5-2020

An empirical analysis on major league soccer player earnings

Aaron Michael Anderson Iehl
University of Northern Iowa

AN EMPIRICAL ANALYSIS ON
MAJOR LEAGUE SOCCER PLAYER EARNINGS

A Thesis Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors

Aaron Michael Anderson Iehl

University of Northern Iowa

May 2020

This Study by: Aaron Michael Anderson Iehl

Entitled: An Empirical Analysis On Major League Soccer Player Earnings

has been approved as meeting the thesis or project requirement for the Designation

University Honors

________ ____

Date Dr. Imam Alam, Professor of Economics, Honors Thesis Advisor

________ ____

Date Dr. Jessica Moon, Director, University Honors Program

Abstract

Application of economic and statistical techniques to the sports world has been on the rise for some time. The major sports in the United States (football, basketball, baseball, and hockey) have been thoroughly studied. However, professional soccer in the United States has emerged as a strong sports market in the last 20 years. This paper applies a statistical technique used by Lucifora and Simmons (2003) on an Italian professional soccer league to examine determinants of professional soccer player wages and assess the overall value added. The model is applied to Major League Soccer to determine how the different league structure impacts the earnings of its players. Empirical study of MLS has shown that league structure impacts how teams assess a player's value to their team through their salary, which is shown to be determined by player age and its square, passing metrics, career goal scoring record, designated player status, US men's national team experience, and whether the player is a foreign-born player.

Table of Contents:

Introduction

Sports are an essential part of life in many ways. Sometimes it requires the absence of something to realize the desirability of that very thing. The Coronavirus Pandemic of 2019-2020 has stripped the world of its access to sports as a means of entertainment. Not only has the Coronavirus dealt a blow to culture as we know it, but a major economic staple has been ripped out as well. In the United States, this is most evident through the suspensions of the traditional kings of sports media markets such as the National Basketball Association (NBA), college basketball's "March Madness," and Major League Baseball (MLB). However, another significant sports league has been put on hold. Major League Soccer (MLS) is an emerging league in a ripe soccer market. A large part of the growth that MLS has experienced came from attracting better players to play in the league. Further, it has retained the good players that currently compete in the league who previously would leave for more competitive leagues. In order to effectively attract and retain talent, sufficient wages must be present to meet the salary demands of the players. However, what determines how much a player is worth to a team? Why should some players earn more than others? These questions are investigated in this study to find the determinants of MLS players' individual salaries. It is found that league structure leads to different salary structure within different leagues. Experience, performance, and reputation determines MLS player salaries.

Background

Worldwide the sports industry is commonly thought to generate $500 billion annually (SportyCo 2017). Within the global sphere, the United States is a massive sports market. However, there has consistently been four dominant professional leagues in the U.S. sports market along with the biggest college sports. Between the NFL, NBA, MLB, NHL, collegiate football and collegiate men's basketball, the majority of revenues are heavily concentrated in these sports and leagues. This was true during the early days when the concept of MLS was still a dream.

MLS is not the first attempt at establishing soccer in the United States. Multiple previous attempts to form a domestic league started but subsequently failed. There simply was not the consumer demand to sustain the expenses incurred. The North American Soccer League (NASL) ran from 1968-1985. Arguably the world's best soccer player ever, Pele, joined the New York Cosmos in the league, but even that was not enough to salvage the league.

As with many successful business ventures, timing is everything. The work of the United States Soccer Federation to secure the 1994 FIFA World Cup in the United States was the perfect storm to attempt to launch a new version of an American based domestic soccer league. The increased attention that came with hosting this major tournament enabled the formation of ten teams planning to begin play in 1996. Although the exposure gained through the World Cup was very helpful in helping to garner enough popularity and financial support, there would need to be something different about the structure of the new league to make it viable in the American sports culture.

At the advent of the league, there were four major steps that the MLS took to set their league up for viability (Delgado 1997). They identified that specific marketing to minority populations in cities around the US would help to create and establish genuine fan bases of the teams. Their focus on hiring players deviated from the style of the old NASL. The NASL preferred superstars from Europe that were nearing the end of their careers, and thus were somewhat "broken down". Older European players along with other early or mid-career English players from lower divisions were the NASL's primary focus. However, the MLS targeted younger and lesser known players, often from Latin American nations, to produce a more appealing and marketable product (Delgado 1997). There was also a much stronger focus to secure strong corporate partners and to establish profitable licensing deals with key corporations such as Adidas, Nike, Pepsi, and Honda (Delgado 1997). Finally, the ownership structure of the MLS was different and unique to each of the other major American sports leagues. The MLS is under a single entity ownership structure. Each entity is owned by the league and the franchises are then operated individually by the league's investors in each franchise. This is very different from not only the other American sports, but nearly every other notable domestic soccer league in the world (Delgado 1997).

The MLS got off the ground in 1996, but the first decade of operation was anything but stable. After a phenomenal inaugural season, attendance and viewership dropped significantly. From the start of the league through 2004 the MLS had lost over $350 million (Holmes 2004). In 2001 there had been two franchises that had folded and prospects for survival were grim. The saving grace for the MLS has been success for the United States in World Cups. In 2002 the United States made a remarkable and very surprising run all the way to the quarterfinals of the tournament (their best ever finish to date). This again provided a spike in popularity of the sport. Following the tournament, the MLS Cup of 2002 had over 61,000 fans in attendance. This was the largest attended MLS game in league history and held that status until 2018.

