Original Research Paper

An Econometric Analysis of Home Field Advantage in the World's Largest Sports: Pre and Post COVID-19 Outbreak

[1,2]**Nazmul Islam, [2]Peter LeRoux and [2]Nicholas Correia**

[1]*Department of Humanities, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh*
[2]*Department of Economics, Colgate University, Hamilton, New York, USA*

Corresponding Author:
Nazmul Islam,
Department of Humanities,
Bangladesh University of
Engineering and Technology
(BUET), Dhaka, Bangladesh.

and

Department of Economics,
Colgate University, Hamilton,
New York, USA.

Email:
nazmul786humbuet@gmail.com

Abstract: The purpose of this study is to measure the home field advantage across sports and within sports. The study also sets forth an analysis of the soccer games played in empty stadiums during the COVID-19 pandemic. We measure whether there is any difference between playing in front of fans and without fans: Something that has not been captured before. We use data for the years 2016-2019 from the top five UEFA soccer leagues, the NBA and the NFL. We find that home field advantage is present in soccer, however, there is no significant difference between its top five leagues or over the last four years. We also find that for both the NBA and the NFL, there is no home advantage, however, there is a difference between each of their respective divisions. Finally, we show that playing in front of no fans for the UEFA leagues because of COVID-19 causes home field advantage to dissipate.

Keywords: Home Field Advantage, COVID-19, Divisional and League Advantage

Introduction

Athletic spectators are amongst the most historic and spirited people all over the world. They have undying passions for their clubs. For instance, every four years, around three million people travel to watch the Fédération Internationale de Football Association (FIFA) World Cup in attendance. Broadcasting data from FIFA.com (2018) shows that 3.572 billion people, over half of the global population, watched the record-breaking 2018 World Cup in Russia in total. Clearly, these expansive audiences must have some impact on how teams play. Players may have different responses to large audiences such that, for example, they may play better in a noisy stadium or buckle under the pressure of tens of thousands of spectators jeering them on. The effect of home field advantage and large groups of fans may give the sporting world some insight as to how their respective sports are influenced by forces outside of the playing field.

In the model set forth by the experiment, we integrate variables such as team salaries, referee behavior and in-game statistics that have not been tested against each other

before. This study not only measures the home field advantage across sports, but also within the different divisional bodies of each sport. Therefore, our analysis will portray any innate differences in home field advantage between divisions or even individual teams of a particular sport. Our research will incorporate similar statistical testing of the European soccer games played under the COVID-19 fan-free regulations. In doing so, our models determine whether there exists any statistically significant difference between playing in front of fans and without fans. In order to make such a comparison, we first come up with our own definition of home field advantage using our own variables and regression analysis. Once we have found a quantitative estimation of this measurement, we can run the same test for the European soccer games played under their respective COVID-19 regulations and compare both measurements of home field advantage.

The main objectives of this study are:

(i) Finding the effect of home field advantage on the outcome of a game in different sports
(ii) Identifying differences in home field advantage within individual sporting leagues as opposed to across different sports

Science Publications

(iii) Calculating any change in home field advantage that arises from the COVID-19 outbreak regulations in sports

Literature Review

Chatterjee *et al*. (1994) uses both linear and logistic regressions to calculate which elements of basketball have the largest impacts on team performance in the National Basketball Association (NBA). The win-loss percentage of any given team represents their dependent variable. Accounting for more than 90% of the variation in their data, field goals, free throws, rebounds and turnovers are some of the independent variables that Chatterjee *et al*. (1994) finds to be statistically significant. Carmichael *et al*. (2000) attempts to capture the performances of teams in the English Premier League (EPL) by replicating a production function used to measure team performance in Major League Baseball (MLB). They include variables such as goal differential and shot differential to measure a given team's success rate. While both of these studies highlight some of the more important elements of basketball and soccer, they fail to account for a crucial element in team performance: Home field advantage. In the research performed by Waters and Lovell (2002) and Trandel and Maxcy (2011), home advantage is proposed as a potential source of influence on many sporting competitions around the world.

Waters and Lovell (2002) discusses some of the characteristics of home field advantage in English soccer in regard to players' retrospective perceptions of confidence and psychology. They claim that playing at home allows players to feel more confident when they are winning a game. However, when the home team is losing, these players feel significantly less confident than if they were to be playing on the road. They also find that football players feel as though "they are not expected to win" when they attend an away game and, thus, feel less pressure. The study concludes that players feel more positive at home, more confident at home than on the road and more anxious on the road.

Trandel and Maxcy (2011) cites that sporting analysts have measured the competitive balance between sporting leagues by comparing the standard deviations of winning percentages of all the teams in a given league and comparing them to that of other leagues. One issue with this method is that leagues have differing numbers of games in their seasons. For example, the MLB plays a 162-game season whereas the National Football League (NFL) plays only 16. It would be inaccurate to compare the winning-percentage standard deviations of these

teams given that, in a short season, each individual game is worth significantly more than it would be in a long season. Thus, any individual result from a short season has a greater impact on a team's record than it would in a long season. Trandel and Maxcy (2011) fixes this problem by dividing the standard deviation of team winning percentages by the standard deviation of the expected outcome of each game if each team were equally likely to win. However, it is never the case in any professional sport that each team has the same chance of winning. In fact, Trandel and Maxcy (2011) shows that the home team tends to win more than 50% of time on average. This contributes to the definition of home field advantage. When this effect is not taken into account, the ratio of standard deviations is smaller than it should be. This system of calculations overestimates the competitive balance of a particular league or sport. Trandel and Maxcy (2011) attempts to adjust for these biases by creating a formula that calculates a "home-advantage-corrected ideal standard deviation." They find that the NBA has the largest home field advantage among North American sports. In correcting for this bias, they show how US sports are competitively imbalanced. They also note that, in doing this adjustment, the competitive difference between the NFL and MLB is modestly decreased. Given that both Waters and Lovell (2002) and Trandel and Maxcy (2011) suggest some potential differences between playing in home and away games as mentioned above, we wish to replicate and expand upon the research made in those articles to measure the performances of sports teams when considering the advantage that teams have when playing at home. This has been performed and quantified in a wide range of literature considering the home field advantage of individual sports, (Clarke and Norman, 1995; Nevill *et al*., 1996; Omotayo, 2003; Pollard *et al*., 2008; García *et al*., 2013; Pollard and Gómez, 2013; Kotecki, 2014; Van Damme and Baert, 2019) and even across multiple sports (Trandel and Maxcy, 2011; Pollard and Gómez, 2014; Pollard *et al*., 2017). Given the extensive amount of research pertaining to the subject, we knew that an experiment involving home field advantage was easily replicable. While the majority of home field advantage research pertains to soccer in specific, we found that we could best set up our study by incorporating aspects of single-sport research and multi-sport comparisons.

Clarke and Norman (1995) focuses on the EPL to analyze the impact of home field advantage on the outcome of soccer matches from 1981-1990. The authors established a dummy variable for each team that competed in the EPL during those years. They found that the difference in home field advantage each year was statistically significant. While these findings may be outdated by now, the magnitude of home field

advantage may have changed since then. This piece of literature does not omit any information that is needed to replicate the data as it presents each computation we would need if we ever felt the need to replicate something similar in our experiment.

Nevill *et al.* (1996) uses the claim from Schwartz and Barsky (1977) that there was a significant relationship in home winning percentage and attendance in the MLB to suggest the importance of their study of the home field advantage in English and Scottish soccer. They found that home teams won 48% of their home matches when playing in front of low-density crowds. In contrast, teams playing in front of high-density crowds won 57% of their home matches. Nevill *et al.* (1996) found that this conflicts with the findings from Dowie (1982) and Pollard (1986) that both absolute crowd size and crowd density, respectively, have no significant effect on home wins. Kotecki (2014) reached a similar conclusion using a logit regression. Their analysis shows that a one standard deviation increase in attendance will result in a 2.7% greater chance that the home team will win. For the reasons set forth in both Nevill *et al.* (1996) and Kotecki (2014), we decided to include a regression of home attendance and home win percentage in our study.

Nevill *et al.* (1996) also highlighted the importance of referee behavior in the outcomes of sporting events as suggested by Agnew and Carron (1994) given that referees have been proven to officiate matches in favor of the home team (Greer, 1983). This study tries to explain that the home field advantage in both English and Scottish soccer is associated with attendance as explained by the refereeing behavior in the 1992 and 1993 seasons. In essence, the authors are trying to disprove the findings of Dowie (1982) and Pollard, 1986). Nevill *et al.* (1996) quantifies referee behavior by recording the number of sending offs and converted penalty kicks in the 1992 and 1993 seasons. The study found, with statistical significance, that leagues with higher attendance exhibited a greater home field advantage than leagues with lower attendance and that the home teams of these higher-attendance leagues also displayed less sending-offs and more penalties converted. These findings conflict with those from Dowie (1982) and Pollard (1986). Nevill *et al.* (1996) suggests this is the case because Dowie (1982) and Pollard (1986) did not test in leagues with lower attendance. In addition, Kotecki (2014) also found that referees tend to call games in favor of the home team. These findings suggest that some measure of referee behavior is going to be necessary in our experiment. While this may be the effect that we are looking to conclude in our paper, we believe that there is a more accurate measure of referee behavior than set forth in Nevill *et al.* (1996). It is important to note that the number of sending offs and converted penalties may, in fact, be correlated with each other since it is very possible to receive a red card and give away a penalty kick in the same play. However, evidence from Altman (2014) tends to suggest that incurring a red card

and a penalty kick simultaneously resulted in no extraneous disadvantage and, instead, scores suggest the two effects had occurred separately. Regardless, we have decided to incorporate a referee behavior variable in our experiment as suggested by Nevill *et al.* (1996) but choose to do so in a form of our own choosing.

