4: Repeat this for all 32 NFL teams.

Step 5: Find league average pass rush percentage for the particular season (summing up all pass rush percentages of individual teams and dividing by 32).

Step 6: Take team average pass rush percentage, and subtract the league average rushing percentage (denote it as normalized average pass rush percentage).

Step 7: Using the new statistic above, take the summation of all opponents "normalized average pass rush percentage" that a particular team plays against (division opponents will be counted twice, because teams play division opponents twice each season), and divide by 16.

A positive coefficient for proportion of knockdowns on quarterbacks by OL of new team is expected because teams that are struggling to protect their quarterback would value a particular offensive lineman more than teams that were successful at protecting their quarterback. Scheme (a dichotomous variable) is expected to have a positive coefficient as an offensive lineman playing in the same scheme indicates that the player's skill set is a good match for what the team asks of its offensive lineman.

Other coefficients that are expected to be negative are: run control and proportion of knockdowns on quarterbacks by other OL teammates. A lower value of run control would indicate a higher-level difficulty of opponent run defense for the OL over the course of the season. Run control is calculated as follows:

Step 1: Find rushing defense by calculating running backs' weighted average YBC per each team that an individual offensive lineman started against, and find the average of NFL team per particular season.

Step 2: Take each team's running backs' weighted average YBC, and from it, subtract the league average running backs' weighted average YBC of that year.

Step 3: Take the sum of the "normalized" values together (up to 16 values), and divide by the number of previous regular season games started by that particular offensive linemen.

The coefficient of other knockdowns allowed proportion is expected to be negative because poor play by one's teammates would hurt the overall production of the offensive line as a unit. This would reflect poorly on the offense as a whole and on the individual offensive lineman. Other knockdowns allowed proportion will be computed as follows:

Step 1: Take the summation of all knockdowns on quarterback allowed by all offensive linemen on a particular team for a particular year, with the exception of the individual offensive lineman being evaluated.

Step 2: Take the given value from step 1, and divide it by the total offensive pass plays for the team in that year (estimated as pass attempts + sacks) in which the individual player participated.

With regards to interaction terms, a positive coefficient on (proportion of knockdowns on quarterbacks by an individual offensive lineman)*(proportion of knockdowns on quarterbacks by OL of new team) or proportion ratio new = (proportion of knockdowns on quarterbacks by an individual offensive lineman)/(proportion of knockdowns on quarterbacks by OL of new team) would indicate that teams place greater value on an offensive lineman who has the same pass-blocking ability as the unit that they played on during the previous season. A negative coefficient would indicate that teams place greater value on an offensive lineman who has better pass-blocking ability relative to the new team's pass protection.

A positive coefficient on the interaction term of (proportion of knockdowns on quarterbacks by an individual offensive lineman)*(proportion of knockdowns on quarterbacks by other OL teammates) or proportion ratio current = (proportion of knockdowns on quarterbacks by an individual offensive lineman)/(proportion of knockdowns on quarterbacks by other OL teammates), would indicate that an individual offensive lineman is positively compensated by playing at the same level as his teammates. Negative coefficients on the terms would indicate that an individual OL is positively compensated by being better in pass protection relative to his teammates.

Other models are illustrated below to address the heterogeneity due to the position. M2 will focus solely on the position of tackles, M3 on guards, and M4 on centers. Thus, M2, M3, and M4 will have the same explanatory variables as the M1 model. This is of interest because certain metrics might better capture a particular position's value.

Models 2, 3, and 4: Wage Determination Model of Particular Positions

(M2, M3, M4)

$$LOG\left(SAL_{ijt}\right) = t_0 + t_1 W_{ij(t-1)} + t_2 X_{ij(t-1)} + t_3 Y_{ij(t-1)} + t_3 Z_{ij(t-1)} + \alpha_{ij(t-1)}$$

Where j denotes whether player i is a tackle, guard, or center at time t, SAL_{ijt} is the average salary, $W_{ij(t-1)}$ is the vector of individual performance statistics, $X_{ij(t-1)}$ is the vector of individual characteristics, $Y_{ij(t-1)}$ is the vector of team performance statistics, and $Z_{ij(t-1)}$ is the vector nonoffensive line position statistics, and $\alpha_{ij(t-1)}$ is the error term. Additionally, proportion ratio new, proportion ratio current, scheme, and interaction terms of (proportion of knockdowns on quarterbacks by an individual offensive lineman)*(proportion of knockdowns on quarterbacks by other OL teammates) and (proportion of knockdowns on quarterbacks by an individual offensive lineman)*(proportion of knockdowns on quarterbacks by OL of new team) at time t-1 are evaluated.

Population and Sample

The population consists of all offensive linemen in the NFL. The sample consists of offensive linemen in the NFL from the years of 2010-2015 who were unrestricted free agents and who signed contracts (not franchise tag or transition tag) in the subsequent years of 2011-2016.

When NFL rookies are drafted, the team that drafted him owns his rights. Thus, the player would sign only with that team unless the team relinquishes the rights to him via trade or release. Unrestricted free agents, unlike rookies who are drafted, unrestricted free agents have the opportunity to sell their services to all 32 NFL teams, assuming they were not designated with what the NFL calls a franchise tag or transition tag. In order to capture a more competitive market that better identifies a player's true value, only unrestricted free agent OL are evaluated.

Data Source

STATS LLC provides data on regular season knockdowns on quarterbacks by an individual offensive lineman, running backs' yards before contact, team yards before contact allowed (to calculate run control), team defensive knockdowns on quarterbacks (to calculate pass control), regular season penalties committed by individual offensive linemen, and offensive snaps (seasons 2013-2015). Pro Football Reference provides data on the number of regular season games started by individual offensive linemen (including position started) and run direction, which is used to calculate the location run proportion metric, and team passing attempts and rushing attempts, which is used to calculate the run-to-pass ratio. Additionally, Pro Football Reference provides the offensive schemes of the particular teams. Football Outsiders provides adjusted line yards, running backs' average yards per carry, power percentage, and stuff percentage of respective teams. NFL.com provides pictures of individual players to determine race, game data on passing attempts, rushing attempts, and sacks to estimate offensive snaps for applicable years. Spotrac is used to provide wages and ages of NFL OL at the time of their signing of a free agent contract for the upcoming season.

Data Analysis

Ordinary least squares with clustered standard errors is utilized in multiple linear regression analysis for all of the models. The team that signed the player in unrestricted free agency acts as a cluster to control for possible team effects.

Results

A total of 391 year × player observations were collected from 244 free-agency NFL offensive linemen who played from 2011 to 2016. There was missing data from some of the observations for reasons such as the free agent offensive lineman not playing a single offensive snap the previous season prior to becoming a free agent. Table 1 illustrates the summary statistics gathered for the response variable and all variables of interest.[1]

Table 1

Descriptive Statistics

Variables	Observations	Mean	Standard Deviation	Minimum	Maximum
Adjusted salary	391	$2.581*10^6$	$2.532*10^6$	$7.000*10^5$	$1.220*10^7$
Individual knockdowns allowed proportion	290	0.022	0.017	0	0.110
Games started	375	8.035	6.548	0	16
Individual proportion of plays with a penalty	289	0.008	0.007	0	0.057
Team rushing yards	373	1815.338	322.654	$1.204*10^3$	$2.762*10^3$
Team passing yards	373	3716.968	602.688	$2.434*10^3$	$5.444*10^3$
Team points scored	373	361.274	70.433	$1.930*10^2$	$6.060*10^3$
Proportion ratio current	281	0.246	0.206	0	1.575
Proportion ratio new	282	0.213	0.168	0	1.067
Adjusted line yards (ALY)	280	3.942	0.315	2.930	4.950

[1] One big regression (OBR) model was created considering all explanatory variables in a single model but was not included in the results section because of multicollinearity issues. Tables for comparison of the OBR model and selected models in the results section are included in Appendix A.

Variables	Observations	Mean	Standard Deviation	Minimum	Maximum
Yards before contact (YBC)	280	1.926	0.292	1.110	2.820
Average yards per carry (AYPC)	280	4.155	0.401	2.950	5.670
Power percentage	280	62.311	8.678	37	83
Stuff percentage	280	19.657	2.930	12	27
Location run proportion	218	0.468	0.128	0.160	0.833
Run-to-pass ratio	367	0.793	0.162	0.487	1.323

Adjusted salary is our response variable of interest. The log of adjusted salary is used as the dependent variable in the models. As for the independent variables, in the models below, the proportion of knockdowns on quarterbacks by an individual offensive lineman will be referred to as “individual knockdowns allowed proportion,” the number of regular season games started by an individual offensive lineman will be referred to as “games started,” and the proportion of penalties committed by an individual offensive lineman will be referred to as “individual proportion of plays with a penalty.” These variables capture the individual performance of an offensive lineman. Regular season rushing yards, passing yards, and total points scored by the offensive lineman’s team measure group performance. They are denoted in the models as “Team rushing yards,” “Team passing yards,” and “Team points scored.” The proportion ratio current and the proportion ratio new variables address whether relative performance of NFL offensive linemen affect an individual’s wages. ALY, running backs' YBC and AYPC, power percentage, and stuff percentage account for other players' performance. Run proportion and run-to-pass ratio consider play selection.