Leading up to the 2006 World Cup there had been a number of expansion teams around the country. Along with this, some of the marquee American players were transferring to big market European teams. This was a good sign in that America was beginning to produce better talent. However, it was struggling to retain the talent when the wages for players were much higher in European leagues. Apart from larger markets for soccer and more established teams, there was another contributing factor to lower wages for players in the MLS. Due to the single entity ownership structure, all salaries are centrally paid out by the league. This has allowed the league to set salary caps on how much each team will have to grant to each of their players. This helps the league limit its expenses on players and also theoretically promotes parity in the league as all teams have equal resources to acquire player talent (Coates, Frick, and Jewell 2016). The drawback to this is the limitation to invest a significant level of allocated capital to some of the world's best players to acquire world class talent. If a significant proportion of your salary goes to a single player, then the rest of the team's players will have less money available, thus the other players will be relatively less talented. This salary structure provides a significant barrier and disincentive to acquiring world class talent. This fed into the commonly stated narrative that the MLS is a low-class league of lesser talent on the international soccer stage. For many soccer fans, they would be more interested in following a European team solely on television with greater quality of play than following a domestic team in the United States with less talented players. The league saw this was a major limitation they would face to becoming a more respected and accepted league in the international soccer community.

To address this issue MLS successfully negotiated a blockbuster move that changed the landscape of the entire league. In 2007 the MLS instituted the Designated Player (DP) rule. This allows teams to exclude three players from their team's standard salary cap restrictions. Therefore, up to three players could be paid exceedingly more than previously possible, and thus the ability to attract better talent was granted. The MLS hoped that they would be able to attract some marquee star power, likely in a skilled European player. David Beckham made their aspirations come true in July 2007 when he signed with the Los Angeles Galaxy (Shapiro, DeSchriver, and Rascher 2014). Beckham was an English midfield player who had starred for some of Europe's biggest clubs and was also a star on the international stage representing England. As a free kick specialist, he held a persona and a collection of skills that added flavor to the MLS that it had not previously had. Beckham's presence brought an immediate surge to the American soccer market and the following of MLS. Leaguewide jersey sales increased 700% and Beckham's club, the LA Galaxy, had their jersey sales rise 5,210%. Online activity rose 117% within the first month of his appearance in the league. In Los Angeles, an additional 11,000 season tickets were sold for fans to watch Beckham play at Galaxy matches (Wahl 2010).

Designated players have continued to be the defining factor of the league for well over a decade. The league has been able to acquire talent and icons that likely would not have been possible without such an exception. There are a number of examples including players coming off of playing on the world stage at the World Cup like David Villa, Andrea Pirlo. The league has included a Ballon d'or winner, awarded to the world's best player for a given year, in Brazil's Kaka. Lastly, one of the biggest soccer personalities the world has ever seen in Zlatan Ibrahimović entered the league with a stunning goal from 40 yards out in his first ever appearance. MLS gained recognition never before possible through the DP rule, but they also used the remaining aspects of the salary cap system to maintain a competitive balance in the league. The LA Galaxy has been the most successful team, but it is certainly no monopoly of success like in the German Bundesliga where Bayern Munich has won seven straight league titles, nor is it an oligopoly like seen in England's Premier League where only 5 different teams have been crowned champions in the last 25 years (Houmes 2017). The MLS has worked hard to achieve two simultaneous goals of retaining the league's competitive parity and attracting better talent for the league, its teams, and the fans.

Literature Review

To understand earnings of individual players it is important to consider how the employers' wage distribution and structure impacts those wages. There has been extensive analysis on professional sports salaries due to the detailed information available. In particular, there are conflicting findings on whether increased salary inequality leads to greater worker effort and increased productivity, or whether the increased inequality is detrimental to the organization. Lazear and Rosen (1981) find evidence that salary inequality does in fact lead to greater levels of effort and productivity. However, Levine (1991) uses cohesion theory to explain findings showing that leveling the playing field among coworkers increases unity. Sports offer a unique arena to analyze the relationships between salaries and productivity, particularly due to the varying wage structures displayed in the various sports and leagues. Some leagues such as the MLB have been thoroughly studied to find that wage inequality is negatively related to team success. Whereas the NBA shows no significant relationship between success and salary dispersion (Berri and Jewell 2004). Soccer has been studied in a similar fashion to determine whether wage dispersion impacted team success. A U-shaped relationship is found in analysis of the German Bundesliga, which implies that teams with very low inequality have lower team success and productivity (Franck and Neusch 2011). Coates, Frick, and Jewell (2012) use the Gini Coefficient to conclude that a negative relationship between salary inequality and production exists for MLS teams. Under the single-entity ownership structure, salary equality is aided by the salary cap system. All but three players are under a league constant salary cap. Therefore, teams are more likely to be equally productive as there is less wage inequality. Since each team has an equal allotment of designated players to be excluded from the salary cap, the league should be relatively balanced in productivity.

The previous studies outline how the unique ownership structure and salary cap system of MLS impacts team performance. At an individual player level, what are the determinants of MLS player salary? Many studies have been conducted by a variety of methods to determine professional soccer player salaries. The majority of this analysis has focused on European leagues and players. The studies are split into two main categories where one group examines team performance as the dependent variable and the other group examines individual player salary as the dependent variable.

Szymanski and Kuypers (1999) conclude that club expenditures have a positive impact on team success within the English Premier League. In a similar manner, Forrest and Simmons (2002) find that a strong positive relationship exists for team salary and performance. Their conclusion applies to the Italian and English leagues but are much less significant for the German League. Due to the very different structure of MLS as compared to the other European leagues, it is difficult to apply similar methods to MLS. MLS has a salary cap system, a single-entity ownership structure, and does not involve a promotion or relegation system. For these reasons, this method was not used to study the MLS players.

The second method examined the determinants of individual player salary. Again, multiple European analyses have been performed to determine that there is a significant relationship between player's earnings and player performance (Torgler and Schmidt 2002). Each of these has taken a unique approach to study different team-based or individual factors in determining player salary. One study in particular by Lucifora and Simmons (2003) uses an ordinary least squares regression to examine earnings of Italian players. By using a cross-sectional sample of data including all players from the top two divisions of Italian soccer (Serie A and Serie B) for the 1994-1995 season, they conclude that a "Superstar Effect" is present which leads to a wage premium for those with very high goal scoring rates over their career. Due to the model including many variables that would be available for MLS players, this Lucifora and Simmons (2003) model is used to construct the model in this analysis of MLS players. Some modifications are made and discussed in the model section of this paper. The different structures of the leagues restrict coefficients from being directly compared, however the Lucifora and Simmons (2003) model is a good model that can be used with more modern and expansive data.