In addition to our other independent variables, we wanted to control for any effects that may arise from in-game statistics. For example, Chatterjee *et al.* (1994) finds that field goals, free throws, rebounds and turnovers are statistically significant in their research. Among these, independent variables are also the assist and steal differentials. In Carmichael *et al.* (2000), the authors attempted to analyze a production function that has been used for measuring MLB teams' success and transpose it to the EPL. They state that they are the first people to create a production function model for the EPL. In the production function, they include variables such as goal differential and shot differential after identifying them as key metrics in a team's performance and, as such, they exemplify statistical significance. The research mentioned above suggests that the inclusion of in-game statistics as independent variables in our study will help us eliminate endogeneity through omitted variable bias from our model.

Pollard and Gómez (2014) tested to see which country in all of the UEFA soccer leagues exhibited the largest home field advantage between the years 2006 and 2012. The study found that it was the greatest in Nigeria (86.82%). Bosnia-Herzegovina, Guatemala, Indonesia, Algeria, Bolivia and Ghana all exhibited home field advantages that were in between 70 and 80%. This evidence suggests that we must control for league in our data because it shows an effect that supports our alternative hypothesis that different leagues have different home field advantages such that $H_A \neq 0$.

In his experiment, Omotayo (2003) uses the number of goals scored as his measure of home field advantage instead of winning percentage. This is useful in a study only considering soccer because it avoids the complicated points system used in soccer leagues around the world. They use differing point totals associated with the potential outcomes of a match; a win, draw, or loss. However, for the purpose of our experiment, we find this method to be fruitless given that it cannot be measured across sports. A goal is the only method of scoring in soccer and it counts as one. In the NBA, one can score one point for a free-throw, two points for a field goal and three points for a three-pointer and teams tend to score over a hundred points a game. In football, a touchdown is worth six points, an extra point is worth one, a field goal is worth three points and a safety is worth two. Here, we can see how the scoring systems between different sports can differ immensely. Thus, we must find a way to measure home field advantage in terms of a percentage rather than points as suggested by Omotayo (2003).

Reade *et al.* (2020) links the COVID-19 pandemic with a significantly smaller home field advantage in sport. Via a series of natural experiments, they are able to test whether social pressure affects referee behavior and outcomes of European soccer matches. They observe that, in any given match played in an empty stadium, referees cautioned the visiting team significantly less often. In fact, they did so by over one third of a yellow card per match. Thus, the authors conclude that the referees favored the home team less in matches played behind closed doors. Even though the spread of COVID-19 altered the myriad ways we consume sporting events, these results highlight a recent theme in sports economics; that home field advantage decreases with the absence of fans. Naturally, this group of literature is fairly limited, due to the recent nature of its subject matter. Our intention is to build upon this literature and incite further research into home field advantage and its many factors.

Conceptual Framework

In order to measure the success of teams in the UEFA, NBA and NFL, we need a dependent variable that represents each sport's metric of success. In the NBA and NFL, the success of a given team's season is measured by their record, the combination of wins and losses accumulated throughout the season. The teams with the most wins are then seeded into the playoffs to compete towards the championship. In the UEFA leagues, success is measured by a team's position on the league table. The table is determined by how many points a team has acquired over the course of the season. In European soccer, when a team wins a game, they are awarded three points to add to their total points tally in the league table. For a draw, each team wins one point and when a team loses, they earn zero points. At the end of the year these points are added up and the team with the most points wins the league. Given that there are two different measures of team success in our three sports, we will have to create two dependent variables; one for UEFA and one for the NBA and NFL. We decided that, since the primary goal of NBA and NFL teams is to acquire as many wins as possible in their season, we will make the winning percentage (*win_percent$_i$*) as a dependent variable for those two leagues. The winning percentage of each team will measure how much their record grows with each game. For the UEFA, we chose to incorporate the natural log of average points per game (*ln_avg_ppg$_i$*) as the dependent variable given that the primary goal of European soccer teams is to gain as many points as they can in a season. The natural log of average points per game will measure how many points a team will earn each game as a percentage. By splitting our dependent variable

into two different measurements, we are accounting for the fact that success is calculated differently in UEFA leagues than it is in the NBA and NFL. That is to say, we are measuring the improvement in each team's season after each game by either how many points or how many wins they acquired from that game.

Given that we are measuring the home field advantage of teams in our respective leagues, we use a dummy dependent variable (*home$_i$*) to describe which games are home and away. If a team is playing at home, then this value will equal one such that home equals one. When a team is away, this value will equal zero.

Nevill *et al.* (1996) and Kotecki (2014) both suggest that we should include a variable for referee behavior in our models. We chose to adopt our own interpretation of such a variable because we found that the measurement for referee behavior set forth in Nevill *et al.* (1996) could potentially generate biasedness. We measure referee behavior in soccer by the differential number of cards received per game (*fouls_diff$_i$*). A yellow card is worth one and a red card is worth two given that if a player receives two yellow cards in a match, he will be given a red card and subsequently sent off the field. For example, if Arsenal average is 2.4 cards per match, they are usually awarded with slightly more than either one red card or two yellow cards in a given match when compared to the other team. We measure referee behavior in the NBA by the differential number of fouls committed per contest (*fouls_diff$_i$*). Similarly, we measure referee behavior in the NFL by the differential number of penalties committed per contest (*penalties_diff$_i$*).

We include a fixed effect for the previous four seasons in our experiment as suggested by Clarke and Norman (1995). This is because Clarke and Norman (1995) find the differences in home field advantage between years to be statistically significant. Even though their research was performed thirty years ago, we still deem it necessary to account for the year in our models. Not only does this remove potential endogeneity from our model, but it also allows us to collect a larger sample size in our data. By including data from four years, 2016-2019, we can test if this relationship found by Clarke and Norman (1995) still holds true today and if there is any statistical significance between these more recent years.

We choose to include various in-game statistics from soccer, basketball and football as independent variables in our study. This idea was presented in Carmichael *et al.* (2000). He proposed a production function that includes variables such as goal differential and shot differential. Eventually, these differentials were found to be statistically significant in his experiment. In our study, we create variables for

both of these metrics for our UEFA regressions (*goal_diff$_i$; shots_diff$_i$*). Similar to Chatterjee *et al.* (1994), we identified that the differentials in score, rebounds, assists, turnovers and steals be given independent variables for our NBA regressions (*score_diff$_i$; rebounds_diff$_i$; assists_diff$_i$; turnover_diff$_i$; steals_diff$_i$*). In the NFL, we came up with our own in-game statistics differentials for our independent variables. They are yards, score, redzone, passer rating and turnovers (*yards_diff$_i$; score_diff$_i$; redzone_diff$_i$; passer_rating_diff$_i$; turnover_margin$_i$*).

Similar to Pollard and Gómez (2014), we included the different leagues and divisions in each sport. In their study, they found that there was a statistically significant difference between European soccer leagues. For this reason, we tested for home field advantage in the five most prominent leagues in the UEFA. They are the Bundesliga, La Liga, the EPL, Serie A and Ligue 1. We then paralleled this practice to the NBA and the NFL where each league is divided into multiple divisions. In the NBA, these are the Atlantic, Central, Southeast, Northwest, Pacific and Southwest divisions. In the NFL, the entire league is divided into two conferences, the American Football Conference (AFC) and National Football Conference (NFC). These conferences are then divided into four separate divisions: The North, South, East and West divisions. We included all of these in our model to check for any statistically significant difference between them.

We also include average_salary$_i$, as the independent variable in the study. This value indicates the average player salary on a given team. The payroll each team incurs usually indicates the amount of talent they have tied to their roster. For example, players like Christiano Ronaldo, Stephen Curry and Aaron Rodgers will make significantly more than an average player. Thus, a team consisting of higher-contract players has a greater chance to win games than a team filled with lower-contract players. We incorporate this into our model to control for any fixed effect that may arise from teams with larger payrolls due to the host of talent attached to their rosters. The idea of these large salaries influencing our data is briefly expanded upon in Sharp (2019). He claims that players with large salaries will welcome the chance to play in away matches because, after their athletic requirements are fulfilled, these high-earning athletes can spend their money and enjoy time with their friends in the away cities. This is particularly convincing when you consider a team from a smaller city such as the Green Bay Packers traveling to a bustling, lavish city like Los Angeles to play the Rams.

Data Description and Analysis

We use panel data for the years 2016-2019 from UEFA, NBA and NFL statistics databases. Before running any of our main descriptive models, we observe the relationship between each sport's home and away advantage. For the UEFA, we have created a scatter plot of every team's average point per game and their home or away status. For the NBA and the NFL, we have created a scatter plot to observe the relationship between home and away win percentages. Keep in mind that home and away status is represented by the *home$_i$* dummy variable. As such, a home team will score a value of 1 and an away team will score a value of 0 for this independent variable. Any team that is away will represent the reference category.

Figure 1 shows the map of country home field advantages as found in Pollard and Gómez (2014). In Fig. 2, we can observe the upward sloping fitted line which indicates that there is a difference in means of the average points a team earns when they play at home versus on the road. As we can see, teams who play at home seem to earn more points per game than teams who play away, on average. In Fig. 3 and 4, we find the same relationship, however, it seems that the trend line for the NFL is not as dramatic as it is for the other two sports. We also observe the relationship between a home team's attendance and the outcome of the game in question. Figure 5, 6 and 7 display the relationship between home team success and home attendance. For the UEFA, we use average points per game and for the NBA and NFL, we use winning percentage. We noticed there is a distinct, positive correlation between attendance and the outcome in case of UEFA and NBA. However, we observe very weak to almost non-existent correlation in case of the NFL. We supposed this is the case because NFL stadiums are usually pretty full especially when compared to NBA stadiums given that fans only have sixteen chances to watch their team play each season (Goodell, 2019).