Based on Table 1, the average proportion of knockdowns allowed by an individual in pass plays is 0.022, while penalty proportion is 0.008, demonstrating a low occurrence of these particular events for offensive linemen who reach free agency. Also, teams on average move the ball twice as far by passing versus running the football. Not surprisingly, average team pass attempts is a greater number than run attempts, based on the mean run-to-pass ratio being lower than 1.

Individual Player Performance and Wage Determination

This section explores the relationship between individual player performance and wage determination. All measurements of individual player performance significantly affect the salary of an NFL offensive lineman, regardless of position, at the significance level of 0.10. The three metrics used are regular season penalty proportion, games started, and the individual knockdowns allowed proportion. The first metric indicates how well offensive linemen avoid being penalized. The second measures overall quality and availability of the player, while the third measures an individual's ability to pass-block effectively. General managers determine wages of offensive linemen, in part because of their individual performance and productivity.

Effect on Salary of Knockdowns Allowed

From a labor productivity standpoint, as the individual knockdowns allowed proportion increases, offensive linemen are less productive. Thus, this variable is expected to create a negative impact on salary. Following the author's expectations, the individual knockdowns allowed proportion has a significant negative effect on a player's open market salary at the 0.05 significance level, regardless of position.[2] Table 2 indicates that if the proportion of knockdowns allowed on a quarterback increases by one standard deviation (0.017), adjusted salary decreases between 13.2% of a standard deviation ($334,186) and 20.6% of a standard deviation ($521,533).

From a practical standpoint, offensive linemen act as insurance on quarterbacks. The fewer times a quarterback is hit, the less likely they are to be injured. This means that general managers value and reward offensive linemen who are able to successfully pass-block and limit the number of knockdowns on their quarterback. Furthermore, considering the average yearly salary for the top 20 paid quarterbacks in 2016 is roughly $21.6 million dollars or 13.9% of the total salary cap (spotrac.com), general managers pay a high premium to offensive linemen who are able to protect their prized possession.

Table 2

Effect on Salary of the Proportion of Individual Knockdowns Allowed

Variables	(1)	(2)	(3)
Individual knockdowns allowed proportion	-0.206*** (-2.93)	-0.175*** (-2.25)	-0.132** (-2.21)
Non-White or Hispanic	-0.137** (-2.04)	-0.176** (-2.42)	-0.097* (-1.68)
Other knockdowns allowed proportion	-0.243*** (-3.77)	-0.225*** (-3.21)	-0.159*** (-2.82)
Position guard	0.020 (0.21)	0.014 (0.13)	0.023 (0.27)
Position tackle	0.130 (1.21)	0.143 (1.21)	0.117 (1.28)

[2] Alternative specifications included interaction components and clustered standard errors but found no statistical significance, and the results remain unchanged.

Variables	(1)	(2)	(3)
Same scheme	0.104 (1.65)	0.085 (1.25)	0.059 (1.08)
Age	-0.206*** (-3.20)	-0.271*** (-3.87)	-0.254*** (-4.48)
Overall pick	-0.232*** (-3.63)	-0.213*** (-3.06)	-0.130** (-2.32)
Games started			0.435*** (7.19)
Observations	204	176	230
R squared	0.245	0.248	0.361

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Games Started

For two reasons, it is expected that the variable of games started would have a positive effect on the salary of an offensive lineman. An offensive lineman starting games demonstrates durability and quality of play for that individual. Table 3 on the following page, shows that the number of regular season games started by an individual offensive lineman has a significant positive effect on a player's open market salary at the 0.01 significance level.[3] Specifically, a standard deviation increase in the number of games started (6.5) increases a player's adjusted salary by 55.6% of a standard deviation ($1,407,634).

Considering that there are 16 games during the regular NFL season, 6.5 games started is roughly 40% of the entire season. With such a short regular season, each game has heightened importance in terms of whether a team makes the playoffs or not. The cliché

[3] Game started*position was insignificant in both models.

statement, “The number one ability is availability” appears to hold true for offensive linemen, based on the large beta coefficient in the results.

Table 3

Effect on Salary of Games Started

Variables	(1)	(2)
Games started	0.556***	0.556***
	(11.64)	(11.68)
Non-White or Hispanic	-0.095*	-0.095*
	(-1.88)	(-1.93)
Position guard	-0.003	
	(-0.04)	
Position tackle	0.052	
	(0.69)	
Same scheme	0.079*	0.079*
	(1.70)	(1.71)
Age	-0.179***	-0.178***
	(-3.76)	(-3.78)
Overall pick	-0.149***	-0.149***
	(-3.08)	(-3.09)
Tackle		0.054
		(1.09)
Observations	301	301
R squared	0.381	0.381

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$.

Effect on Salary of Penalties

Linemen who are less prone to being penalized are more productive overall on the field. Thus, it is expected that individual proportion of plays with a penalty would have a negative impact on the salary of an offensive lineman. Following expectations, Table 4 illustrates that the previous regular season penalty proportion denoted "Pen/Plays" has a significant negative effect on a player's open market salary at the 0.01 significance level. A standard deviation increase in the penalty proportion (0.0075) decreases a player's adjusted salary between 14.4% of a standard deviation ($364,567) and 22.3% of a standard deviation ($564,572).

Further evaluating, the goal of a team's offense is to continuously "move the chains," which means that the team on offense gets a first down. Repeatedly gaining first downs gives the offense a greater chance to score. Additionally, it allows the team's defense to be well-rested and can tire the opponent's defense. Considering that it takes an offense moving the ball 10 yards forward to earn a first down in a total of four downs, a holding penalty of 10 yards by an offensive lineman doubles the necessary distance necessary to successfully move the chains. Thus, general managers clearly take note of these types of errors when evaluating offensive linemen.

Table 4

Effect on Salary of Penalties

Variables	(1)	(2)	(3)
Individual proportion of plays with a penalty	-0.195*** (-2.85)	-0.144* (-1.96)	-0.223*** (-3.48)
Pass control	-0.124* (-1.90)	-0.118* (-1.71)	-0.123* (-1.95)
Run control	0.112* (1.72)	0.170** (2.47)	0.133** (2.09)
Non-White or Hispanic	-0.122* (-1.79)	-0.136* (-1.90)	-0.116* (-1.77)
Position guard	0.019 (0.20)	0.022 (1.21)	0.044 (0.48)

Variables	(1)	(2)	(3)
Position tackle	0.150	0.165	0.111
	(1.48)	(1.52)	(1.14)
Same scheme	0.086	0.070	0.035
	(1.33)	(1.03)	(0.56)
Age	-0.209***	-0.271***	-0.187***
	(-3.19)	(-3.85)	(-2.93)
Overall Pick	-0.275***	-0.250***	-0.271***
	(-4.23)	(-3.60)	(-4.33)
Observations	208	180	229
R squared	0.213	0.239	0.183

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Team Performance and Wage Determination

This section explores the relationship between team performance and wage determination. Team performance positively impacts the wages of an NFL offensive lineman, regardless of position. Team rushing yards is a product of a variety of skills at different positions, such as the elusiveness of the running back and the ability of his teammates, especially offensive linemen, to run-block for him. Likewise, team passing yards is a product of the skills of many players, such as an offensive line's ability to pass-block for the quarterback; a quarterback's ability to throw accurately; and the ability of wide receivers, tight ends, and running backs to get open and catch the ball. Even more pieces come into play because points scored is the sum of many components, such as field position, special teams, and the ability of the offense to move the ball and ultimately score. Ultimately, higher team performance is valued by a general manager when determining salary.

Effects on Salary of Passing and Rushing Yards

Both team rushing yards and passing yards measure different aspects of offensive production in terms of moving the ball successfully. Thus, it is expected that both have a positive effect on the salary of an offensive lineman. According to Table 5, a standard deviation increase in team rushing yards (323) improves the adjusted salary of a player between 19.6% of a standard deviation ($496,216) and 28.6% of a standard deviation ($724,070), while an increase of a standard deviation of team passing yards (603) beta coefficient is between 10.7% ($270,894) and 21.0% ($531,660). Total rushing yardage has a greater impact on salary than passing yards, when comparing the coefficients of each model separately.