Data

As the popularity of MLS has increased, so has the availability of data on MLS players and franchises. Similar to other sports, the collection of data down to the most granular of metrics is analyzed and stored. In a sport where the slightest edge is all it takes to win; the advantage could be found in these detailed data metrics.

Four data categories are used in the paper examining Italian players' salaries (Lucifora and Simmons 2003). These include measures of experience, performance, reputation, and team qualities. All of these categories are examined to see if they are able to explain the determinants of MLS player salaries.

Table 1: Variables and Descriptive Statistics (541) Observations			
Variable	**Description**	**Mean**	**Std Deviation**
ln(Salary)	Current Annualized Base Salary preseason values 2019. Expressed in log values	12.25	1.04
Experience			
Age	Age (in years)	25.32	4.32
AgeSQ	Quadratic term in age (in years)	659.91	223.63
Apps	MLS appearances during 2018 season	18.24	10.24
AppsSQ	Squared term of Apps	437.48	371.58
PreviousApps	Accumulated MLS appearances prior to 2018 season	49.09	71.27
PreviousAppsSQ	Squared term of PreviousApps	7480.30	17939.32
Performance			
F	Forward*	0.301	
M	Midfielder*	0.311	
D	Defender*	0.388	
ForGlsP90	Goals per 90 minutes in MLS by forward in 2018	0.090	0.20
MidGlsP90	Goals per 90 minutes in MLS by midfielder in 2018	0.034	0.11
DefGlsP90	Goals per 90 minutes in MLS by defender in 2018	0.013	0.05
ForAstP90	Assists per 90 minutes in MLS by forward in 2018	0.044	0.13
MidAstP90	Assists per 90 minutes in MLS by midfielder in 2018	0.031	0.09
DefAstP90	Assists per 90 minutes in MLS by defender in 2018	0.025	0.16
SuccDribP90	Successful Dribbles completed per 90 minutes in MLS in 2018	1.01	1.09
ForPasses	Passes completed by forwards in MLS in 2018	81.11	180.04
MidPasses	Passes completed by midfielders in MLS in 2018	196.98	424.21
DefPasses	Passes completed by defenders in MLS in 2018	218.59	392.71
MidTklIntP90	Defensive tackles and interceptions by midfielders per 90 minutes in MLS in 2018	0.842	1.50
DefTklIntP90	Defensive tackles and interceptions by defenders per 90 minutes in MLS in 2018	1.213	1.70
CrdYP90	Yellow cards received per 90 minutes in MLS in 2018	0.183	0.26
CrdRP90	Red cards received per 90 minutes in MLS in 2018	0.014	0.05
TwoCrdYP90	Second yellow cards received per 90 minutes in MLS in 2018	0.005	0.03
PK	Penalty Kick goals scored in MLS in 2018	0.198	0.76
PKatt	Penalty Kick attempts in MLS in 2018	0.251	0.90
ForStrikeRate	Career goals per game by forward in MLS	0.064	0.14
MidStrikeRate	Career goals per game by midfielder in MLS	0.025	0.07
DefStrikeRate	Career goals per game by defender in MLS	0.012	0.03
ForStrikeRateSQ	Square of ForStrikeRate	0.024	0.09
MidStrikeRateSQ	Square of MidStrikeRate	0.005	0.02
DefStrikeRateSQ	Square of DefStrikeRate	0.001	0.01
Reputation			
Superstar1	More than 0.40 goals per match in MLS*	0.031	
Superstar2	0.25 to 0.40 goals per match in MLS*	0.059	
DP	Designated Player*	0.098	
Domestic	Domestic, or US Citizen player*	0.470	
USAIntl	US National Team player in career*	0.107	
USAcaps	Career appearances with US National Team	2.37	13.13
Team Qualities			
LogAttendance	Log of home attendance in 2018	9.94	0.31
ln(LagFranchiseRev)	Franchise revenue in 2017	2.14E+07	6473252.00
ln(FranchiseRev)	Franchise revenue in 2018	2.45E+07	7483622.00
Pts	Season point total (Win = 3, Draw = 1, Loss = 0)	47.07	12.86
GD	Goal Differential, Goals Scored less Goals Against)	-0.366	16.56
Sources: Sports-Reference.com, MLSsoccer.com, TransferMarkt.us			
*** Dummy variable indicating relative proportion in the sample.**			

Descriptive statistics for the variables are shown in Table 1. The sample includes 541 outfield players. Players are 25 years old on average and have an average of 49 previous appearances in MLS detailing their experience as MLS professionals. The mean salary for the league is \$208,981 whereas the median is \$180,000 showing a right skew to the distribution of salary. There is quite the range of salaries in the sample. Salaries start as low as the league minimum of \$54,500 for a number of low-quality players and extend all the way up to former World Cup champion, Bastian Schweinsteiger's, base salary of \$6,100,000. In Table 1, player salaries are expressed as the natural log of salaries.

To gather the necessary data for each of these data categories, a number of sources were necessary. One helpful aspect of the MLS being a single-entity system is that large amounts of data are centrally stored. The main website for the MLS, *mlssoccer.com*, had a wealth of data that accounted for a number of the variables in the four categories. In particular, the MLS website provided information about each player's career and history in the league, their in-game performance measure such as goals and assists, along with some of the characteristics about each player's team. As the official site of MLS, the data is both secure and reliable.

It is important to ensure that the data for the model is reliable and accurate. To do this, the data from the MLS website was cross referenced with data from a database titled *FBref.com*. This was where the most intricate and detailed information for every player in the MLS was included. Further, it matched all of MLS's official statistics for each player and included a host of additional figures. This site was particularly used for the detailed performance statistics of the players. This site launched in 2018, but has data extending much further back. This group maintains similar data bases with much longer history relating to statistics for the NFL, MLB, and NBA. Matt Gelb (2020) notes how this company has recorded over 1 billion page views across their websites since they have launched their numerous sports reference data bases. This shows that the site is frequented and has been used by a number of researchers and other statisticians in the sports world. They have not had any significant data issues reported to date, which speaks to the integrity the source carries.