Lastly, we analyze referee bias for the home team. In Fig. 8, 9 and 10, we illustrate the relationship between the fouls or penalties awarded to the home and away teams in any given contest. These figures demonstrate that, for all three sports, the home team is less likely to be called for a foul. This brings to light an unverifiable question, do teams actually commit less fouls and play a more error-free game at home, or do the referees produce a bias and call less fouls on the home team? Fig. 8, 9 and 10 suggest the ladder, but it is certainly not proven.
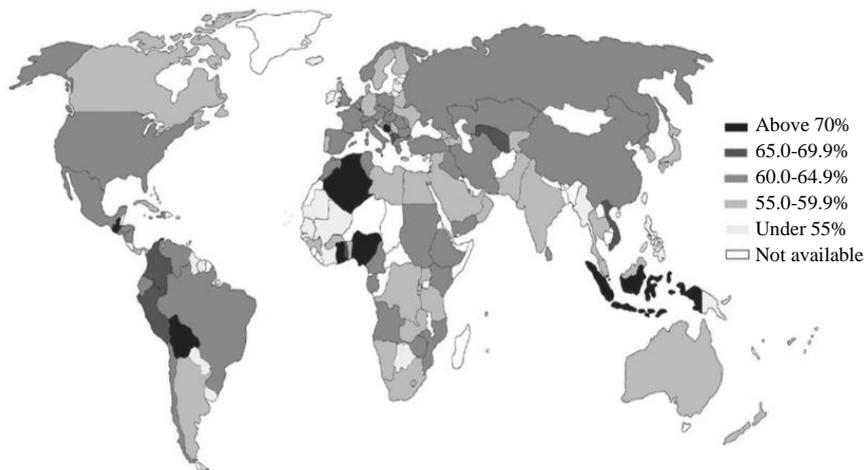
Fig. 1: Map of country home field advantages as found in Pollard and Gómez (2014)
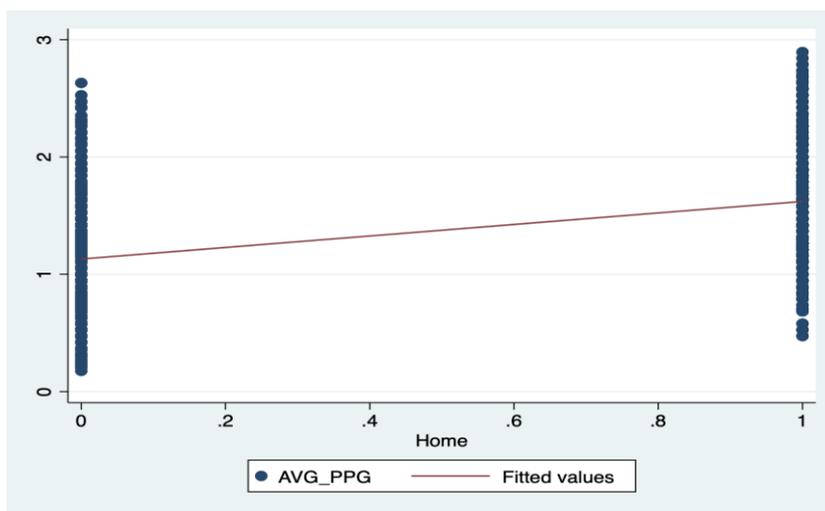
Fig. 2: UEFA points per game for home and away teams
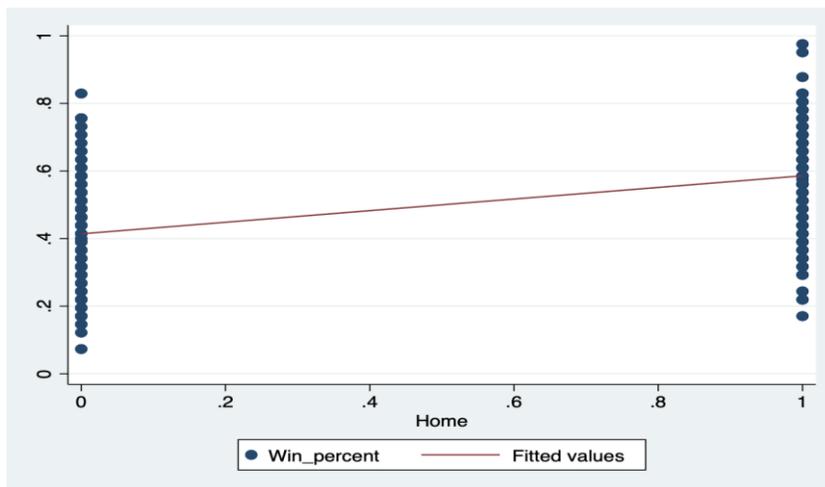
Fig. 3: NBA home and away win percentages
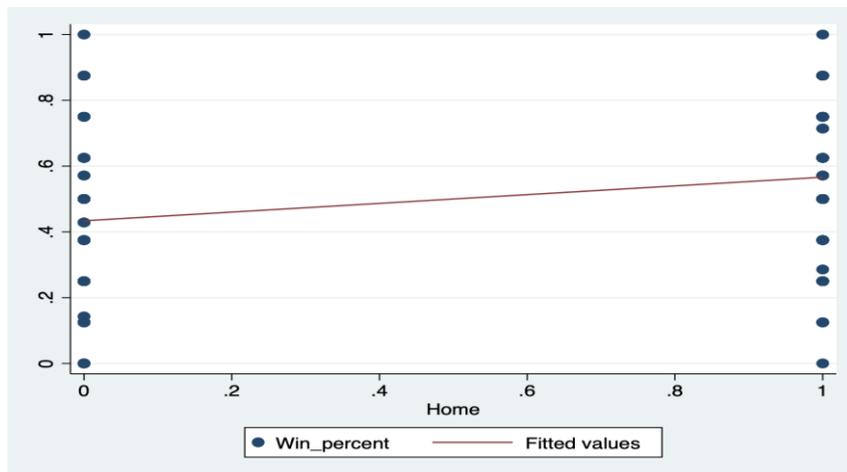
Fig. 4: Graph of NFL home and away win percentages
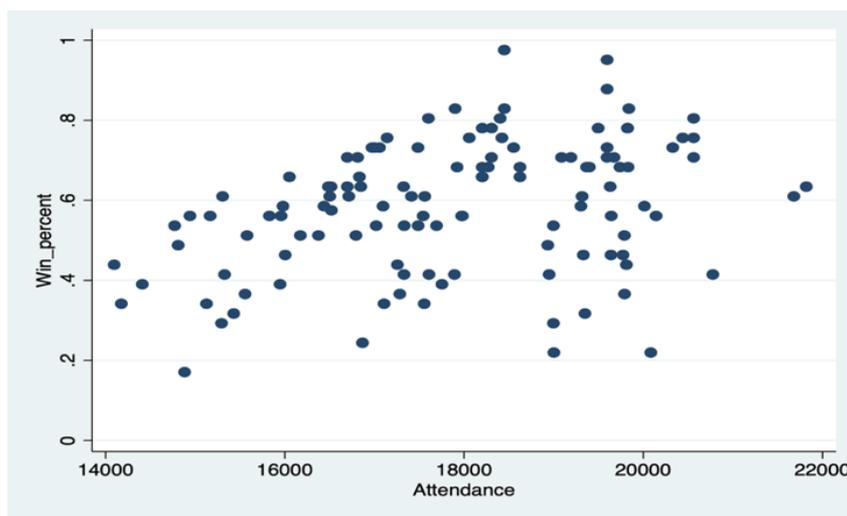
Fig. 5: UEFA attendance (in 1000s) and home average points per game
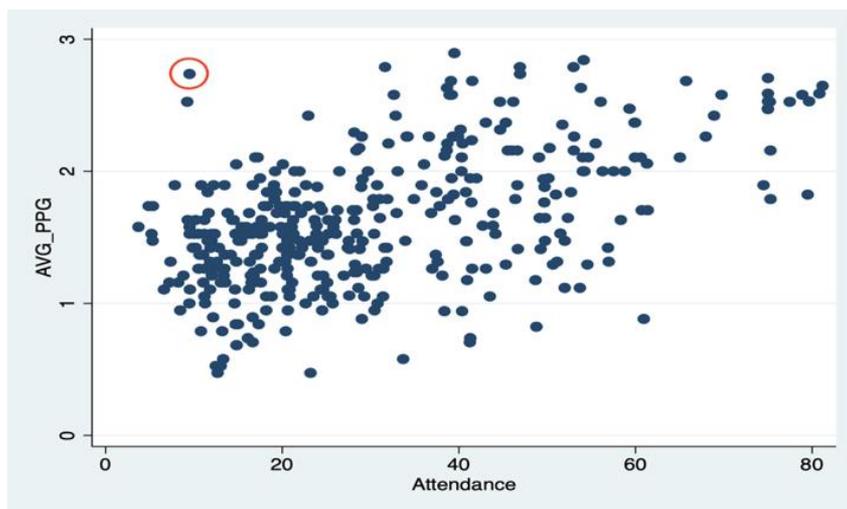
Fig. 6: NBA attendance and home win percentage
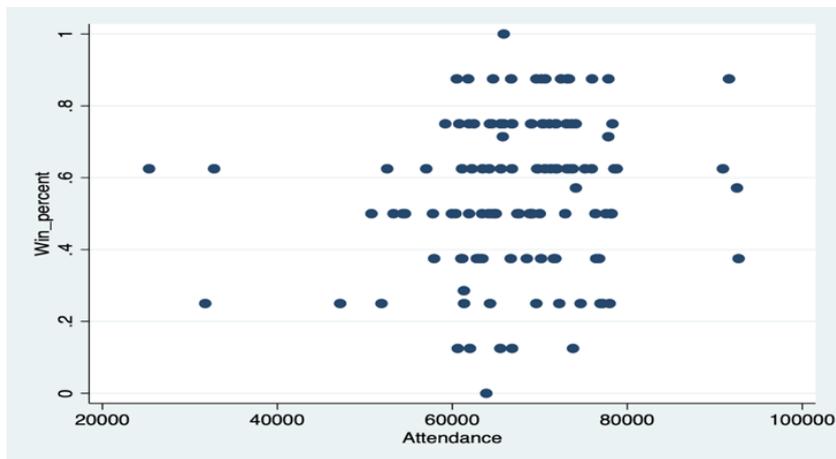
Fig. 7: NFL attendance and home win percentages
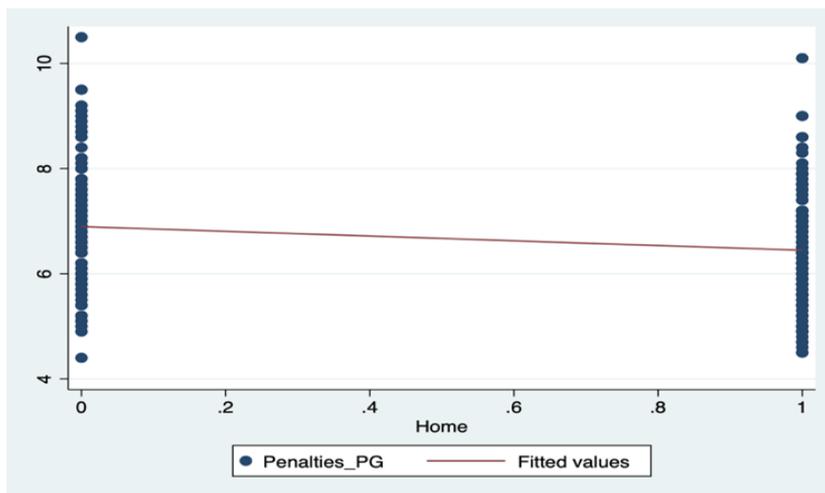
Fig. 8: NFL difference in penalties between home and away team
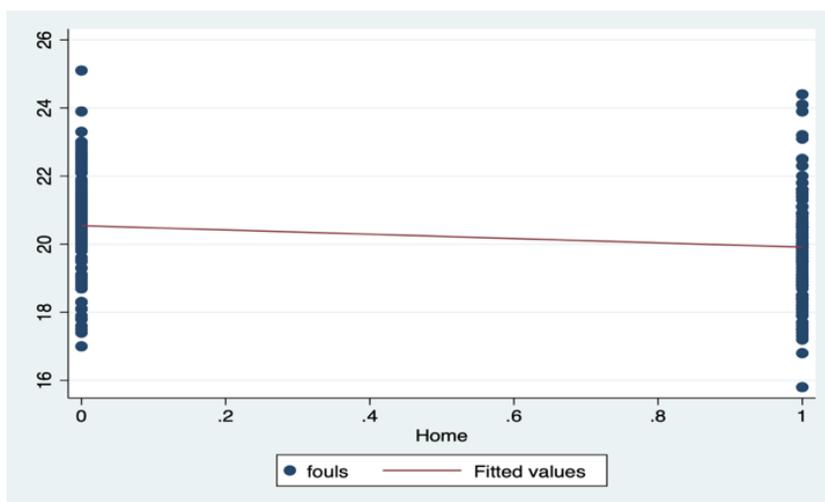
Fig. 9: NBA difference in fouls between home and away team
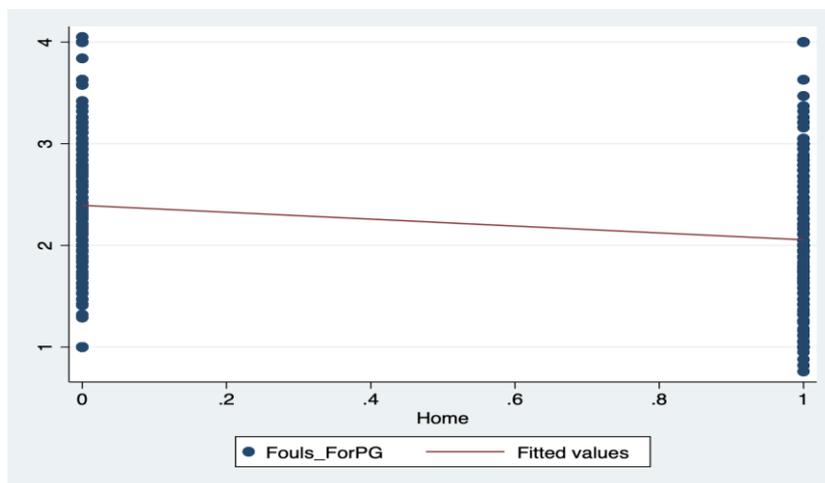
Fig. 10: UEFA differences in fouls between home and away team