However, by examining statistics from the 2016 NFL season, teams pass the ball on average 58% of the time, and the average total passing yards per team is roughly 2.21 times that of total rushing yards. Taking the ratio of total passing yards to rushing yards into consideration, it appears that general managers might value passing yards as much or even more than rushing yards (by multiplying the coefficient of team passing yards by 2.21). This would support the idea of offensive linemen being particularly valued for their ability to protect the quarterback because it is primarily the quarterback who throws the ball on pass plays.

Table 5

Effects on Salary of Passing and Rushing Yards

Variables	(1)	(2)	(3)
Team rushing yards	0.262***	0.286***	0.196***
	(4.07)	(4.16)	(4.14)
Team passing yards	0.210***	0.201***	0.107**
	(3.17)	(2.80)	(2.23)
Non-White or Hispanic	-0.121*	-0.124*	-0.085*
	(-1.81)	(-1.74)	(-1.72)
Position guard	-0.025	-0.016	-0.012
	(-0.27)	(-0.16)	(-0.17)
Position tackle	0.047	0.065	0.033
	(0.48)	(0.63)	(0.45)
Same scheme	0.097	0.075	0.070
	(1.55)	(1.12)	(1.54)
Age	-0.202***	-0.257***	-0.185***
	(-3.14)	(-3.72)	(-3.99)
Overall pick	-0.289***	-0.284***	-0.163***
	(-4.48)	(-4.05)	(-3.43)
Games started			0.581***
			(12.36)
Observations	208	180	301
R squared	0.228	0.254	0.418

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Points Scored

It is expected that total team production reflects well on the individual productivity of an offensive lineman. Thus, the metric of team points scored would have a positive significant effect on the salary of an offensive lineman. Following the researcherr's expectations, team points scored has a positive impact on the salaries of NFL offensive linemen at the significance level of 0.01. According to Table 6, a standard deviation increase in team points scored (70) improves an offensive lineman's adjusted salary from 21.4% of a standard deviation ($541,787) to 31.6% of a standard deviation ($800,022), depending on the model.

Furthermore, the average team in 2016 scored 364 points per game. Using the 2016 Miami Dolphins as an average offense, which scored 363 points, a 70-point increase would move them from 17th in team points scored to fourth in the NFL (Pro Football Reference). With such a jump, we might expect a larger increase in salary. However, team points scored has many variables to consider, such as the quality of play of the quarterback, running back, and wide receiver. Thus, the relatively small increase could be due to this metric's lack in precision with regards to measuring an individual offensive lineman's actual play.

Table 6

Effect on Salary of Points Scored

Variables	(1)	(2)	(3)
Team points scored	0.311***	0.316***	0.214***
	(4.99)	(4.72)	(4.77)
Non-White or Hispanic	-0.113*	-0.123*	-0.078
	(-1.70)	(-1.74)	(-1.95)
Position guard	-0.003	0.007	0.021
	(-0.03)	(0.08)	(0.31)
Position tackle	0.087	0.102	0.054
	(0.91)	(1.00)	(0.74)
Same scheme	0.101	0.077	0.080*
	(1.61)	(1.16)	(1.79)

Variables	(1)	(2)	(3)
Age	-0.183***	-0.233***	-0.187***
	(-2.88)	(-3.41)	(-4.06)
Overall pick	-0.281***	-0.276***	-0.159***
	(-4.44)	(-4.03)	(-3.40)
Games started			0.566***
			(12.26)
Observations	208	180	301
R squared	0.239	0.260	0.426

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Relative Performance and Wage Determination

This section explores the relationship between relative performance and wage determination. Relative performance has mixed results in the determination of wages. On one hand, the difference between a player's proportion of knockdowns allowed on quarterbacks and his offensive line teammates has no impact on the individual player's salary. Thus, whether an offensive lineman's performance is distinct from his teammates is not of importance in determining his salary. However, the difference in knockdown proportion between a player and his new team has a significant negative effect on salary. A player whose pass-blocking performance is better relative to his new team is compensated more ceteris paribus.

Effect on Salary of the Difference between a Player's Proportion of Knockdowns Allowed on Quarterbacks and His Offensive Line Teammates

It is thought by the author that a difference between a player's proportion of knockdowns allowed on quarterbacks and his offensive line teammates would have an impact on salary because the offensive line works as a singular unit. However, to the author's surprise, proportion ratio current is not significant at the significance level of 10%. Table 7 illustrates that there is no relationship in wage determination between player pass-blocking performance and teammate pass-blocking proportion.

This indicates that a general manager is not influenced by the play of an individual offensive lineman's teammates in pass blocking. In other words, the pass-blocking ability of individual offensive linemen is being separately evaluated from his teammates. Thus, a star left tackle on the open market who excels in pass blocking will be compensated the same whether on a team with fellow OL teammates who are also great at pass blocking or with teammates who are poor at pass blocking.

Table 7

Effect on Salary of the Difference between a Player's Proportion of Knockdowns Allowed on Quarterbacks and His Offensive Line Teammates

Variables	(1)	(2)	(3)
Proportion ratio current	-0.072	-0.021	-0.061
	(-0.98)	(-0.26)	(-0.98)
Non-White or Hispanic	-0.142**	-0.155**	-0.095
	(-2.02)	(-2.05)	(-1.58)
Pass control	-0.158**	-0.161**	-0.121**
	(-2.37)	(-2.25)	(-2.14)
Position guard	0.030	0.006	0.047
	(0.30)	(0.05)	(0.55)
Position tackle	0.123	0.111	0.136
	(1.09)	(0.90)	(1.45)
Same scheme	0.119*	0.102	0.067
	(1.78)	(1.43)	(1.19)

Variables	(1)	(2)	(3)
Age	-0.197***	-0.260***	-0.257***
	(-2.91)	(-3.56)	(-4.38)
Overall pick	-0.257***	-0.221***	-0.143**
	(-3.82)	(-3.04)	(-2.46)
Games started			0.472***
			(7.89)
Observations	204	176	225
R squared	0.163	0.183	0.334

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5. Proportion ratio current = (proportion of knockdowns on quarterbacks by an individual offensive lineman)/(proportion of knockdowns on quarterbacks by other OL teammates).

Effect on Salary of the Difference between the Player's Proportion of Knockdowns Allowed on Quarterbacks and His New Team

It is expected by the author that the greater the relative quality differential between individual knockdowns allowed proportion and new team knockdowns allowed proportion, the greater the wage would be for the individual lineman. Thus, the variable proportion ratio new (defined in the table notes) would have a negative coefficient. Following the author's expectations, the values in the first row of Table 8 illustrate that proportion ratio new is negative and significant at the 0.10 level in all models. If proportion ratio new increases by one standard deviation (0.168), adjusted salary decreases between 17.5% of a standard deviation ($443,050) and 10.6% of a standard deviation ($268,362). Thus, there is a salary penalty for players with a relatively higher proportion of knockdowns allowed in the previous season when compared to their performance with their new team.

Alternatively, we could state that an offensive lineman who excels in pass blocking relative to his new team will be paid more. From a labor economics perspective, this makes sense in the open market because a team that struggles to protect its quarterback would have a greater need to address this aspect of their team. Signing an offensive lineman in free agency is one way for an organization to address this deficiency.

Table 8

Effect on Salary of the Difference between the Player's Proportion of Knockdowns Allowed on Quarterbacks and His New Team's Proportion of Knockdowns Allowed on Quarterbacks

Variables	(1)	(2)	(3)
Proportion ratio new	-0.167**	-0.175**	-0.106*
	(-2.37)	(-2.29)	(-1.71)
Non-White or Hispanic	-0.139**	-0.171**	-0.091
	(-2.01)	(-2.31)	(-1.54)
Pass control	-0.142**	-0.134*	-0.113**
	(-2.14)	(-1.88)	(-2.00)
Position guard	0.054	0.047	0.051
	(0.54)	(0.44)	(0.60)
Position tackle	0.162	0.188	0.147
	(1.51)	(1.63)	(1.60)
Same scheme	0.119*	0.098	0.072
	(1.82)	(1.40)	(1.27)
Age	-0.202***	-0.274***	-0.262***
	(-3.03)	(-3.80)	(-4.49)
Overall pick	-0.258***	-0.223***	-0.142***
	(-3.88)	(-3.13)	(-2.46)
Games started			0.457***
			(7.58)

Observations	205	177	226
R squared	0.183	0.208	0.339

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5. Proportion ratio new = (proportion of knockdowns on quarterbacks by an individual offensive lineman)/(proportion of knockdowns on quarterbacks by OL of new team).