The salaries for the individual players were obtained from a press release from the MLS Players Association. The full guide was accessed from the labor union’s website and included each player's base salary and the annualized guaranteed compensation for all MLS players during the 2018 season. The subsample of data used for this study excludes goalkeepers. Their position is very different, and performance is based on very different metrics.

The last source of the data contributed to the team qualities category. The average revenue for the franchises within the single-entity system were acquired from an international soccer data source known as “TransferMarkt”. The website works to collect and analyze the transfer fees exchanged between clubs for the rights to sign players to a contract with their new team. They also had data on the total revenues of each MLS team showing how much money each team brought in as revenue by year. This hopefully would help to show the interest that the team was gathering from the audience around them.

While the sources of data are accurate and reliable, some of the data assumptions have limited the full picture of MLS players. For example, the data set used only includes players that appeared in matches during 2018. Therefore 3 types of players are potentially omitted. First is players who are young and developing. They may be under contract but require improvement before they appear in any matches. Next are retired players who may have a contract guarantee or bonuses that are to be paid out during the 2018 season. However, since they are not participating in any matches, they are not deemed current players. Finally, injured players are omitted from the regression. It was not possible to objectively select some injured players to include and others to omit because the type of injury and corresponding injuries are unknown. Some players could have career ending injuries, whereas others could continue to be dominant forces of the league. One example is Seattle Sounders FC forward Jordan Morris who missed the entire 2018 season but has been one of the league's most impactful players.

Model

An ordinary least squares (OLS) regression is used in this analysis to study the effects of the experience, performance, reputation, and team qualities variables on the earnings of MLS players in 2018. This follows the method used by Lucifora and Simmons (2003) in analyzing the earnings of Italian players in the 1994-1995 season. While a similar method is used there are some deviations in the updated model from the original. The deviations will be more thoroughly explained later in the discussion, but many are the result of increased detail of available data measured on the current league. Some original measures are not available in the current league in a similar regard and are thus omitted. Further, it is important to note that coefficients should be directly compared as the structure of the soccer leagues are very different for three reasons. First, the MLS is a single-entity ownership structure where Italy's Serie A consists of independent organizations. Secondly, league competition differs from year to year since there is a promotion and relegation system in the Italian Federation. Lastly, the system for salary distribution faces different limits in the MLS compared to Serie A. Serie A had no sort of salary restriction in 1994-1995, whereas the MLS uses a soft cap system with 3 DPs that are not included in the salary cap.

The dependent variable is the log wages (lnSalary) of individual MLS players. The annualized base salary is used to reduce the complexity of the model by excluding things such as signing bonuses. The model (Model 1) is then regressed with four separate vectors of independent variables corresponding to the aforementioned data categories of experience, performance, reputation, and team qualities.

Model 1:

$$\ln(Salary)_i = \alpha_0 + \alpha_1 EXPERIENCE_i + \alpha_2 PERFORMANCE_i + \alpha_3 REPUTATION_i + \alpha_4 TEAMQUALITIES_i + \varepsilon_i$$

The experience category is broken down to include a player's age (Age) and their age squared (AgeSQ). It is expected that wages would increase with age at a decreasing rate. This is similar to the expectation of wages of nearly all other industries. However, as the age increases and players near retirement, they cease to accumulate additional earnings and new contracts. In fact, their salaries may even begin to decrease as they are less productive to the team. Therefore, it is expected that age would have a positive coefficient and the square of age would be negative.

Also, in the experience category are the number of appearances in matches for the player during the 2018 season (Apps) and the square of that number of appearances in matches (AppsSQ). The more that a player appears for the team, the more their value is revealed. Similarly, the number of career appearances (PreviousApps) for a player shows how valuable they have been to teams over the long term. Both the appearances in 2018 and the career appearances prior to the 2018 season are expected to have a positive sign. The square of each of the 2018 appearances and previous career appearances (PreviousAppsSQ) are included as well. Once players near the maximum number of possible games during the season there is less variance in salary from players who participate in all matches compared to those who may miss only one to two matches due to rest, injury, or tactics. Therefore, a negative coefficient for the AppsSQ and PreviousAppsSQ is expected.

The next category of explanatory variables is the performance category. These are used to assess a player's impact on the game based upon their individual contributions to the team. The more they contribute and are integral to team success, the more they are expected to earn. This section closely follows the category used by Lucifora and Simmons (2003). However, as previously mentioned, it is slightly expanded to include additional measures that were not available while they conducted their research.

While a lot goes into winning a soccer match, at the final whistle, only one measure determines whether the result is a win, loss, or draw. Simply stated, the team with more goals wins. Goals are most assuredly the most important variable to winning games, but are they the most important in determining salary? The number of goals scored per ninety minutes played during the 2018 season is included in the performance category. By looking at a per minutes basis it looks at the productivity over the time that the player is able to impact the match. Some players may enter as substitutes frequently, but if they can score in their limited minute allocation, they can be very valuable players to the team. Further, the goals per ninety minutes is separated out by the three major position groups of forwards, midfielders, and defenders (ForGlsP90, MidGlsP90, and DefGlsP90). It is better to compare how people are performing in terms of goal scoring relative to their positional peers. Each position has varied roles and responsibilities on the field. Forwards are more often scoring goals for their teams. In different situations of the game, all positions do have the opportunities to score goals, and therefore each position's goal scoring rate is included in the model. For each position however, it is expected that the variables coefficient will be positive indicating that scoring more goals increases a player value to the team, and thus earnings.

In a similar way, it is expected that assists per ninety minutes would also have a positive coefficient. Assists are recorded when a player makes the final pass to another player who soon after scores a goal. This helps to measure players creativity, facilitation, and enabling other players to score goals. Assist variables are also split by position (ForAstP90, MidAstP90, and DefAstP90) to account for the fact that it is more likely for forwards and midfielders to generate assists as they spend more time in the attacking third of the field. Defenders, however, may commonly make long passes that lead to goals, so they are included in this equation.