Empirical Models and Estimation Methods

In our study, we ran a bivariate model, a pooled multivariate regression, a one-way fixed effect regression and a two-way fixed effect regression with both fixed effects models using the Least Squares Dummy Variable (LSDV) method. For the bivariate model, we have created a home dummy variable that represents either home or away. Our bivariate model for the five main UEFA leagues, the NFL and the NBA are:

i. UEFA: $\text{Ln}(avg\_ppg) = \alpha + \alpha 1\ home_i + \varepsilon_i$
ii. NFL: $Win\_percent = \beta_0 + \beta_1\ home_i + \vartheta_i$
iii. NBA: $Win\_percent = \gamma_0 + \gamma_1\ home_i + \delta_i$

In regards to the NBA, game results are binary in that a team must either win or lose as per the (NBA Rulebook, 2019). If a contest results in a draw at the end of regulation, successive overtimes will be played until a winner is decided. For the NFL, their rulebook (Goodell, 2019) also has an overtime rule; however, only one overtime is played and if the score remains tied, then the end result is a draw. This rule has supplied the league with very few draws as seen in historical NFL team records. Draws rarely happen over the course of an NFL season which is why, as per the rule book, the NFL seeds its teams using their record, a measure of winning percentage, much like the NBA. Therefore, for the NFL and NBA regressions, we will use winning percentage (*win_percent*) as the dependent variable.

Win percentage in the UEFA leagues would not encapsulate the home field advantage given that matches end in draws quite frequently. If we want to take these draws into account, we cannot use win percentages as we did with the NBA and NFL. Given that soccer measures team success in terms of points rather than winning

records, we will use the natural log of average points per game (ln(*avg_ppg*)) to represent the dependent variable.

We have created the following multivariate models for each sport:

iv. UEFA: $\ln(avg\_ppg)\text{-} = \alpha_0 +$
$\alpha_1\ home_i + \alpha_2\ goal\_diff_i +$
$\alpha_3\ shots\_diff_i +$
$\alpha_4\ fouls\_diff_i +$
$\alpha_4\ average\_salary_i + \varepsilon_i$

v. NFL: $Win\_persent_i = \beta_0 +$
$\beta_1\ home_i + \beta2\ yards\_diff_i +$
$\beta_3\ score\_diff_i +$
$\beta_4\ redzone\_diff_i$
$\beta_5\ passer\_rating\_diff_i +$
$\beta_6\ turnover\_margin_i +$
$\beta_7\ average\_salary_i +$
$\beta_8\ penalties\text{-}diff_i + \vartheta_i$

vi. NBA: $Win\_persent_i = \gamma_0 +$
$\pi_1\ home_i + \gamma_2\ score\_diff_i +$
$\gamma_3\ rebound\_diff_i +$
$\gamma_4\ assists\_diff_i +$
$\gamma_5\ turnover\_diff_i +$
$\gamma_6\ steals\_diff_i + \gamma_7\ fouls\_diff_i +$
$\gamma_8\ average\_salary_i + \delta_i$

Since these multivariate models include other independent variables, they may avoid potential sources of endogeneity that stem from Omitted Variable Bias (OVB). In these regressions, we will discern how each variable affects the percentage change in a given team's

average points per game or win percentage. All the above variables are measured on an average per game basis for every team for each season and are subtracted by the collective average of each team's opponents for each season as well. For each of the following definitions, the information came from each respective sport's official rulebook. For the UEFA regression, the $goal\_diff_i$ variable represents the goal difference between a team and their opponent on a per game basis. A goal is scored when the ball completely crosses the goal line which is positioned between the two goal posts. $Shots\_diff_i$ is the difference between the number of shots each team has on goal per game. Shots are considered to be a directional attempt to score a goal. $Fouls\_diff_i$ are the differential number of yellow cards and red cards a team receives in a match. As per the FIFA.com (2015), a yellow card is seen as a foul that comes from an aggressive infraction and cautions the player to not commit another offense. A red card is when the foul is exceedingly aggressive and the player is then thrown out of the game and the team cannot substitute for that player. If a team receives a red card, they will play with 10 men versus the 11 opposing players for the remainder of the game. A player can get a red card from also earning two yellow cards in a single match. Lastly, $average\_diff_i$ is each team's average salary per player. This is calculated by taking a team's total salary and dividing it by the number of players on the roster; this is measured in pounds for all teams.

For the NFL, $yards\_diff_i$ is the average difference in a team's total yards gained and yards allowed per game. In the NFL, yards are an offensive metric that measure how far a team travels down the 100-yard playing field. $Score\_diff_i$ is the difference in the score. In the NFL, scores can be counted from a touchdown, an extra point, a two-point conversion, a field goal, or a safety. A touchdown occurs when a team crosses the goal line with possession of the ball and results in 6 points. An extra point is awarded to the team if they kick the ball through the goalposts after a touchdown is scored. A two-point conversion is successful if a team crosses the goal line after a touchdown is scored instead of conducting an extra point. A field goal is when a team kicks the ball through a goal post after failing to reach first down to earn 3 points. $Redzone\_diff_i$ is the difference in red zone opportunities. As per the NFL handbook (Goodell, 2019), a redzone opportunity is when a team is within twenty yards to the endzone. $Passer\_rating\_diff_i$ is the difference in an individual team's quarterback passer rating and that of the opposing quarterback. Passer rating, as defined by NFL statisticians, measures a quarterback's performance in a game. NFL statisticians and analysts such as DaSilva (2017) have also revealed that the quarterback is the most important position on the team and, thus, have the greatest influence on the outcome of a game. Therefore, we have considered quarterback performance in our model. The variable $turnover\_margin_i$ is the difference in takeaways between each team. A takeaway is when a defending team intercepts the quarterback's pass or recovers the ball after it has been fumbled by the offensive team. $Average\_salary_i$ for the NFL is measured in the same way as the UEFA's average salary except it is counted in dollars. Lastly, $penalties\_diff_i$ is the difference between penalties drawn by a team and the penalties they commit. Penalties are given to a team if they break the gameplay rules set forth in the NFL handbook (Goodell, 2019).