Individual Performance of Nonoffensive Line Positions and Wage Determination

This section explores the relationship between performance of nonoffensive line positions and wage determination. Using a variety of metrics, such as adjusted line yards (illustrated below in the following paragraph), many of the performance variables have a positive significant impact. Players such as running backs who contribute on a rushing play impact the salary of an offensive lineman, regardless of position, and general managers look at joint productivity as a factor when determining a player's salary.

Effect on Salary of Adjusted Line Yards

Adjusted Line Yards (ALY) is a unique metric developed by Football Outsiders that assigns credit to the offensive line on carries by a running back. A larger ALY indicates a more successful run-blocking offensive line. Thus, ALY is expected to have a positive impact on salary and should be a better evaluator of an offensive lineman's run blocking ability than average yards per carry (AYPC). Following the positive correlation expectations, Table 9 illustrates a positive significant effect on salary of an offensive lineman at the 0.01 significance level. A standard deviation increase in ALY (0.32) would improve the player's adjusted salary between 21.4% of a standard deviation ($541,787) and 23.4% of a standard deviation ($592,421).

The unstandardized coefficients of ALY are 1,691,000 and 1,850,000, respectively. Thus, a one-unit increase in ALY will increase an offensive lineman's salary between those two values, depending on the model. Comparing the unstandardized coefficients of the AYPC are 1,398,000 and 1,448,000 based on standardized coefficients calculated in Section 4.4.3, it appears that ALY might indeed be a better evaluator of run-blocking ability because the value of the ALY coefficient is greater than AYPC.

Table 9

Effect on Salary of ALY

Variables	(1)	(2)	(3)
ALY	0.234*** (3.64)	0.219*** (3.26)	0.214*** (3.92)
Run control	0.110* (1.71)	0.175** (2.59)	0.106* (1.94)
Non-White or Hispanic	-0.165** (-2.45)	-0.155** (-2.18)	-0.114** (-1.98)
Position guard	-0.021 (-0.22)	-0.020 (-0.20)	0.002 (0.02)
Position tackle	0.040 (0.41)	0.076 (0.73)	0.070 (0.83)
Same scheme	0.073 (1.14)	0.065 (0.97)	0.019 (0.35)
Age	-0.190*** (-2.94)	-0.253*** (-3.66)	-0.243*** (-4.31)
Overall pick	-0.279*** (-4.32)	-0.254*** (-3.71)	-0.166*** (-2.96)
Games started			0.483*** (8.59)
Observations	208	180	230
R squared	0.216	0.253	0.369

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Running Backs' Yards before Contact (YBC)

It is expected by the author that running backs' YBC would have a positive significant effect on the salary of an offensive lineman because it reflects a player's ability to maintain a run block. Table 10 demonstrates that YBC has a significant positive effect on salary at the significance level of 0.05. In particular, Model 1 suggests that YBC has a lesser effect on the salary of an offensive tackle versus the salary of a center, based on the significant interaction effect at the 0.10 significance level. From a general management perspective, they value offensive tackles more based on their pass-blocking than run-blocking ability versus centers. In light of the above results of Model 1, an NFL where passing is more common than running the football and considering that 15 of the top 20 paid offensive linemen based on average salary of 2016 were offensive tackles, while only one center made the top 20 list, we have further evidence of offensive linemen being more valued for their pass-blocking ability than their run-blocking ability.

Table 10

Effect of Salary of Running Backs' YBC

Variables	(1)	(2)	(3)
YBC	0.474**	0.170**	0.155***
	(2.57)	(2.42)	(2.73)
Run control	0.112*	0.166**	0.095*
	(1.70)	(2.39)	(1.67)
Non-White or Hispanic	-0.161**	-0.144**	-0.099*
	(-2.35)	(-2.00)	(-1.69)

Variables	(1)	(2)	(3)
Position guard	1.006 (1.43)	-0.027 (-0.27)	-0.010 (-0.12)
Position tackle	1.338* (1.90)	0.084 (0.80)	0.080 (0.93)
YBC*guard	-1.100 (-1.51)		
YBC*tackle	-1.355* (-1.86)		
Same scheme	0.065 (1.01)	0.056 (0.83)	0.025 (0.44)
Age	-0.181 (-2.74)	-0.238*** (-3.36)	-0.227*** (-3.95)
Overall pick	-0.281*** (-4.28)	-0.257*** (-3.69)	-0.169*** (-2.97)
Games started			0.476*** (8.32)
Observations	208	180	230
R squared	0.206	0.233	0.347

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Running Backs' AYPC

As a similar metric for run blocking to ALY and YBC, the AYPC of running backs reflects an offensive lineman's ability to run-block. The better a player blocks, the larger the running backs' AYPC ceteris paribus. The findings reflect these expectations that offensive linemen play a key role in the performance of a running back and are demonstrated in Table 11. Specifically, the table shows that a standard deviation increase in AYPC (0.40) increases adjusted salary 22.1% of a standard deviation ($559,509) to 22.9% of a standard deviation ($579,763).

Table 11

Effect on Salary of Running Backs' AYPC

Variables	(1)	(2)	(3)
AYPC	0.229***	0.222***	0.221***
	(3.59)	(3.30)	(4.02)
Run control	0.120*	0.179***	0.099*
	(-2.04)	(2.67)	(1.81)
Non-White or Hispanic	-0.151**	-0.144**	-0.104*
	(-2.24)	(-2.03)	(-1.81)
Position guard	-0.045	-0.043	-0.024
	(-0.48)	(-0.44)	(-0.30)
Position tackle	0.025	0.053	0.057
	(0.26)	(0.51)	(0.68)
Same scheme	0.075	0.061	0.021
	(1.17)	(0.91)	(0.38)
Age	-0.204***	-0.271***	-0.247***
	(-3.14)	(-3.92)	(-4.40)
Overall pick	-0.292***	-0.268***	-0.182***
	(-4.51)	(-3.89)	(-3.24)
Games started			0.483***
			(8.59)
Observations	208	180	230
R squared	0.215	0.254	0.371

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Power Percentage

Looking at successful rushing attempts in short-distance situations on third down and fourth down, it is expected by the author that power percentage (which measures this) of the offensive lineman's team would have a negative impact on salary of that individual. After all, in these crucial downs, offensive linemen play a key role in the success of the play. However, contrary to the researcher's expectations, Table 12 shows that power percentage has no significant effect on salaries at the 0.10 significance level. For general managers, this finding supports the idea that other players on offense are being held accountable for their play in short-distance situations and that what offensive lineman do prior to reaching these circumstances is of more importance.

Table 12

Effect on Salary of Power Percentage

Variables	(1)	(2)	(3)
Power percentage	0.055	0.062	0.062
	(0.84)	(0.90)	(1.11)
Run control	0.140**	0.205***	0.127**
	(2.12)	(2.98)	(2.27)
Non-White or Hispanic	-0.146**	-0.147**	-0.105*
	(-2.10)	(-2.02)	(-1.77)
Position guard	-0.020	-0.025	0.003
	(-0.21)	(-0.25)	(0.04)
Position tackle	0.058	0.085	0.092
	(0.58)	(0.80)	(1.06)
Same scheme	0.083	0.069	0.029
	(1.26)	(1.01)	(0.51)
Age	-0.199***	-0.269***	-0.245***
	(-2.97)	(-3.76)	(-4.20)

Variables	(1)	(2)	(3)
Overall pick	-0.283***	-0.253***	-0.169***
	(-4.23)	(-3.56)	(-2.89)
Games started			0.476***
			(8.20)
Observations	208	180	230
R squared	0.167	0.210	0.328

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Stuff Percentage

Stuff percentage of the offensive lineman's team, which measures an offense's inability to move the football forward on a running play, is expected by the researcher to have a negative impact on the salary of an offensive lineman because it represents plays that do not achieve their desired result. Furthermore, a running play, in which a running back fails to move the ball across the line of scrimmage, could be due to a blown block by a particular offensive lineman. As expected, stuff percentage has a negative significant effect on the salary of an offensive lineman at the 0.05 significance level. Table 13 shows that a standard deviation increase in the stuff percentage (2.93%) will decrease an offensive lineman's adjusted salary 15.4% of a standard deviation ($389,884) to 17.4% of a standard deviation ($440,518), depending on the model.