An addition to the model used for developing the current model is the inclusion of dribbling and passing statistics. Dribbling includes the number of successful dribbling runs made by a player per ninety minutes (SuccDribP90). Better players will be able to complete more of these in a match. Therefore, it is expected that the coefficient for this related to earnings would be positive. Passing is another crucial element to the game. Each position is needed in the passing game. However, it is common for different positions to have different types of passing strategies. Forwards are often playing with more opponents surrounding them, and thus need to make quick short passes. Midfielders are sometimes in situations, but also make longer passes to all parts of the field. Defenders are most likely to be making fewer, but longer passes to their teammates. Since the type of passes differ, and the average number of passes per match vary by position. Forwards average 21 passes per match, midfielders 41 per match and defenders 37 per match. This led to having the passes per ninety minutes broken up by position. For each position, passing helps the team in all circumstances, so passes variables are expected to have a positive coefficient.

A very important part of team performance is defending. This model improves on the Lucifora and Simmons (2003) model as it includes position specific metrics accounting for defending. A combined measure of a player's tackles and interceptions per ninety minutes measures their ability to defend against opponents (MidTklIntP90 and DefTklIntP90). Tackles are when a player stops the player from advancing on the dribble without fouling the player. Interceptions are when a player gains possession of the ball while an opponent was attempting to pass it to a teammate. Both of these are successful measures, so it is expected that the combined measure would have a positive coefficient. Defending is an aspect most important to midfielders and defenders, so the position specific variable is only included for these two positions.

Within the game of soccer, there are various forms of discipline for breaking the laws of the game. Anytime a player commits a foul the referee classifies the infringement within three categories and pairs it with appropriate punishment. The foul could be simply a careless common foul resulting in a free kick for the opponent, but no further discipline. It could be a reckless foul or tactical foul which would result in being shown a yellow card as a caution against any similar behavior. The third foul could be a dangerous challenge with excessive force resulting in a sendoff where the player is shown a red card, ejected from the match, and the team plays with one fewer player for the remainder of the match. This is the most severe punishment a player could receive for their actions on the soccer field. These signals of disciplinary action for the players could indicate erratic behavior that may put the team at risk. Players with questionable behavior may be detrimental to the team, so it is expected that these variables would have negative coefficients.

If a team is fouled in the opponent's penalty area, which is the area closest to the goal, they are then awarded a penalty kick (PK). These situations give the attacking team a tremendous opportunity to convert the PK to score a goal. This opportunity does not come absent of coinciding pressure for the selected player. Oftentimes the best player is the one selected to take the kick as a signal of confidence as the team relies on this advantageous opportunity. The players that successfully convert their attempts will continue to be trusted in the future and are key to helping the team win. Therefore, both the confidence displayed in being selected as a PK taker and the resulting goals from those opportunities will be expected to have a positive coefficient in predicting salaries.

Performance over the current season is important to see how indicative that performance is of a player's salary. However, their performance over their entire career is also important. This is why the variable StrikeRate is included. StrikeRate measures a player's accumulation of goals for every match they have appeared in. This signifies the long-term productivity of the player. It is anticipated that a positive coefficient will result for the position specific strike rates. The more productive a player is at accumulating goals, the higher the wages they could in turn expect. The square of StrikeRate is also included to be compared with the results in Lucifora & Simmons (2003). This concludes the performance measures included in the model to measure what metrics of performance impact earnings.

The third category of data variables in the model relate to player reputations. These reputations are built up overtime and are bolstered by their performance. However, there is an importance in finding ways to be noticed and appreciated by teammates, coaches, and fans around the world. Garnering a reputation as an impactful player would theoretically lead to a rise in player earnings. A "Superstar" dummy variable is assigned to players with an impressive strike rate throughout their career. For players who have played a minimum of 10 MLS games, those with a goals per match rate between 0.25 and 0.40 are in the Superstar2 category. Those with a goal per match rate greater than 0.40 are in the Superstar1 category. Due to the career success and positive reputation built through scoring goals, it is expected that each of these dummy variables have a positive coefficient. These players are the most impactful in scoring goals and are likely to receive the most attention of any player on their team and perhaps in the entire league.

As discussed in the background section, the MLS is unique in how the system allows designated players (DP). Clearly these three players are deemed worthy of a wage premium relative to their majority of their teammates and the coefficient in the model will be expected to be positive for the league's DPs.

Soccer is truly an international game. Since the beginnings of the MLS, the league knew that there was talent to be found in foreign players that would add value to their team. In order to scout this talent and convince them to leave their native country to play in an American league, it is likely going to take a wage premium to convince them to stay. The model includes a dummy variable marking the domestic, or American, players in the league. Due to the heightened expenses and wage demands to acquire foreign players, it is expected that variable for domestic players will have a negative coefficient.

International players are actively involved in the MLS. A number of them feature on the international stage through international competitions. The best players from each nation play as representatives on the national team. For the American players who are selected to play for the national team it is expected that they are superior players and demand superior wages. A dummy variable is used for only American players to measure if they have any experience playing for the national team. The coefficient for those who have played is expected to be positive. Further, the level of participation and national team experience differs for those that are selected. The number of international appearances is also included. Again, a positive coefficient is expected, as the more appearances indicates a higher quality player who can expect to earn more in MLS wages. These reputation variables all attempt to see whether greater recognition and exposure correlate with increased earnings. All variables are expected to have positive coefficients. If this is true, it may imply that a greater and larger reputation is correlated to earning more for playing in the MLS.