For the NBA, $score\_diff_i$ is the average difference between the two scores of a game. As per the NBA Rulebook, 2019, a team scores two points if the ball goes in the basket from within the three-point line. A three-pointer occurs if the player scores from beyond the three-point line. A foul shot is taken from the free throw line when a foul is committed. Each foul shot is worth one point. $Rebounds\_diff_i$ is the average difference between each team's rebounds during a game. A rebound is awarded to a player if he grabs the ball from a missed shot. $Assists\_diff_i$ is the average difference between the two team's assist totals per game. An assist is awarded to a player if the player he passes the ball to scores a point. $Turnaover\_diff_i$ is the average difference of two team's turnover average per game. As defined by the NBA, a turnover occurs when a team loses possession of the ball before a shot attempt is made. $Steals\_diff_i$ is the difference between each team's number of steals per game. A steal is when a team takes the ball from the opposing team without the ball going out of bounds or being scored. $Fouls\_diff_i$ is the difference between the number of fouls a team draws versus what they commit per game. Lastly, $Average\_salary_i$ is the average salary per player for every team which is measured in the same way as the NFL.

We also use fixed effects models (one-way fixed effect models and two-way fixed effect model) to fight endogeneity for each sport since we believe that there may be underlying fixed effects associated with various aspects of each sport. Since each league or division has a different difficulty on average, we figure that some teams may benefit from home field advantage more than others do. That is to say, as UEFA leagues or NBA/NFL divisions become more competitive, we predict that the highly anticipated, close matches featuring teams at the top of the standings may have a greater/lesser home field advantage and, thus, will affect the dependent variables. For example, the EPL often features around six teams that are good enough to compete for the league title whereas a league like the Bundesliga may only have one or two. Thus, there are more tightly contested, high-profile matches in the EPL than there are in the Bundesliga, on average. In the NFL, teams are split up in divisions of four

and each team must play each division rival twice, once at home and once on the road. Therefore, with only 16 games and 32 teams, not every team plays against each other and, thus, one team could have a much more difficult schedule than another. In addition, in the NBA, each team plays against each division rival four times, but a team may only play other teams twice throughout the course of the season. Therefore, we predict that the strength of schedule matters for the NFL and NBA as well and must be accounted for appropriately. Following from this, we will run the following regressions to account for differences in league and division:

vii. UEFA: $\ln(avg\_ppg)$- = α_0+
 α_1 *home*$_i$+α_2 *goal_diff*$_i$+
 α_3 *shots_diff*$_i$+
 α_4 *fouls_diff*$_i$+
 α_4 *average_salary*$_i$+ π_i + ε_i

viii. NFL: *Win_persent*$_i$ = β_0+
 β_1 *home*$_i$+β2 yards_diff$_i$+
 β_3 *score_diff*$_i$+
 β_4 *redzone_diff*$_i$
 β_5 *passer_rating_diff*$_i$+
 β_6 *turnover_margin*$_i$+
 β_7 *average_salary*$_i$+
 β_8 *penalties-diff*$_i$+ λ_i + ϑ_i

ix. NBA: *Win_persent*$_i$ = γ_0+
 π_1 *home*$_i$+γ_2 *score_diff*$_i$+
 γ_3 *rebound_diff*$_i$
 γ_4 *assists_diff*$_i$+
 γ_5 *turnover_diff*$_i$+
 γ_6 *steals_diff*$_i$+γ_7 *fouls_diff*$_i$+
 γ_8 *average_salary*$_i$+ μ_i + δ_i

In the UEFA model, we consider league fixed effect which is represented by π_i with the Bundesliga as our reference category. This model allows us to control for any differences in league home field advantage. The λ_i and μ_i variables for both the NFL and the NBA represent the fixed effect for the differing divisions in each respective sport. For the NBA, the Northwest division is our reference category and for the NFL, the AFC East is our reference category.

Nevertheless, our data may still be pooled in these regressions which is why we decide to run a two-way fixed effect model for each sport. In addition to concerns we have about an underlying fixed effect across the

different leagues and divisions of our three sports, we also figured that there may be another fixed effect in the different years we were testing. We are going to run the following regressions to control for this:

x. UEFA: $\ln(avg\_ppg)$- = α_0+
 α_1 *home*$_{it}$+α_2 *goal_diff*$_{it}$+
 α_3 *shots_diff*$_{it}$+
 α_4 *fouls_diff*$_{it}$+
 α_4 *average_salary*$_i$+ π_i + κ_t + ε_{it}

xi. NFL: *Win_persent*$_i$ = β_0+
 β_1 *home*$_i$+β_2 *yards_diff*$_{it}$+
 β_3 *score_diff*$_{it}$+
 β_4 *redzone_diff*$_{it}$
 β_5 *passer_rating_diff*$_{it}$+
 β_6 *turnover_margin*$_{it}$+
 β_7 *average_salary*$_{it}$+
 β_8 *penalties-diff*$_{it}$+ λ_i + ϕ_t + v_{it}

xii. NBA: *Win_persent*$_i$ = β_0+
 π_1 *home*$_i$+β_2 *yards_diff*$_{it}$+
 π_3 *score_diff*$_{it}$+
 π_4 *redzone_diff*$_{it}$
 π_5 *passer_rating_diff*$_{it}$+
 π_6 *turnover_margin*$_{it}$+
 π_7 *average_salary*$_{it}$+
 π_8 *penalties-diff*$_{it}$+ μ_i + σ_t + δ_{it}

In this model, the fixed effects are represented by κ_t, ϕ_t and σ_t respectively and account for the different years in our regression; 2016, 2017, 2018 and 2019. The reference category is 2016.

The COVID-19 pandemic has caused major sports to rethink how they are going to carry out the rest of their seasons. The Serie A, Italian Soccer League was the first sporting body to suggest the idea of playing the rest of the season's games behind closed doors with no fans, effectively nullifying any home field advantages. In light of these new developments, we also measure the causal effect of home field advantage that sports teams receive with no fans in their home stadium. For the UEFA contests played under the COVID-19 regulations, we use the same model as we had before. However, Ligue 1 is not included given that Ligue 1 ended their season in March while the others restarted their seasons in the summer months. We also cosider a two-way fixed effect for this regression. We do not include COVID-19 regressions for the NBA and the NFL since they have yet to start their respective seasons. The NBA plans to continue their season in a bubble in Orlando, thus meaning there are no

home games while UEFA leagues still play games at their home stadiums. Below is our regression for the UEFA COVID-19 model:

xiii. UEFA-COVID-

19: $\ln(avg\_ppg)_{it} - = \Omega_0 +$
$\Omega_1\ home_{it} + \Omega_2\ goal\_diff_{it} +$
$\Omega_3\ shots\_diff_{it} +$
$\Omega_4\ fouls\_diff_{it} +$
$\Omega_5\ average\_salary_{it} + Z_i + \rho_t + \delta_{it}$

In the above two-way fixed effects regression, we included Z_i to represent the fixed effect across only four of our leagues since Ligue 1 was not included in the data set. The Bundesliga is the reference category. The variable ρ_t represents the fixed effect across the different months when matches have taken place. These months are May, June and July; and May is the reference category. The only league that played in May is the Bundesliga. All four leagues played in June and only the Serie A, EPL and La Liga have continued their seasons into July. Even though the final Serie A match took place on Sunday, August 2nd, it will be grouped in with July data because the round of matches started in July.

We also conducted some supplementary models to aid our understanding of the variable relationships at hand. These supplementary regressions are included in the appendix. We have decided to include any regressions that determine home field advantages for each division and each UEFA league. In order to find the effect of home field advantage for each league, we will run a series of multivariate regressions to see the exact impact each UEFA league or division has. In addition, we have included extra regressions that showed relationships between two given variables such as wins and attendance that can also be found in the appendix.

Empirical Results and Analysis

For our models, we used a linear model as we are looking at season totals versus the probability to win one game. If we were looking at the probability to win one game, we would use a tobit model. The first regression we ran in our experiment was a simple bivariate regression depicting the relationship between a team's average points per game or win percentage and their home/away status. The results of this regression can be found in Table 1 for the UEFA, Table 2 for the NFL and Table 3 for the NBA. These three tables show that playing at home has a significant advantage than when playing on the road for the bivariate model. Table 1 represents that one additional UEFA home game leads to a 41.6% increase in the number of points compared to the game on the road. For the NFL (Table 2), the bivariate model displays that playing at home means there is a 13.3% better chance at

winning an additional game when at home than on the road. Lastly, for the NBA (Table 3), there is a 17.2% better chance at winning an additional game at home than on the road. These models indicate that there is a substantial advantage when playing at home. We also ran multivariate regression models for all three of the sports.

We see a drastic change in the home status variable's coefficient in the multivariate models. For every additional UEFA home game, ceteris paribus, the home team has a 5.28% increase in the amount points a team wins for each individual home game versus away game. For the NFL, each additional home game is associated with a 3.2% increase in the chance to win if the game is at home versus away. Both multivariate regressions are significant at the 95% confidence level. However, for the NBA, there is only a 1.5% increase for a team to win an additional home game versus a road game. This metric is not significant at the 95% confidence level. It is clear there is no statistically significant difference between playing at home or the road for the NBA. However, this data seems to be pooled for all three sports since it groups all the UEFA leagues and the divisions for the NBA and NFL together in one group.