Table 13

Effect on Salary of Stuff Percentage

Variables	(1)	(2)	(3)
Stuff percentage	-0.167**	-0.174**	-0.154***
	(-2.59)	(-2.56)	(-2.78)
Non-White or Hispanic	0.134**	0.197***	0.122**
	(2.06)	(2.92)	(2.22)
Other knockdowns allowed proportion	-0.146**	-0.136*	-0.099*
	(-2.13)	(-1.91)	(-1.69)
Position guard	-0.030	-0.037	-0.009
	(-0.32)	(-0.37)	(-0.11)
Position tackle	0.054	0.086	0.089
	(0.55)	(0.83)	(1.03)
Same scheme	0.072	0.063	0.019
	(1.11)	(0.93)	(0.34)
Age	-0.176***	-0.243***	-0.228***
	(-2.66)	(-3.46)	(-3.98)
Overall pick	-0.275***	-0.243***	-0.159***
	(-4.19)	(-3.51)	(-2.79)
Games started			0.482***
			(8.42)
Observations	208	180	230
R squared	0.191	0.236	0.348

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Play Selection of a Team and Wage Determination

This section explores the relationship between play selection of a team and wage determination. Play selection is found to have no significant effect on the wages of different positions on the offensive line. Decision makers on the field, such as coaches and quarterbacks, do not impact offensive players' salary. Neither do team tendencies, as illustrated in the next section. The offensive lineman's salary is determined by their play on the field.

Effect on Salary of Location Run Proportion

It is thought that the more a team runs to the left or right or over the gap of a particular offensive linemen position as measured by the variable of interest, more value would be placed on the performance of that offensive lineman. He would be a key blocker on such plays. However, Table 14 illustrates that regular-season location run proportion of the offensive lineman's team has no impact on the salary of the individual at the significance level of 0.10.

This result might in large part be due to the metric not accounting for schematics such as pulling offensive linemen for particular plays. For instance, a play that ends up being run several gaps away from the particular offensive lineman would not be measured in the metric as illustrated in the picture below. However, the pulling offensive lineman might be a key block related to the success of a run play. Thus, the run proportion metric fails to account for certain key blocks of an offensive lineman such as the following diagrammed play.

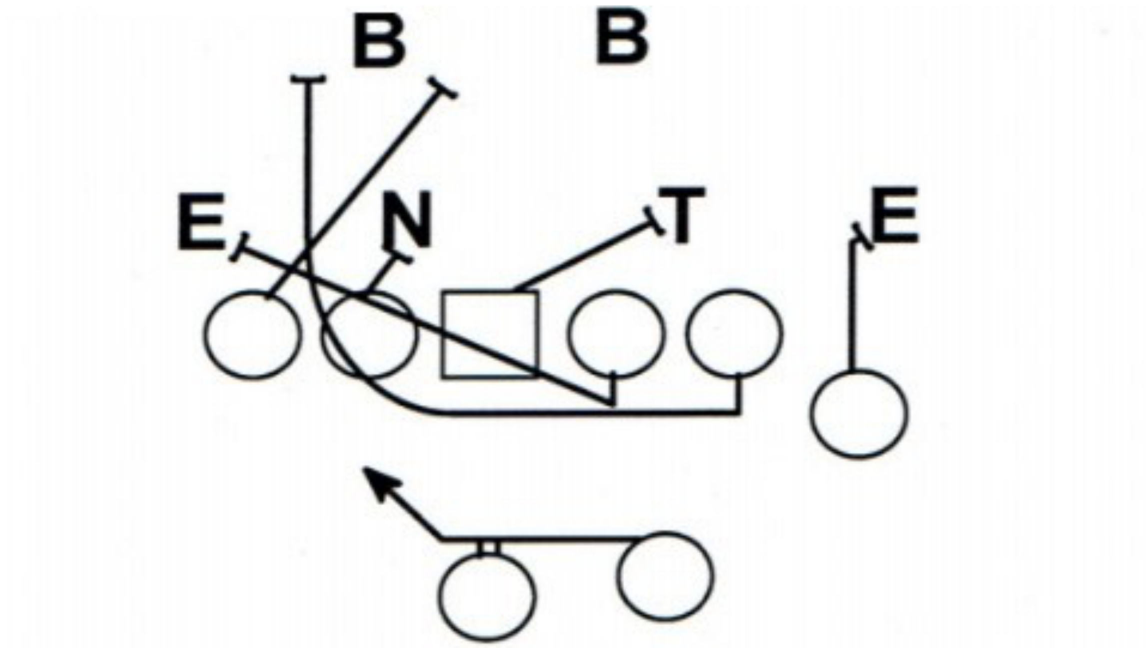


Table 14

Effect on Salary of Location Run Proportion

Variables	(1)	(2)	(3)
Location run proportion	0.007 (0.09)	-0.026 (-0.33)	-0.013 (-0.20)
AYPC	0.310*** (4.26)	0.304*** (3.93)	0.281*** (4.53)
Non-White or Hispanic	-0.177** (-2.36)	-0.195** (-2.47)	-0.118* (-1.85)
Age	-0.179** (-2.41)	-0.242*** (-3.08)	-0.220*** (-3.44)
Overall pick	-0.285*** (-3.90)	-0.294*** (-3.79)	-0.190*** (-3.00)
Games started			0.468*** (7.26)
Observations	158	137	179
R squared	0.208	0.237	0.353

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Effect on Salary of Run-to-Pass Ratio

Based on research demonstrating that a running back was paid more for specialization, a team that values running the ball or passing the ball more might place a higher value on offensive linemen who possess a unique skill set as a run blocker or pass blocker. Thus, the coefficients of the run-to-pass ratio would be of interest to the researcher. While the coefficients were positive in all three models of Table 15, the ratio of regular-season run attempts to pass attempts on the offensive lineman's team has no impact on the salary of the offensive lineman at the significance level of 0.10. Thus, general managers appear to focus solely on what the player does on the field versus the team tendency to run or pass the football.

Table 15

Effect on Salary of Run-to-Pass Ratio

Variables	(1)	(2)	(3)
Run-to-pass ratio	0.060	0.076	0.079
	(0.92)	(1.11)	(1.44)
Run control	0.142**	0.192***	0.118**
	(2.16)	(2.77)	(2.10)
Pass control	-0.126*	-0.124*	-0.095*
	(-1.92)	(-1.80)	(-1.70)
Non-White or Hispanic	-0.135**	-0.135*	-0.083
	(-2.05)	(-1.94)	(-1.47)
Age	-0.206***	-0.277***	-0.255***
	(-3.15)	(-4.01)	(-4.53)

Variables	(1)	(2)	(3)
Overall pick	-0.299***	-0.277***	-0.177***
	(-4.54)	(-4.02)	(-3.09)
Games started			0.467***
			(8.15)
Observations	207	179	230
R squared	0.170	0.212	0.335

Standardized beta coefficients; t statistics in parentheses; *** $p < .01$, ** $p < .05$, * $p < .1$. Model 1 is a subsample in which games started is greater than 3. Model 2 is a subsample in which games started is greater than 5.

Conclusion

How does one measure the value of an offensive lineman? This paper is one of the first to examine factors that impact the wages of National Football League (NFL) offensive linemen. The research questions: How are the wages for different positions in the NFL offensive line affected by individual player performance? How are the wages of different positions in the NFL offensive line affected by team performance? How does relative performance of NFL offensive linemen affect their wages? How does the individual performance of nonoffensive line positions affect the pay of different positions in the NFL offensive line? How does the play selection of a team affect pay of different positions in the NFL offensive line? The author addressed these questions by utilizing multivariate linear regression with adjusted salary as the response variable for the years 2011-2016. To measure these variables, data from resources such as STATS LLC and Spotrac was utilized.

Summary of Major Findings

Starting at the position level, an offensive tackle's run-blocking ability is less important than a center's ability to run-block. And regardless of an offensive lineman's position, the number of games started consistently has a significant impact on salaries in models, highlighting the importance of an offensive lineman being available to play. With the offensive line working as a unit on every offensive snap of the ball, the pass-blocking ability of an individual offensive lineman is separately evaluated from that of his teammates by personnel evaluators such as general managers. However, the difference between the player's proportion of knockdowns allowed on quarterbacks and his new team's proportion of knockdowns allowed on quarterbacks was negatively significant, a statistic that highlights the economical concept of diminishing returns. Lastly, the nonsignificant power percentage findings in models support the idea that other players on offense are being held accountable for their play in short-distance situations.

Summary of Findings

The author found that various factors at the individual and team levels consistently impact the salary of an NFL offensive lineman. These factors include individual player performance, team performance, and individual performance of nonoffensive linemen. Relative performance variables demonstrated mixed results, while play selection had no impact on salary.