The fourth and final data category is Team Qualities. This considers a number of variables to determine if particular teams may have an advantage by offering a wage premium to certain players. The log of attendance is measured, and a positive coefficient is expected. If a team brings in more ticket revenue, they would have more to share with the players. Franchise revenue and its lag are similar inputs that measure a broader sample of revenue streams beyond simply tickets. These could include advertisements, jersey sales, and many other licensing deals. It is expected that both franchise revenue and its lag would have positive coefficients if they are statistically significant. Lastly, some team performance characteristics are included. Team success measured in points and the goal differential are included. The more points a team has, the better they performed over the course of the year. Similarly, goal differential is the net of goals scored and goals conceded. The more that a team outscored opponents over the course of the year, the more success they experienced. A positive coefficient is expected to be found with both team points and the team's goal differential.

Results

Since the empirical model used was an extension of that used by Lucifora and Simmons (2003), it was expected to yield statistically significant results. This expectation was confirmed through the F-test of the model which was shown to reject the null hypothesis that each coefficient was equal to zero with a p-value of (0.000). The overall fit of the model is important to examine as well. The adjusted R^2 of the model is 0.674 which means when accounting for the number of variables in the model, the model explains 67.4% of the variation in the data sample.

Upon running the regression, many of the expected results were statistically confirmed and are displayed in Table 2. Coefficients and corresponding p-values are listed for every variable in the model. Every additional year of age for a player increases the log salary by 0.1959. This result is statistically significant at the 99% confidence level. There is weak evidence of convexity to log salary with age. For every one unit increase in the square of age, the log salary falls by 0.0024. This is significant only at the 90% level and is thus a much weaker association. However, at the 90% confidence level we are able to conclude that MLS player salaries are increasing at a decreasing rate. None of the measures relating to experience or appearances in MLS had any statistical impact on MLS earnings. From this sample of data, it shows that experience in the league is not important for a player to earn more money. The lack of significance contrasts to Italy's Serie A as it was shown that both appearances and career appearances were statistically significant (Lucifora and Simmons, 2003). For the MLS, the insignificance could be due to a number of players being recruited from foreign leagues to improve the teams. Higher salaries are needed to attract these players, particularly if they are joining the league as a designated player. However, some of the league's best players are very tenured in the league and are also high wage earners. Some of the designated players are very experienced in MLS itself. While it is not clear what the reason for statistical insignificance is, these could be a few of the reasons.

For the MLS data, it is surprising to find that all of the position related statistics measuring goals and assists are statistically insignificant. In fact, the majority of performance variables included in the model are insignificant. There are six statistically significant performance variables at the 95% confidence level. To begin, each of the categories measuring passing are statistically significant at the 99% confidence level. This shows they are good predictors. Further, it shows that each position has a unique relationship between how they pass in relation to their resulting wages. Forwards gain the most relative to other positions for an additional pass. Forwards increase their log earnings by 0.001 for every additional pass. Midfielders and defenders increase their log earnings by 0.0009 and 0.0008, respectively. These coefficients take into account total passes over the entire 2018 season, of which most players made hundreds of passes.

Table 2 : Estimates of Players' Log Earnings Function			
Variable	**1 (All Players)**	**2 (Omit DPs)**	**3 (Only DPs)**
Age	0.1959 (0.004)	0.1732 (0.018)	0.2628 (0.145)
AgeSQ	-0.0024 (0.072)	-0.0019 (0.182)	-0.0033 (0.337)
Apps	-0.0034 (0.781)	-0.0043 (0.736)	0.155 (0.124)
AppsSQ	-0.0004 (0.228)	-0.0003 (0.365)	-0.0043 (0.047)
PreviousApps	0.0003 (0.755)	0.0005 (0.659)	0.0004 (0.914)
PreviousAppsSQ	0.000 (0.457)	-3.09E-06 (0.472)	-1.57E-05 (0.293)
ForGlsP90	0.2033 (0.441)	0.1353 (0.615)	0.9944 (0.573)
MidGlsP90	0.2213 (0.607)	-0.0885 (0.846)	-0.2732 (0.911)
DefGlsP90	-1.183 (0.238)	-1.18 (0.243)	0.7643 (0.92)
ForAstP90	0.2198 (0.37)	0.3453 (0.169)	-3.175 (0.015)
MidAstP90	0.0089 (0.982)	-0.1493 (0.72)	1.016 (0.412)
DefAstP90	-0.1204 (0.55)	-0.1117 (0.584)	-8.096 (0.66)
SuccDribP90	-0.0536 (0.062)	-0.0408 (0.17)	-0.0727 (0.541)
ForPasses	0.001 (0.001)	0.0008 (0.047)	0.0011 (0.095)
MidPasses	0.0009 (0.000)	0.0008 (0.000)	0.0014 (0.001)
DefPasses	0.0008 (0.000)	0.0008 (0.000)	0.0008 (0.601)
MidTklIntP90	-0.0092 (0.74)	-0.0055 (0.849)	-0.3809 (0.072)
DefTklIntP90	-0.0031 (0.902)	-0.0023 (0.928)	-
CrdYP90	0.0438 (0.743)	0.0352 (0.797)	1.319 (0.154)
CrdRP90	0.4452 (0.488)	0.4566 (0.489)	-5.056 (0.217)
TwoCrdYP90	-0.5175 (0.611)	-0.4586 (0.656)	-16.77 (0.169)
PK	-0.2398 (0.08)	-0.1305 (0.502)	-0.5049 (0.013)
PKatt	0.2339 (0.042)	0.1642 (0.302)	0.512 (0.011)
ForStrikeRate	2.874 (0.001)	3.079 (0.001)	2.828 (0.286)
MidStrikeRate	1.498 (0.403)	1.836 (0.346)	-8.358 (0.107)
DefStrikeRate	2.916 (0.176)	2.997 (0.174)	-
ForStrikeRateSQ	-2.047 (0.037)	-2.376 (0.027)	-4.893 (0.067)
MidStrikeRateSQ	0.9728 (0.839)	3.973 (0.498)	21.86 (0.045)
DefStrikeRateSQ	-3.675 (0.574)	-3.994 (0.544)	-
Superstar1	0.2589 (0.258)	0.7802 (0.057)	-0.0041 (0.993)
Superstar2	0.1905 (0.2)	0.2551 (0.133)	-0.189 (0.537)
DP	0.8557 (0.000)	-	-
Domestic	-0.5366 (0.000)	0.1732 (0.018)	0.1009 (0.803)
USAIntl	0.3389 (0.007)	-0.0019 (0.182)	-
USAcaps	0.0033 (0.208)	-0.0043 (0.736)	0.0002 (0.964)
LogAttendance	0.1565 (0.183)	-0.0003 (0.365)	-0.0946 (0.84)
LogLagFranchiseRev	0.0349 (0.65)	0.0005 (0.659)	-0.332 (0.068)
LogFranchiseRev	-0.0743 (0.599)	-3.09E-06 (0.472)	0.3258 (0.441)
Pts	-0.0033 (0.561)	0.1353 (0.615)	-0.0517 (0.018)
GD	0.0028 (0.539)	-0.0885 (0.846)	0.0562 (0.006)
Adj R^2	0.674	0.549	0.839
No. of Observations	541	489	52
Note: Dependent Variable = ln(Salary). P-Values are in parentheses. A constant term is included in each regression			