In order to account for any fixed effects amongst the different leagues and divisions of our sports to fight endogeneity, we need to incorporate the five UEFA leagues and the NBA and NFL divisions we have defined in our model. To do so, we perform a one-way fixed effect model using the LSDV approach for all three sports. These regressions are captured in all three of the tables. In the fixed effect model for the UEFA leagues, our reference league is the Bundesliga. For the NFL, our reference division is the AFC East and for the NBA, our reference division is the Northwest.

In our one-way fixed effect model for the UEFA, we can see that the coefficient of home field advantage is 0.0473 and is statistically significant at a 95% confidence level. This implies that one additional UEFA home game leads to a 4.7% increase in the number of points earned as opposed to an away game. However, we see no significant differences in home field advantage between the different leagues. While the NBA and NFL do display a significant home field advantage as entire leagues, we found that there is yet a significant difference between the individual divisions in the NFL and the NBA. For the NBA, there is a significant difference between the Northwest division and the Central division at a 95% confidence level. For the NFL, there is a significant difference between the AFC East and the AFC North. These examples illustrate that there are different levels of competitiveness for each division.

We also tested for the year each season was played since Clarke and Norman (1995) observed significant change in home field advantage over time. Therefore, we extended our model to become a two-way fixed effect regression including year dummy variables considering 2016 as the reference category. The last column in Table 1, 2 and 3 demonstrates the results of the two-way fixed effect models.

In controlling for the year, our main explanatory variable coefficients for all three sports exhibited very minimal changes in home field advantage. For example, the UEFA leagues coefficient ranged from 0.0473 to 0.0472. Nevertheless, this metric is still statistically significant at the 95% confidence level. The other home advantage coefficients for the NBA and the NFL had similar impacts when controlling for each individual year. Neither of these coefficients were statistically significant. Because there was no noticeable change in our main explanatory variable ($home_i$) for any of the sports, we can safely say that there is no significant difference in home field advantage for different years with regard to all three sports.

After these four main regressions, we decided to test which UEFA league or NBA/NFL division had the largest home field advantage. To do so, we ran a group of five regressions to represent the five UEFA leagues, eight regressions for the NFL divisions and six regressions for the NBA divisions. Eventually, we found that controlling for the different leagues or divisions does have a significant impact on home field advantage. In testing for division within the NFL especially, we observed that some divisions were statistically significant at the 95% confidence level while others were not. These results can be found under Appendix A, B, C and F.

Table 4 highlights our COVID-19 regression results. We ran bivariate, multivariate, one-way fixed effect and two-way fixed effect models for any contests played under some form of regulation arising from the novel Coronavirus outbreak. The only sport who has done so has been soccer. The five UEFA leagues, except for Ligue 1, do not allow spectators to be present at matches as Ligue 1 discontinued their season. In running these regressions, we can test how important spectators are in the quantification of home field advantage.

Table 4 demonstrates that there is no significant home

Table 1: UEFA regression results

Ln_avg ppg	Bivar.	Multivar.	Fixed Effects (One-way)	Fixed Effects (two-way)
home	0.416***	0.0528***	0.0473***	0.0472***
	(0.0294)	(0.0164)	(0.0168)	(0.0168)
goal_diff		0.517***	0.522***	0.522***
		(0.0147)	(0.0150)	(0.0150)
Shotsdiff		0.0137**	-0.0122*	-0.0122*
		(0.00644)	(0.00650)	(0.00651)
Fouls_diff		-0.0137	0.0133	0.0134
		(0.0108)	(0.0108)	(0.0108)
Average salary		-1.94e-08***	-2.48e-08***	-2.48e-08***
		(5.73e-09)	(6.65e-09)	(6.65e-09)
EPL			0.0291	0.0291
			(0.0227)	(0.0227)
La Liga			-0.000613	-0.000616
			(0.0211)	(0.0211)
Ligue 1			-0.00563	-0.00564
			(0.0212)	(0.0212)
Serie A			-0.00338	-0.00338
			(0.0210)	(0.0210)
2017				-0.0249
				(0.0185)
2018				-0.00803
				(0.0185)
2019				-0.0115
				(0.0185)
Constant	0.0172	0.199***	0.198***	0.209***
	(0.0208)	(0.0105)	(0.0172)	(0.0206)
Obs	784	784	784	784
R^2	0.204	0.843	0.844	0.844

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

field advantage when there are no fans present. fact, there is a slight negative effect when playing at home given that our main explanatory variable is -0.00232, -0.00109 and -0.000412 for our multivariate, one-way

fixed effect and two-way fixed effect models, respectively. When analyzing our two-way fixed effect model, it is practically zero thus alluding to there being no home field advantage at all.

Table 2: NFL regression results

Win_percent	Bivar.	Multivar.	Fixed effects (One-way)	Fixed effects (two-way)
Home	0.133***	0.0320**	0.0299*	0.0299*
	(0.0280)	(0.0162)	(0.0163)	(0.0164)
yards		-0.000187	-0.000150	-0.000150
		(0.000279)	(0.000281)	(0.000283)
score		0.0235***	0.0235***	0.0235***
		(0.00306)	(0.00309)	(0.00311)
redzone		0.01030	0.009580	0.00959
		(0.01390)	(0.01410)	(0.0142)
Passer rating		0.000965	0.000959	0.000960
		(0.000857)	(0.000865)	(0.000871)
turnovers		0.0376**	0.0369**	0.0369**
		(0.01810)	(0.0182)	(0.0183)
average		1.44e-05	4.71e-08*	4.71e-08*
		(0.00589)	(2.53e-08)	(2.54e-08)
penalties		3.67e-08	0.00159	0.00159
		(2.41e-08)	(0.00610)	(0.00614)
AFC North			-0.0647**	-0.0647**
			(0.0294)	(0.0296)
AFC South			-0.0476	-0.0476
			(0.0299)	(0.0301)
AFC West			-0.0458	-0.0458
			(0.0296)	(0.0298)
NFC East			-0.0432	-0.0432
			(0.0296)	(0.0298)
NFC North			-0.0515*	-0.0515*
			(0.0303)	(0.0305)
NFC South			-0.0528*	-0.0528*
			(0.0305)	(0.0307)
NFC West			-0.0343	-0.0343
			(0.0296)	(0.0298)
2017				0.000485
				(0.0207)
2018				0.00292
				(0.0207)
2019				0.00219
				(0.0207)
Constant	0.434***	0.490***	0.533***	0.532***
	(0.0198)	(0.0109)	(0.0230)	(0.0264)
Obs	256	256	256	256
R^2	0.081	0.758	0.764	0.764

Note. Standard errors in parentheses
***Statistical significance at 99% confidence
**Statistical significance at 95% confidence
*Statistical significance at 90% confidence

Table 3. NBA regression results

Win_Percent	Bivar.	Multivar.	Fixed Effects (One-way)	Fixed Effects (two-way)
home	0.172***	0.0150*	0.0126	0.0127
	(0.0202)	(0.00835)	(0.00835)	(0.00840)
score		0.0309***	0.0308***	0.0308***
		(0.00116)	(0.00117)	(0.00117)
rebounds		-0.00154	-0.00118	-0.00119
		(0.00151)	(0.00164)	(0.00166)
assists		-0.00144	-0.000841	-0.000841
		(0.00179)	(0.00182)	(0.00183)
turnover		0.00419	0.00387	0.00387
		(0.00356)	(0.00359)	(0.00362)
steals		0.00291	0.00262	0.00262
		(0.00476)	(0.00483)	(0.00486)
fouls		0.00106	0.00134	0.00134
		(0.00234)	(0.00238)	(0.00240)
average_salary		0.000599	0.000507	0.000507
		(0.000395)	(0.000397)	(0.000400)
Atlantic			0.0107	0.0107
			(0.0121)	(0.0122)
Central			0.0255**	0.0255**
			(0.0118)	(0.0118)
Southeast			0.0108	0.0108
			(0.0120)	(0.0121)
Pacific			0.0182	0.0182
			(0.0122)	(0.0122)
Southwest			0.00676	0.00676
			(0.0125)	(0.0126)
2017				0.000111
				(0.00927)
2018				-0.000154
				(0.00927)
2019				7.88e-05
				(0.00927)
Constant	0.414***	0.492***	0.482***	0.482***
	(0.0143)	(0.00528)	(0.00912)	(0.0108)
Obs.	240	240	240	240
R^2	0.232	0.923	0.925	0.925
	(0.0198)	(0.0109)	(0.0230)	(0.0264)

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Table 4. UEFA Covid-19 regression results

ln_avg_ppg	Bivar.	Multivar.	Fixed Effects (One-way)	Fixed Effects (two-way)
home	0.0857	-0.00232	-0.00109	-0.000412
	(0.0853)	(0.0542)	(0.0544)	(0.0545)
goal_diff		0.429***	0.429***	0.428***
		(0.0227)	(0.0228)	(0.0231)
shots_diff		-0.00292	-0.00351	-0.00387
		(0.00510)	(0.00513)	(0.00516)
fouls_diff		0.0182	0.0174	0.0174
		(0.0237)	(0.0237)	(0.0238)
average_salary		1.54e-09	1.75e-09	2.18e-09
		(1.44e-08)	(1.44e-08)	(1.45e-08)
EPL			0.0551	0.102
			(0.0738)	(0.100)
La Liga			0.0771	0.120
			(0.0715)	(0.0969)
Serie A			0.113	0.161
			(0.0734)	(0.100)
June				-0.0318
				(0.106)
July				-0.0821
				(0.121)
Constant	0.252***	0.151***	0.0881	0.103
	(0.0607)	(0.0375)	(0.0587)	(0.0759)
Obs.	259	259	259	259
R^2	0.004	0.650	0.654	0.655

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

However, there is a limitation as each team has only played eight to eleven matches under COVID-19 regulations instead of the full 38-game season schedule. However, since there is no statistically significant effect from playing at home during these games, we fail to reject the null hypothesis. Therefore, we find that spectators must be in attendance in order for there to be a significant advantage for the home team.