For individual player performance, which captures individual productivity, regular season penalty proportion, games started, and proportion of knockdowns allowed were found to be significant. As expected, penalty proportion and knockdowns allowed had a negative association on salary, while games started was positively related. The findings demonstrate that individual labor components in a complex team sport such as football impact how a player is compensated. These results agree with findings in the literature done by others, such as Jones and Walsh (1988) and Idson and Kahane (2000). Those studies used various individual metrics in hockey and found that career average points scored per game and average penalty minutes per game were positive and significant in most of their wage models.

In terms of the effects of individual performance of nonoffensive line positions on the wages of offensive linemen, the present analysis indicated that all metrics of interest, with the exception of the power percentage, were positive and significant. Thus, the determination of an offensive lineman's salary also depends on the assistance that he receives from teammates, such as a running back. These findings advance work done on quarterbacks' wages by Berri and Simmons (2009), who utilized offense salary to proxy for quality of a quarterback's teammates. Offense salary was found to be positive and significant in most of the models.

To address team performance, passing yards, rushing yards, and points scored were analyzed. All were found to be positive and significant in the respective models. These team-level metrics demonstrate how joint production determines the wages of an individual offensive lineman by NFL general managers. These results support the findings of Simmons and Berri (2009), which determined that rushing yards and receiving yards of running backs positively impacted their salaries. For their other study (Berri & Simmons, 2009) on wage discrimination of quarterbacks, passing yards was found to be positive and significant.

For player performance relative to their teammates, effect on salary of the difference between a player's proportion of knockdowns allowed on quarterbacks and his offensive line teammates was not significant. This nonsignificance suggests that complementary effects in football might not be as important as in the sport of hockey, according to the study of Idson and Kahane (2000). They found a significant negative interaction between individual average penalty minutes per game and team average penalty minutes per game. That finding demonstrated diminishing returns in terms of roster composition of enforcers. However, diminishing returns is supported as the effect on salary of the difference between the player's proportion of knockdowns allowed on quarterbacks and his new team's proportion of knockdowns allowed on quarterbacks in my results was significant. Lastly, both metrics in the NFL that reflected play selection (run proportion and run-to-pass ratio) were found to be insignificant.

Managerial Applications

General managers and decision makers in the NFL can find value in the models as an additional multitool kit to help them evaluate players. Additionally, significant and nonsignificant coefficients of the explanatory variables can be used to look at possible market inefficiencies, in which a general manager could gain a competitive advantage. In the research, the nonsignificance of the power percentage variable could highlight the undervaluing of an offensive lineman's ability to succeed on third down and fourth down and on short-yardage plays. A general manager can use this material to compensate a player below the value of his actual production. In baseball, this idea is known as the moneyball hypothesis. The term originated after a reporter wrote about baseball executive Billy Beane's tactics for building championship team in the American League West division during the 2002 season, in spite of the lack of means to pay for star power.

Another implication is for labor economists who use American football to study concepts such as joint production, individual production, and racial discrimination. The research provides empirical evidence for all three of those concepts mentioned. Football, being a team sport, reflects the challenges of many business organizations in terms of its numerous interactions between teammates. Thus, these concepts can be applied to better understand the functioning and dynamics of the markets that influence wage labor.

Recommendations for Future Research

This research advanced the methods of determining wages for NFL offensive linemen. Future development of better tracking methods is needed to fully enumerate individual offensive lineman performance. At the individual level, a run-blocking metric would be beneficial. A publicly available plus/minus run-blocking metric, in which an individual offensive lineman would be graded a plus for a successful run block and a minus for an unsuccessful run block, would give the researcher an indicator of individual performance to evaluate an offensive lineman's run-blocking ability. To gather this data, game film would be evaluated by a person with strong familiarity of all of the running plays of the specific team, in particular, the individual goals of each offensive lineman during each run play. Additionally, controls that could better evaluate the talent level of the opponent as a run defender and pass defender would be beneficial. One way to do this would be to tally the number of individual matchups between an offensive lineman and particular pass rushers he faces during every game of the season. Doing this for all passing plays of the lineman with a tool to evaluate each pass rusher's ability (another metric) would enable us to gain a better understanding of the level of difficulty of individual opponents. Similarly, this could be done for all running plays in which the offensive lineman participates.

Building on enumerating individual offensive lineman performance to reach the goal of marginal production, one needs to be able to allocate individual contributions to joint production. Using a combination block on a running play as an example of joint production, an expert on offensive line play (such as an NFL offensive line coach) would need to break down the level of difficulty of both offensive lineman's responsibilities. Finally, one needs to be able to weight the significance of each individual's action as it relates to the outcome of the game. An expert would need to track and identify an objective value of each block in which an individual lineman participates. For example, a combination block between a left tackle and left guard might yield an outcome of 8 yards rushing by a running back. The left tackle, based on the difficulty of the combination block, its relative importance, and his success could be credited with 2 yards of the outcome. Using the marginal production function and a revenue function that addresses potential reverse causalities, one can estimate what a player is worth to a particular team in the open market. This method can be extended to all positions in the NFL such that you can compare which positions on an NFL team are most important in terms of winning.

Final Remarks

The analysis has shined an additional light on wage determination in sport and was one of the first studies to focus on the market of NFL offensive linemen. Wage determination models shed light on productivity and what is valued by organizations as a whole. Thirty-two NFL organizations vie against each other every season to assemble the most competitive roster in hopes of ultimately winning a Super Bowl. One of the pieces to building a roster is the offensive line, and part of the construction occurs during the free agency market. This study has revealed some of the factors that go into determining what an offensive lineman is paid in free agency. Many factors impact the wages of a player, such as age, the number of games started, and the proportion of individual knockdowns allowed.

Understanding the impact of these factors, NFL general managers have a multitool kit at their disposal for the evaluation of this position as they consider using free agency in roster construction. Further research would benefit general managers' goals of gaining more insight to estimate how much they should pay their employees and of possible market inefficiencies in the offensive line. These market inefficiencies can enable an organization to gain a competitive advantage. Extending this research beyond offensive linemen to every other position on a team would be highly beneficial for an NFL organization.

Additionally, economists have a launching point for future research on the labor components of a complex team sport. Because the population consists only of players in free agency, an economist could better evaluate the impact that performance has on wages. Further research might ultimately unlock the marginal revenue products of players and enable economists to look into possible racial discrimination.

Appendix A

Comparison of OBR Model and Selected Models from Research Questions

Table I

Comparison of OBR Model and Selected Models from Research Question #1

Variables	OBR Model	Research Question #1A	Research Question #1B	Research Question #1C
Individual knockdowns allowed proportion	**-6.364** (-2.10)**	**-0.206*** (-2.93)**		
Games started	**0.073*** (7.44)**		**0.556*** (11.68)**	
Individual proportion of plays with a penalty	**0.620 (0.09)**			**-0.223*** (-3.48)**
Team rushing yards	0.0007** (2.49)			
Team passing yards	-.0003** (-2.11)			
Points scored	0.004 (3.36)			
Adjusted line yards	0.196 (1.08)			
Run-to-pass ratio	-1.334** (-2.29)			
Non-White or Hispanic	-0.159* (-1.67)	-0.137** (-2.04)	-0.095* (-1.93)	-0.116* (-1.77)
Position guard	0.030 (0.23)	0.020 (0.21)		0.044 (0.48)
Position tackle	0.266* (1.75)	0.130 (1.21)		0.111 (1.14)
Same scheme	0.079 (0.90)	0.104 (1.65)	0.079* (1.71)	0.035 (0.56)
Age	-0.082*** (-4.36)	-0.206*** (-3.20)	-0.178*** (-3.78)	-0.187*** (-2.93)
Overall pick	-0.258*** (-3.88)	-0.232*** -3.63	-0.149*** (-3.09)	-0.271*** (-2.93)

Variables	OBR Model	Research Question #1A	Research Question #1B	Research Question #1C
Tackle (as dichotomous variable)			0.054 (1.09)	
Other knockdowns allowed proportion		-0.243*** (-3.77)		
Pass control				-0.123* (-1.95)
Run control				0.133** (2.09)
Observations	228	204	301	229
R^2	0.430	0.245	0.381	0.183

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1.