The other performance variables are more focused on dribbling and goal scoring. First is the number of penalty kick attempts. For every additional attempt, log earnings increase by 0.2339. This shows that it is the trust of coaching staff and teammates to attempt the penalty kick is more significant in determining earnings than converting the opportunity itself. While the season's goal scoring rate is not statistically significant, the goal scoring rate for a player's career is. This variable can conceptually be thought of as similar to Age as it ought to increase at a decreasing rate. The career goal rate is significant only for forward players, but forwards' log earnings increase by 2.874 for every one unit increase in the career strike rate ratio. The square of the forwards' strike rate is a negative 2.047. This implies that there is a diminishing return to high goal scoring rates in the MLS. This could arise due to non-designated players scoring goals at a prolific rate, but then having their salary capped. It also could arise from an upper limit on a sustainable goal scoring rate. Due to the nature of soccer, it is very difficult to maintain the success of consistently scoring at such a high rate. Even Josef Martinez, one of the most prolific goal scorers in MLS history has a career goal scoring rate of 0.88. This shows that even the game's best cannot be counted on for a goal in every game they participate in. Team executives have certainly developed a system to pay their players, but due to the insignificance of a number of variables, it is clear that player earnings are not based solely off of the performance variables included in this model. Table 3 is included below with a summary of the statistically significant variables used in Model 1 over the dataset with all players within the sample.

Table 3: Statistically Significant Variables in the Model		
Variable	**Coefficient**	**P-Value**
Age	0.1959	0.004
Age SQ	-0.0024	0.072
For Passes	0.0010	0.001
Mid Passes	0.0009	0.000
Def Passes	0.0008	0.000
PK att	0.2339	0.042
DP	0.8557	0.000
Domestic	-0.5366	0.000
USA Intl	0.3389	0.007

A player's reputation was one of the main factors tested in Lucifora and Simmons (2003). They were able to conclude that a superstar effect was present in Serie A during their study. In this study, no such premium is statistically significant for a similar marker based solely on the StrikeRate for players in MLS. However, this could be due to a clearer designation of the elite players being established in MLS as compared to Serie A. That is the designated player status, which is determined by teams for players they wish to pay a premium. Clearly, these players earn more, and as the results show, this variable is very statistically significant with a p-value near zero. Being a designated player is expected to raise the log wages of players by 0.8557. Some factors to the game of soccer cannot be measured or quantified. Things such as leadership, tact, personality, popularity, tactical knowledge, and passion for the game cannot be found in any single number and all impact the reputation of a player. However, many of these factors are things teams look for in designated players. Hopefully the players they select exemplify a number of these traits to help their teams win. If they do help their teams win more games than they would have otherwise, it is not a problem to compensate these designated players with a wage premium.

Two other aspects of a player's reputation come with their international recognition. First of all, due to the need to incentivize foreign players to relocate, a wage premium is often needed to acquire these foreign players. The results confirm that domestic, or US born players are expected to earn 0.5366 less in log earnings than those of another nationality. However, some American players are a step above the rest and are given the opportunity for both higher competition and greater recognition. This comes through representation at the US Men's National Team level. They are able to travel the world to play international competition along with competing at some of the biggest sporting events in the world including FIFA's World Cup. Any player in the MLS that has represented the US at the international level is expected to earn 0.3389 more in log wages. This is clearly a boost to reputation, but also is an indicator of a better player who likely has earned a higher wage. Similar to the MLS appearances, it is only whether a player has represented the US national team at all that is statistically significant, the number of appearances, whether 1 or 100, is statistically irrelevant.

The last vector of variables in the regression model measured the team qualities for each player. The model showed that for this data set, none of the included team qualities had any statistical significance on a player's log salary. Attendance, franchise revenues, a team's record, and their level of scoring compared to their peers ceased to make any matter. Due to the fact that the MLS is a single-entity ownership structure, this is more reasonable than with other leagues or ownership systems. In a way, the playing field is leveled across the league for players. MLS's self interest lies in advancing the league as a whole rather than just one particular team. Therefore, it will want players from all teams to receive similar wages so great salary inequality does not arise to lower production as was discussed by Levine (1991).

The results of this model have shown that only some aspects of a player's performance on the field statistically impact their earnings. Older players earn more, but the increasing rate of salary decreases convexly with age. Perhaps some of the most important things to determine salary in MLS are whether a player is a designated player, a foreign player, or a representative of the US Men's National team.

Model 1 was also used on two subsets of the data split on one defining characteristic as shown in Table 4. That is whether the players are designated players or not. Column 2 uses the model for all players held under the salary cap restrictions. Column 3 shows the model results for the league's 52 designated players. The summarized measures of model fit show that the model explains much more of the designated players' variation in the data than those players who are not designated players with respective adjusted R^2 of 0.549 and 0.839. Therefore, while both models could be improved, the data shows that this model could be best used to show the determinants of a designated player, rather than a standard MLS player.