Lastly, since we used winning percentage as the dependent variable for both the NBA and NFL, we ran a regression with winning percentage instead of average points per game for the UEFA leagues to test if it changed our findings. This analysis is located in the appendix. Since we had to exclude any tied matches from this data set, we do not mention its results in our paper given that we believe this test is not as accurate as our preferred model that uses average points per game as the dependent variable.

We hypothesize that these results are due to several factors, two of which are crowd-density and increased TV viewership. The literature has found that dense crowds increase home field population. This, in tandem with the fact that NBA crowds have historically been less densely populated than European soccer crowds, proves that our results could be factual. TV viewership of sports has increased dramatically in the US given its recent increases in HDTV sales which far exceed those of the UK. Perhaps the shift from in-person spectating to television-spectating has decreased home field advantage in America relative to Europe. Again, our findings support this.

When running the models without goal or score differential like Single Reade *et al.*, (2020), we saw no changes in our main coefficients being significant or not which is seen in Appendix O-R. With all of that said, the models we ran (Table 1 - 4) originally hold up.

When analyzing the results of the tables, the fixed-effect model can be used rather than the random-effects model because of the Hausman test not giving us a significant relationship for any of the tables. When looking at Appendix K-N, there is no difference between the two models. Because of this relationship, the fixed effect model can be used. When running the random effects model, we used cross sectional data across the different years. However, we do not see a difference with the fixed effects model.

When it came to the tobit model, we have seen similar results with the coefficient on our main variable of interest, home, not seeing any change as well. This can be seen in Appendix G-J. Therefore, with the coefficient staying the same for all the models. We can see that there is no reason to use the tobit model over the OLS fixed-effect model.

Conclusion

In regard to our findings, we have found that strength of schedule is a determining factor for home field advantage in American sports as both the NBA and NFL have divisions that do have a home field advantage. When analyzing the one-way fixed effect model in Table 2, NFL home field advantage is no longer statistically significant compared to the multivariate regression model. The reason being is that only one of the divisions in the NFL reject the null hypothesis and have a home field advantage; this result can be seen in our appendix Table B and C, while the other seven divisions do not display significant home field advantage. We believe that this division has home field advantage due to the strength of schedule for these two divisions and because there are rivalries within them. Pollard and Gómez (2013) suggest that passionate fans could yield to their team winning more at home. Therefore, in the AFC North, home field advantage would be more prevalent as Stein (2012) reports that the Steelers and the Ravens have one of the biggest rivalries in the entire NFL and these two teams are known to be top tier teams for the past few years. These two elements equate to more passionate fans at each home as even when they do not play against one another, winning can put one team ahead in the standings over the other. Although the NFL and the NBA both do not statistically have a home field advantage as a league, individual divisions still have one.

We also believe that the reason that there is home field advantage in the UEFA leagues compared to that of the NBA and the NFL is due to a high-density crowd causing teams to play better that is supported by Nevill *et al.* (1996) and Kotecki (2014). As shown with Fig. 5, a multitude of NBA teams have low attendance numbers even though NBA stadium sizes have almost the same capacity. Also, for the 2018-2019 season, USA Today (2019) reports that a record number of sellouts for NBA games was recorded which was 760 out of 2460 games. This translates to low density crowds for NBA teams. However, in the EPL, Hoskin (2020) reports that each team averages a stadium capacity of 90% and the top seven teams average well above 99%: Virtually a sell out every game. Given the larger crowd densities in the EPL, we can potentially observe why European leagues exhibit a home field advantage while the basketball teams do not. For the NFL, Fig. 6 displays that almost every team has the same amount of game attendees, thus alluding to a high-density crowd to every game as every NFL stadium is relatively the same size. However, with only 16 games, we hypothesize that teams might be more desperate to win on the road thus causing home field

advantage to not be prevalent in most divisions. For UEFA leagues, they have a high-density crowd at almost every game and they also have 38 games which is more than the NFL's thus meaning the road team is not as desperate to win thus causing them to have home field advantage. In addition, the COVID-19 games are played with no fans which means there is a low-density number of attendees present at a game. Our COVID-19 regression results support the importance of having a high-density crowd for a home game as home field advantage is not statistically significant.

Sharp (2019) states that increased TV viewership because of cheaper HDTVs and increased access to internet streaming services may incentivize sports fans to watch their favorite teams play from home. This technological factor may play a role in calculating home field advantage. This assumption makes even more sense with our data when considering the distribution of HDTVs in both Europe and the UK. While The Benton Institute for Broadband & Society reports that, as of 2015, over 80% of American households own an HDTV (Frayer, 2015), the Statista Research Department claims that only 57% of UK households own an HDTV (Vailshery, 2021). Given the large distribution of HDTVs in the US, it would make more sense that American sports fans watch more games from home than fans from the UK. Perhaps therefore we do not observe statistical significance of home field advantage in American sporting leagues such as the NBA and NFL and we do in European soccer.

Nevertheless, after running numerous tests, including many independent variables and fixed effects, we can conclude that there is enough statistically significant evidence to conclude that playing at home will result in more average points per game than playing away for UEFA soccer teams. However, the same cannot be said for the NFL and the NBA. For our UEFA model, both our one-way fixed-effect and two-way fixed-effect models had a coefficient of 0.047 and a p-value of less than 0.05. This would suggest that there is, in fact, an innate advantage to playing in your own stadium as a team earns 4.7% more points per game at home than on the road. However, for the NBA and NFL, we found that neither of the fixed effect models had enough statistically significant evidence to reject the null hypothesis, thus indicating that home field advantage is not present in those two leagues. We know this to be true because, in our experiment, we ran multiple regressions so that we could eliminate any possible sources of endogeneity in order to receive an unbiased estimator for each model.

In addition to proving the existence of home field advantage, we also were able to draw other, supplementary conclusions from our experiment. We found that each UEFA league does exhibit its own home field advantage. That is to say, the utility each team receives from playing at home differs from league to league as our main explanatory variable, $home_i$, did change from 0.0477 to 0.0529. Therefore, we can conclude that there is a difference between the home field advantage of UEFA's top leagues. Additionally, we found that some divisions in the NFL and NBA exhibit home field advantage while others do not as the AFC North and the Central division both have a p-value that is less than 5%. The rest of the divisions for NBA and NFL do not display this and consequently fail to reject the null hypothesis which indicates that there is no home field advantage. Therefore, home field advantage may not be present in all the divisions and the league as a whole, but it is still present in some of these divisions. We also found that playing in front of no fans for the UEFA leagues because of COVID-19 causes home field advantage to dissipate as it is not statistically significant.

We believe that this research contributes a model to the current literature that not only measures the home field advantage across various sports, but also within the divisional sectors of each sport. In testing across sports, like Gómez *et al.* (2011), we are able to deduce which sports rely more on their fans to contribute home field advantage. What's more, these observed trends in home field advantage indicate room for potential new coaching strategies to be implemented in match preparation. For example, English soccer teams may need to prepare more for away matches and American basketball teams may not need to practice as hard before an away game and can spare some time for rest and recovery. In addition to these strategic recommendations, our paper also suggests potential inmate differences between sport that have yet to be realized. For example, Gómez *et al.* (2011) concludes that Spanish rugby employed a relatively high home field advantage that is likely due to "the continuous, aggressive and intense nature of the sport." More specifically, our model captures the innate differences in home field advantage within divisions as described above. This has not been tested before. Differences in home field advantage with respect to divisional rivalry matches would be of the utmost importance to sporting teams and coaches because they stress the need to prepare more for divisional contests rather than out-of-division games. For example, the Philadelphia Eagles would

benefit from extended practice sessions before playing the Dallas Cowboys more so than they would against the San Diego Chargers regardless of team record, based on the fact that they share a division with the Cowboys. The regressions that were tested to find this relationship incorporate independent variables such as team salaries, referee behavior and in-game statistics that have not been used in the same analysis before. Our study includes an analysis of the soccer games played under the empty stadium regulations during the COVID-19 pandemic. Thus, our experiment can measure any statistically significant difference between playing in front of fans and without fans.

In gathering research and performing our study, we came across some potential limitations of our experiment. First, we did not account for the distance between home and away teams as Pollard *et al*. (2008) and Van Damme and Baert (2019) did in their respective studies. They calculated the distance that away teams travel to play in a match and various other measurements of distances such as language, culture and climate. Including a variable representing some category of distance in our model would potentially increase the legitimacy of our experiment by limiting potential sources of endogeneity. However, we chose to neglect this intuition because we found that there are not very many distances, outside of location, we would incorporate into our model for the NBA and NFL given that both leagues take place within a singular country. Second, we did not account for any historical or political effect surrounding the games. Upon controlling for competitive balance, crowd size and distance traveled, Pollard and Gómez (2013) found that home field advantage is the largest in the Balkan regions of Europe. They suggest that the heightened sense of territoriality among these historically conflict-stricken countries may have an effect on the athletic gameplay of the region. Perhaps the increased passion for their region drives the fans to appear at more matches or even interact with the players more during the games. Political and historical influences were not controlled for in our study which could lead to potential endogeneity. However, we think this would only exhibit a very small, if not negligible, effect in the NBA and NFL given that there has not been much conflict within the USA since the Civil War.

Author's Contributions

Nazmul Islam: Designed the research plan, organized the study and supervised all experiments; coordinated the data-analysis and edited the writing of the manuscript; coordinated the mouse work.

Peter LeRoux and Nicholas Correia: Designed the research plan; participated in all experiments, data-analysis and contributed to the writing of the manuscript; mouse work.