Table II

Comparison of OBR Model and Selected Models from Research Question #2

Variables	OBR Model	Research Question #2ab	Research Question #2c
Individual knockdowns allowed proportion	-6.364** (-2.10)		
Games started	0.073*** (7.44)		
Individual proportion of plays with a penalty	0.620 (0.09)		
Team rushing yards	**0.0007** (2.49)**	**0.286*** (4.16)**	
Team passing yards	**-.0003** (-2.11)**	**0.201*** (2.80)**	
Points scored	**0.004 (3.36)**		**0.316*** (4.72)**

Variables	OBR Model	Research Question #2ab	Research Question #2c
Adjusted line yards	0.196 (1.08)		
Run-to-pass ratio	-1.334** (-2.29)		
Non-White or Hispanic	-0.159* (-1.67)	-0.124* (-1.74)	-0.123* (-1.74)
Position guard	0.030 (0.23)	-0.016 (-0.16)	0.007 (0.08)
Position tackle	0.266* (1.75)	0.065 (0.63)	0.102 (1.00)
Same scheme	0.079 (0.90)	0.075 (1.12)	0.077 (1.16)
Age	-0.082*** (-4.36)	-0.257*** (-3.72)	-0.233*** (-3.41)
Overall pick	-0.258*** (-3.88)	-0.284*** (-4.05)	-0.276*** (-4.03)
Observations	228	180	180
R^2	.430	0.254	0.260

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1.

Table III

Comparison of OBR Model and Selected Model from Research Question #4

Variables	OBR Model	Research Question #4a
Individual knockdowns allowed proportion	-6.364** (-2.10)	
Games started	0.073*** (7.44)	
Individual proportion of plays with a penalty	0.620 (0.09)	
Team rushing yards	0.0007** (2.49)	
Team passing yards	-.0003** (-2.11)	
Points scored	0.004 (3.36)	

Variables	OBR Model	Research Question #4a
Adjusted line yards	**0.196 (1.08)**	**0.234*** (3.64)**
Run-to-pass ratio	-1.334** (-2.29)	
Non-White or Hispanic	-0.159* (-1.67)	-0.165** (-2.45)
Position guard	0.030 (0.23)	-0.021 (-0.22)
Position tackle	0.266* (1.75)	0.040 (0.41)
Same scheme	0.079 (0.90)	0.073 (1.14)
Age	-0.082*** (-4.36)	-0.190*** (-2.94)
Overall pick	-0.258*** (-3.88)	-0.279*** (-4.32)
Run control		0.110* (1.71)
Observations	228	208
R^2	.430	0.216

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1.

Table IV

Comparison of OBR Model and Selected Model from Research Question #5

Variables	OBR Model	Research Question #5b
Individual knockdowns allowed proportion	-6.364** (-2.10)	
Games started	0.073*** (7.44)	0.467*** (8.15)
Individual proportion of plays with a penalty	0.620 (0.09)	
Team rushing yards	0.0007** (2.49)	
Team passing yards	-.0003** (-2.11)	

Variables	OBR Model	Research Question #5b
Points scored	0.004 (3.36)	
Adjusted line yards	0.196 (1.08)	
Run-to-pass ratio	**-1.334**** **(-2.29)**	**0.079** **(1.44)**
Non-White or Hispanic	-0.159* (-1.67)	-0.083 (-1.47)
Position guard	0.030 (0.23)	
Position tackle	0.266* (1.75)	
Same scheme	0.079 (0.90)	
Age	(-0.082)*** (-4.36)	-0.255*** (-4.53)
Overall pick	-0.258*** (-3.88)	-0.177*** (-3.09)
Run Control		0.118** (2.10)
Pass control		-0.095* (-1.70)
Observations	228	230
R^2	.430	0.335

Standardized beta coefficients; t statistics in parentheses; *** p < .01, ** p < .05, * p < .1.

References

Alamar, B. C., & Weinstein-Gould, J. (2008). Isolating the effect of individual linemen on the passing game in the National Football League. *Journal of Quantitative Analysis in Sports*, *4*(2).

Andersen, T., & Croix, S. J. L. (1991). Customer racial discrimination in major league baseball. *Economic Inquiry*, *29*(4), 665-677.

Barker, M., Wyatt, T. J., Johnson, R. L., Stone, M. H., O'bryant, H. S., Poe, C., & Kent, M. (1993). Performance factors, psychological assessment, physical characteristics, and football playing ability. *The Journal of Strength & Conditioning Research*, *7*(4), 224-233.

Baumer, B. S., Jensen, S. T., & Matthews, G. J. (2015). openWAR: An open source system for evaluating overall player performance in Major League Baseball. *Journal of Quantitative Analysis in Sports*, *11*(2), 69-84.

Bellemore, F. A. (2001). Racial and ethnic employment discrimination promotion in major league baseball. *Journal of Sports Economics*, *2*(4), 356-368.

Bent, D. P., & Sommers, P. M. (2014). Pay and performance in the National Hockey League, 2011-2012. *Atlantic Economic Journal*, *42*(1), 109.

Becker, G. S. (1971). *The Economics of Discrimination 2nd ed*. (University of Chicago Press, Chicago).

Berri, D. J. (1999). Who is 'most valuable'? Measuring the player's production of wins in the National Basketball Association. *Managerial and Decision Economics, 20*(8), 411-427.

Berri, D. J. (2007). Back to back evaluations on the gridiron. *Statistical thinking in sports*, 241-262.

Berri, D. J. (2008). A simple measure of worker productivity in the National Basketball Association. *The business of sport, 3*, 1-40.

Berri, D. J., Brook, S. L., & Schmidt, M. B. (2007). Does one simply need to score to score? *International Journal of Sport Finance*, *2*(4), 190-205.

Berri, D. J., & Simmons, R. (2009). Race and the evaluation of signal callers in the National Football League. *Journal of Sports Economics, 10*(1), 23-43.

Berri, D. J., & Simmons, R. (2011). Catching a draft: On the process of selecting quarterbacks in the National Football League amateur draft. *Journal of Productivity Analysis*, *35*(1), 37-49.

Blass, A. A. (1992). Does the baseball labor market contradict the human capital model of investment? *The Review of Economics and Statistics*, 261-268.

Bradbury, J. C. (2008). Statistical performance analysis in sport. *The business of sport, 3*, 41-56.

Brophy, R. H., Chehab, E. L., Barnes, R. P., Lyman, S., Rodeo, S. A., & Warren, R. F. (2008). Predictive value of orthopedic evaluation and injury history at the NFL combine. *Medicine & Science in Sports & Exercise*, *40*(8), 1368-1372.

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.

Bryson, A., Frick, B., & Simmons, R. (2013). The returns to scarce talent: footedness and player remuneration in European soccer. *Journal of Sports Economics*, *14*(6), 606-628.

Bryson, A., Rossi, G., & Simmons, R. (2014). The migrant wage premium in professional football: A superstar effect? *Kyklos*, *67*(1), 12-28.

Burdekin, R. C., Hossfeld, R. T., & Smith, J. K. (2005). Are NBA fans becoming indifferent to race? Evidence from the 1990s. *Journal of sports economics*, *6*(2), 144-159.

Burnett, N. J., & Van Scyoc, L. J. (2015). Compensation discrimination for defensive players: applying quantile regression to the National Football League market for linebackers and offensive linemen. *Journal of Sports Economics*, *16*(4), 375-389.

Carroll, B., Palmer, P., & Thorn, J. (1989). *The hidden game of football*. Warner books.

Carroll, B. N., Palmer, P., Thorn, J., & Pietrusza, D. (1998). *The hidden game of football: The next edition*. Total Sports.

Cover, T. M., & Keilers, C. W. (1977). An offensive earned-run average for baseball. *Operations Research*, *25*(5), 729-740.

Daniel, M. L., Brown, B. S., & Gorman, D. (1984). Strength and anthropometric characteristics of selected offensive and defensive university-level football players. *Perceptual and motor skills*, *59*(1), 127-130.

Depken, C. A., & Ford, J. M. (2006). Customer-based discrimination against major league baseball players: Additional evidence from All-star ballots. *The Journal of Socio-Economics*, *35*(6), 1061-1077.

Dhar, A. (2011). *Drafting NFL wide receivers: Hit or miss*? Technical report. Department of Statistics, University of California at Berkeley.

Doran, J., & Doran, D. R. (2004). *Inequality in pay: A study of wage disparity in the NFL.* Technical report. Department of Finance, Florida State University.

Fizel, J. (2006). *Handbook of sports economics research*. ME Sharpe.

Futterman, Matthew (2017, February 28). The 2017 Draft: Secrets of the NFL Combine. *Wall Street Journal* pp. A.14.

Gandelman, N. (2009). Selection biases in sports markets. *Journal of Sports Economics*, *10*(5), 502-521.

Gerrard, B. (2007). Is the Moneyball approach transferable to complex invasion team sports? *International Journal of Sport Finance, 2*(4), 214.

Gius, M., & 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.

Groothuis, P. A., & Hill, J. R. (2004). Exit discrimination in the NBA: A duration analysis of career length. *Economic Inquiry*, *42*(2), 341-349.

Hakes, J. K., & Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis. *The Journal of Economic Perspectives, 20*(3), 173-185.