Table 4	All Players	No DPs	Only DPs
Adj R^2	0.674	0.549	0.839
No. of Observations	542	489	52

One reason this data may not be the best measure of wages is that wages are not a continuously updated variable. Oftentimes a salary is a multi-year agreement. If a player either under performs or outperforms what their salary was negotiated for, their performance will not be fully reflective of their current earnings. To improve these findings, a panel regression could be set up with multiple years of data, to look at how the determinants studied in this paper compare with salaries overtime.

Conclusion

Sports are fun and fascinating events to watch but winning requires many different resources coming together. The better players that a team has, the more likely they will be able to win additional games. Team executives could use a model, such as the one used in this research, to examine a player's experience, performance, and reputation to determine whether a salary figure is worth offering to a player. The model is best applied specifically to designated players to predict their salary. Of the variables used in the models, the most significant determinants were age and its square (experience), passing and career strike rate (performance), and status as either designated players, domestic players, or US national team players (reputation). These variables were most statistically significant in determining the salary of MLS players.

The MLS shows no signs of stunted growth, as additional expansion teams are being introduced each year and the popularity of the sport is on the rise. As the league continues to grow, a model such as this one will likely need to be improved to account for different league characteristics that may arise. For further analysis, it would likely be helpful to consider a way to include homegrown talent subsidies into salary determinant models. Other additions could be including a player's all-star team selections too. In this model, the international experience was limited to US National Team players. However, a number of foreign-born MLS players represent their own countries as well. Including this fact could help improve the model and remove some, but not all, of the omitted variable bias in the results. One day sports will return, and hopefully their vigor will be stronger than ever. With a return to the sports world, and the added stress this pandemic has put on sporting administrations, it will be even more advantageous to further understand the various determinants of player salaries.

Reference List

"2018 Major League Soccer Player Stats." FBref.com. Accessed November 10, 2019. https://fbref.com/en/comps/22/1759/stats/2018-Major-League-Soccer-Stats.

Berri, David J., and R. Todd Jewell. "Wage Inequality and Firm Performance: Professional Basketballs Natural Experiment." *Atlantic Economic Journal* 32, no. 2 (2004): 130–39. https://doi.org/10.1007/bf02298830.

Coates, Dennis, Bernd Frick, and Todd Jewell. "Superstar Salaries and Soccer Success." *Journal of Sports Economics* 17, no. 7 (December 3, 2016): 716–35. https://doi.org/10.1177/1527002514547297.

Delgado, Fernando. "Major League Soccer." *Journal of Sport and Social Issues* 21, no. 3 (August 1, 1997): 285–97. https://doi.org/10.1177/019372397021003005.

Franck, Egon, and Stephan Nüesch. "The Effect of Wage Dispersion on Team Outcome and the Way Team Outcome Is Produced." *Applied Economics* 43, no. 23 (2011): 3037–49. https://doi.org/10.1080/00036840903427224.

Forrest, David, and Robert Simmons. "Team Salaries and Playing Success in Sports: A Comparative Perspective." *Sportökonomie*, 2002, 221–38. https://doi.org/10.1007/978-3-663-07711-4_12.

Gelb, Matt. "Baseball-Reference at 20: After Spawning a 1B Page-View..." The Athletic, January 3, 2020. https://theathletic.com/1502056/2020/01/02/baseball-reference-at-20-after-spawning-a-1b-page-view-business-whats-next-for-the-website-of-record/.

Holmes, Stanley. "Soccer: Time To Kick It Up A Notch." Bloomberg.com. Bloomberg, November 22, 2004. https://www.bloomberg.com/news/articles/2004-11-21/commentary-soccer-time-to-kick-it-up-a-notch.

Houmes, John. "European Football... It's Time for a Salary Cap-Part One." Bleacher Report, September 30, 2017. https://bleacherreport.com/articles/23882-European-Football-its-Time-for-a-Salary-Cap---Part-1-170508.

Lazear, Edward P., and Sherwin Rosen. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89, no. 5 (1981): 841–64. https://doi.org/10.1086/261010.

Levine, David I. "Cohesiveness, Productivity, and Wage Dispersion." *Journal of Economic Behavior & Organization* 15, no. 2 (1991): 237–55. https://doi.org/10.1016/0167-2681(91)90031-r.

Lucifora, Claudio, and Rob Simmons. 2 "Superstar Effects in Sport: Evidence from Italian Soccer." *Journal of Sports Economics* 4, no. 1 (February 2003): 35-55. https://doi.org/10.1177/1527002502239657.

"Major League Soccer - Top Market Values 2018." Transfermarkt. Accessed October 18, 2019. https://www.transfermarkt.us/major-league-soccer/marktwerte/wettbewerb/MLS1.

Shapiro, Stephen L, Timothy Deschriver, and Daniel A. Rascher. "The Beckham Effect: Examining the Longitudinal Impact of a Star Performer on League Marketing, Novelty, and Scarcity." *SSRN Electronic Journal*, May 2014. https://doi.org/10.2139/ssrn.3421674.

SportyCo. "How Big Is the Sports Industry?" Medium. SportyCo, November 13, 2017. https://medium.com/sportyfi/how-big-is-the-sports-industry-630fba219331.

"Stats." MLSsoccer.com. Accessed November 10, 2019. https://www.mlssoccer.com/stats/.

Szymanski, Stefan, and T Kuypers. "The English Football Industry: Profit, Performance and Industrial Structure." *International Review of Applied Economics* 11, no. 1 (1997): 135–53. https://doi.org/10.1080/02692179700000008.

Torgler, Benno, and Sascha L. Schmidt. "What Shapes Player Performance in Soccer? Empirical Findings from a Panel Analysis." *Applied Economics* 39, no. 18 (2007): 2355–69. https://doi.org/10.1080/00036840600660739.

Wahl, Grant. "Grant Wahl: New Afterword from The Beckham Experiment." Sports Illustrated, June 1, 2010. https://www.si.com/more-sports/2010/06/01/beckham-experiment.