Hamilton, B. H. (1997). Racial discrimination and professional basketball salaries in the 1990s. *Applied Economics*, *29*(3), 287-296.

Hartman, M. (2011). Competitive performance compared to combine performance as a valid predictor of NFL draft status. *The Journal of Strength & Conditioning Research*, *25*, S105-S106.

Haupert, M. J. (2009). Player pay and productivity in the reserve clause and collusion eras. *Nine: A Journal of Baseball History and Culture*, *18*(1), 63-85.

Idson, T. L., & Kahane, L. H. (2000). Team effects on compensation: an application to salary determination in the National Hockey League. *Economic Inquiry*, *38*(2), 345-357.

Jenkins, J. A. (1996). A reexamination of salary discrimination in professional basketball. *Social Science Quarterly*, *77,* 594-608.

Jewell, R. T., Brown, R. W., & Miles, S. E. (2002). Measuring discrimination in major league baseball: evidence from the baseball hall of fame. *Applied Economics*, *34*(2), 167-177.

Jones, J. C. H., Nadeau, S., & Walsh, W. D. (1999). Ethnicity, productivity and salary: player compensation and discrimination in the National Hockey League. *Applied Economics*, *31*(5), 593-608.

Jones, J. C., & Walsh, W. D. (1987). The World Hockey Association and player exploitation in the National Hockey League. *Quarterly Review of Economics and Business*, *27*(2), 87-101.

Jones, J. C., & Walsh, W. D. (1988). Salary determination in the National Hockey League: the effects of skills, franchise characteristics, and discrimination. *Industrial & Labor Relations Review*, *41*(4), 592-604.

Kahn, L. M. (1992). The effects of race on professional football players' compensation. *Industrial & Labor Relations Review*, *45*(2), 295-310.

Kahn, L. M., & Shah, M. (2005). Race, compensation and contract length in the NBA: 2001-2002. *Industrial Relations: A Journal of Economy and Society*, *44*(3), 444-462.

Kahn, L. M., & Sherer, P. D. (1988). Racial differences in professional basketball players' compensation. *Journal of Labor Economics*, *6(1),* 40-61.

Kanazawa, M. T., & Funk, J. P. (2001). Racial discrimination in professional basketball: Evidence from Nielsen ratings. *Economic Inquiry*, *39*(4), 599-608.

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.

Kuzmits, F. E., & Adams, A. J. (2008). The NFL combine: does it predict performance in the National Football League? *The Journal of strength & conditioning research*, *22*(6), 1721-1727.

Lewis, M. (2004). *Moneyball: The art of winning an unfair game.* New York: W.W. Norton.

Lewis, M. (2007). *The blind side: Evolution of a game*. WW Norton & Company.

Lindsey, G. R. (1959). Statistical Data Useful for the Operation of a Baseball Team. *Operations Research*, *7*(2), 197-207.

Lindsey, G. R. (1963). An investigation of strategies in baseball. *Operations Research*, *11*(4), 477-501.

McGee, K. J., & Burkett, L. N. (2003). The National Football League Combine: A Reliable Predictor of Draft Status? *The Journal of Strength & Conditioning Research*, *17*(1), 6-11.

Medoff, M. H. (1976). Monopsonistic exploitation in professional baseball. *Quarterly Review of Economics and business*, *16*(2), 113-121.

Michael, J. (2014). *The game before the money: Voices of the men who built the NFL*. U of Nebraska Press.

Mincer, J. (1958). Investment in human capital and personal income distribution. *The Journal of Political Economy*, *66(4),* 281-302.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* New York: Columbia University Press.

Mogull, R. G. (1973). Football salaries and race: some empirical evidence. *Industrial Relations: A Journal of Economy and Society*, *12*(1), 109-112.

Mogull, R. G. (1981). Salary discrimination in professional sports. *Atlantic Economic Journal*, *9*(3), 106-110.

Mulholland, J., & Jensen, S. T. (2014). Predicting the draft and career success of tight ends in the National Football League. *Journal of Quantitative Analysis in Sports*, *10*(4), 381-396.

New York Times (1982). Average NFL salary is $90,102, survey says. Retrieved from http://www.nytimes.com/1982/01/29/sports/average-nfl-salary-is-90102-survey-says.html

Parekh, S., & Patel, A. (2017). The NFL Combine as a Predictor of On-field Success. *Foot & Ankle Orthopaedics*, *2*(3), 2473011417S000315.

Pasteur, R. D., & Cunningham-Rhoads, K. (2014). An expectation-based metric for NFL field goal kickers. *Journal of Quantitative Analysis in Sports*, *10*(1), 49-66.

Quirk, J., & Fort, R. D. (1997). *Pay dirt: The business of professional team sports*. Princeton University Press.

Robbins, D. W. (2010). The National Football League (NFL) combine: does normalized data better predict performance in the NFL draft? *The Journal of Strength & Conditioning Research*, *24*(11), 2888-2899.

Robbins, D. W. (2011). Positional physical characteristics of players drafted into the National Football League. *The Journal of Strength & Conditioning Research*, *25*(10), 2661-2667.

Robst, J., VanGilder, J., Coates, C. E., & Berri, D. J. (2011). Skin tone and wages: Evidence from NBA free agents. *Journal of Sports Economics*, *12*(2), 143-156.

Salfino, M. (2015, February 19). The Count: Great Moments in NFL Combine Failure. *Wall Street Journal* pp. D.6.

Santo, C., & Mildner, G. (2010). *Sport and public policy: Social, political, and economic perspectives*. Human Kinetics.

Sawyer, D. T., Ostarello, J. Z., Suess, E. A., & Dempsey, M. (2002). Relationship between football playing ability and selected performance measures. *Journal of Strength and Conditioning Research*, *16*(4), 611-616.

Schollaert, P. T., & Smith, D. H. (1987). Team racial composition and sports attendance. *The Sociological Quarterly*, *28*(1), 71-87.

Scott, F. A., Long, J. E., & Somppi, K. (1983). Free agency, owner incentives, and the National Football League Players Association. *Journal of Labor Research*, *4*(3), 257-264.

Scully, G. W. (1973). Economic discrimination in professional sports. *Law and Contemporary Problems*, *38*(1), 67-84.

Scully, G. W. (1974a). Pay and performance in Major League Baseball. *The American Economic Review*, *64*(6), 915-930.

Scully, G. W. (1974b) “The Economics of Discrimination in Professional Sports: The Case of Baseball,” in Noll, R. G. (Ed.). (1974). *Government and the sports business:*

Papers prepared for a conference of experts, with an introduction and summary (No. 8). Brookings Institution.

Sierer, S. P., Battaglini, C. L., Mihalik, J. P., Shields, E. W., & Tomasini, N. T. (2008). The National Football League Combine: Performance differences between drafted and nondrafted players entering the 2004 and 2005 drafts. *The Journal of Strength & Conditioning Research*, *22*(1), 6-12.

Simmons, R., & Berri, D. J. (2009). Gains from specialization and free agency: the story from the gridiron. *Review of Industrial Organization, 34*(1), 81-98.

Sommers, P. M., & Quinton, N. (1982). Pay and performance in Major League Baseball: The case of the first family of free agents. *The Journal of Human Resources*, *17*(3), 426-436.

Somppi, K. (1985). Salary vs. marginal revenue product under monopsony and competition: The case of professional basketball. *Atlantic Economic Journal*, *13*(3), 50-59.

Stimel, D. (2009). A statistical analysis of NFL quarterback rating variables. *Journal of Quantitative Analysis in Sports*, *5*(2). doi:10.2202/1559-0410.1034

Tainsky, S., & Winfree, J. A. (2010). Discrimination and demand: The effect of international players on attendance in Major League Baseball. *Social Science Quarterly*, *91*(1), 117-128.

von Dohlen, P. (2011). Tweaking the NFL's quarterback passer rating for better results. *Journal of Quantitative Analysis in Sports*, *7*(3).

Weimar, D., & Wicker, P. (2017). Moneyball revisited: Effort and team performance in professional soccer. *Journal of Sports Economics*, *18*(2), 140-161.

Wolfson, J., Addona, V., & Schmicker, R. H. (2011). The quarterback prediction problem: Forecasting the performance of college quarterbacks selected in the NFL draft. *Journal of Quantitative Analysis in Sports*, *7*(3).

Wicker, P., Prinz, J., Weimar, D., Deutscher, C., & Upmann, T. (2013). No pain, no gain? Effort and productivity in professional soccer. *International Journal of Sport Finance*, *8*(2), 124.

Yurko, R., Ventura, S., & Horowitz, M. (2018). nflwar: A reproducible method for offensive player evaluation in football (extended edition). Technical report. Department of Statistics, Carnegie Mellon University.