Ethics

This article is original and contains unpublished study. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Agnew, G. A., & Carron, A. V. (1994). Crowd effects and the home advantage. *International Journal of Sport Psychology*. https://psycnet.apa.org/record/1995-03795-001.

Altman, D. (2014). "Are All Red Cards Created Equal?" The New Yorker, 6 Nov. www.newyorker.com/sports/sporting-scene/red-cards-created-equal.

Carmichael, F., Thomas, D., & Ward, R. (2000). Team performance: the case of English premiership football. *Managerial and decision Economics*, 21(1), 31-45. https://onlinelibrary.wiley.com/doi/abs/10.1002/1099-1468(200001/02)21:1%3C31::AID-MDE963%3E3.0.CO;2-Q.

Chatterjee, S., Campbell, M. R., & Wiseman, F. (1994). Take that jam! An analysis of winning percentage for NBA teams. *Managerial and Decision Economics*, 15(5), 521-535. https://onlinelibrary.wiley.com/doi/abs/10.1002/mde.4090150514.

Clarke, S. R., & Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. *Journal of the Royal Statistical Society*: Series D (The Statistician), 44(4), 509-521. https://rss.onlinelibrary.wiley.com/doi/abs/10.2307/2348899

DaSilva, C. (2017). Ranking the importance of every NFL position, from QB to long snapper. *Fox Sports*. https://www.foxsports.com/nfl/gallery/nfl-position-importance-ranking-value-every-player-053117

Dowie, J. (1982). Why Spain should win the world cup. *New Scientist*, 94(10), 693-695. Retrieved from ResearchGate. https://www.researchgate.net/publication/284600911_Why_Spain_should_win_the_World_Cup

Frayer, L. (2015). HDTV Sets Now in Over 80 percent of US households. *Benton Institute for Broadband & Society*. https://www.benton.org/headlines/hdtv-sets-now-over-80-percent-us-households

FIFA.com (2015/2016). Laws of the game. https://digitalhub.fifa.com/m/3f3e15cc1ab8977b/original/datdz0pms85gbnqy4j3k-pdf.pdf

FIFA.com (2018). More than half the world watched record-breaking 2018 world cup. https://www.fifa.com/tournaments/mens/worldcup/2018russia/media-releases/more-than-half-the-world-watched-record-breaking-2018-world-cup

García, M. S., Aguilar, Ó. G., Marques, P. S., Tobío, G. T., & Romero, J. J. F. (2013). Calculating home advantage in the first decade of the 21th century UEFA soccer leagues. *Journal of human kinetics*, 38, 141. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3827749/

Gómez, M. A., Pollard, R., & Luis-Pascual, J. C. (2011). Comparison of the home advantage in nine different professional team sports in Spain. *Perceptual and Motor Skills*, 113(1), 150-156. https://journals.sagepub.com/doi/abs/10.2466/05.PMS.113.4.150-156.

Goodell, R. (2019). Official playing rules of the national football league. https://operations.nfl.com/media/3831/2019-playing-rules.pdf

Greer, D. L. (1983). Spectator booing and the home advantage: A study of social influence in the basketball arena. *Social psychology quarterly*, 252-261. https://www.jstor.org/stable/3033796

Hoskin, R. (2020). Ranking how full each Premier League stadium has been this season. *Give Me Sport*. www.givemesport.com/1537005-ranking-how-full-each-premier-league-stadium-has-been-this-season

Kotecki, J. (2014). Estimating the effect of home court advantage on wins in the NBA. https://digitalcommons.iwu.edu/econ_honproj/124/

NBA Rulebook (2019). NBA.com Retrieved from official.nba.com/rulebook/.

Nevill, A. M., Newell, S. M., & Gale, S. (1996). Factors associated with home advantage in English and Scottish soccer matches. *Journal of sports sciences*, 14(2), 181-186. https://www.tandfonline.com/doi/abs/10.1080/02640419608727700

Omotayo, O. O. (2003). Exploratory study of the home advantage factor in professional soccer players' goal scoring. Kenyatta University Institutional Repository. https://ir-library.ku.ac.ke/handle/123456789/13723

Pollard, R. (1986). Home Advantage in Soccer: A Retrospective Analysis. *Journal of Sports Sciences,* vol. 4, no. 3, 237–248. https://www.tandfonline.com/doi/abs/10.1080/02640418608732122

Pollard, R., & Gómez, M. Á. (2013). Variations in home advantage in the national basketball leagues of Europe. *Revista de Psicología del Deporte*, 22(1), 263-266. https://www.redalyc.org/pdf/2351/235127552035.pdf

Pollard, R., & Gómez, M. A. (2014). Components of home advantage in 157 national soccer leagues worldwide. *International Journal of Sport and Exercise Psychology*, 12(3), 218-233. https://www.tandfonline.com/doi/abs/10.1080/1612197X.2014.888245

Pollard, R., Prieto, J., & Gómez, M. Á. (2017). Global differences in home advantage by country, sport and sex. *International Journal of Performance Analysis in Sport*, 17(4), 586-599. https://www.tandfonline.com/doi/abs/10.1080/24748668.2017.1372164

Pollard, R., Silva, C. D., & Medeiros, N. C. (2008). Home advantage in football in Brazil: differences between teams and the effects of distance traveled. *Revista Brasileira de Futebol (The Brazilian Journal of Soccer Science)*, 1(1), 3-10. https://rbf.ufv.br/index.php/RBFutebol/article/viewFile/6/4

Reade, J., Schreyer, D., & Singleton, C. (2020). Eliminating supportive crowds reduces referee bias. SSRN 3743972. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743972

Schwartz, B., & Barsky, S. (1977). The Home Advantage. Social Forces, vol. 55, no. 3, 641–661. www.researchgate.net/publication/273585873_The_Home_Advantage

Sharp, W. (2019). NFL Home-Field Advantage: It's Not As Important As It Once Was: *Sharp Football. Sharp Football Analysis*. www.sharpfootballanalysis.com/analysis/nfl-home-field-advantage/

Stein, M. (2012). The Biggest Team Rivalries in the NFL. https://bleacherreport.com/articles/1308656-the-biggest-team-rivalries-in-the-nfl/

Trandel, G. A., & Maxcy, J. G. (2011). Adjusting winning-percentage standard deviations and a measure of competitive balance for home advantage. *Journal of Quantitative Analysis in sports*, 7(1). https://www.degruyter.com/document/doi/10.2202/1559-0410.1297/html

USA Today (2019). NBA Sets Record with 760 Sellouts during 2018-19 Season. AP News Department. https://www.usatoday.com/story/sports/nba/2019/04/11/nba-sets-record-with-760-sellouts-during-2018-19-season/39331283/

Van Damme, N., & Baert, S. (2019). Home advantage in European international soccer: Which dimension of distance matters? *Economics*, 13(1). https://www.degruyter.com/document/doi/10.5018/economics-ejournal.ja.2019-50/html

Vailshery, S. L. (2021). Household penetration of smart TV sets in the United Kingdom (UK) from 2014 to 2020. *Statista*. https://www.statista.com/statistics/654074/smart-tv-penetration-in-uk-households/

Waters, A., & Lovell, G. (2002). An examination of the homefield advantage in a professional English soccer team from a psychological standpoint. *Football studies*, 5(1), 46-59. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.612.7047&rep=rep1&type=pdf

Appendices

Appendix A: UEFA league's home field advantage results

Ln_avg_ppg	Bundesliga	EPL	La Liga	Ligue 1	Serie A
home	-0.0213	0.0902**	0.0901*	0.00537	0.0681**
	(0.0431)	(0.0351)	(0.0505)	(0.0342)	(0.0322)
goal_diff	0.553***	0.522***	0.537***	0.459***	0.572***
	(0.0428)	(0.0256)	(0.0432)	(0.0338)	(0.0337)
shots_diff	-0.00740	-0.0350***	-0.0179	0.0100	-0.0134
	(0.0199)	(0.0115)	(0.0228)	(0.0166)	(0.0111)
fouls_diff	0.0302	0.0207	-0.00702	0.0441**	-0.0335
	(0.0304)	(0.0258)	(0.0232)	(0.0221)	(0.0236)
average_salary	-6.50e-08***	8.82e-09	-2.70e-08*	-4.70e-08***	-2.39e-09
	(2.16e-08)	(1.46e-08)	(1.39e-08)	(1.64e-08)	(1.57e-08)
2017	-0.0403	-0.0356	-0.0187	-0.0322	-0.0181
	(0.0458)	(0.0407)	(0.0473)	(0.0358)	(0.0359)
2018	-0.0191	-0.0144	0.0327	-0.0117	-0.0137
	(0.0455)	(0.0410)	(0.0480)	(0.0359)	(0.0362)
2019	-0.0328	-0.00750	0.0562	-0.000816	-0.0311
	(0.0457)	(0.0420)	(0.0499)	(0.0359)	(0.0370)
Constant	0.262***	0.194***	0.156***	0.239***	0.205***
	(0.0391)	(0.0330)	(0.0402)	(0.0305)	(0.0297)
Observations	144	160	160	160	160
R-squared	0.839	0.860	0.817	0.859	0.889

Note. Standard errors in parentheses
***Statistical significance at 99% confidence
**Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix B: NFL AFC divisional home field advantage

Win_percent	AFC east	AFC north	AFC south	AFC west
home	0.0174	0.118**	0.0401	-0.00416
	(0.0594)	(0.0471)	(0.0522)	(0.0513)
yards_diff	-0.000540	0.00195**	0.000626	-0.00150
	(0.00107)	(0.000841)	(0.00115)	(0.00109)
score_diff	0.0171	0.0122	0.0130	0.0328***
	(0.0149)	(0.00891)	(0.00921)	(0.00833)
redzone_diff	0.0337	-0.0840	0.0245	0.0168
	(0.0581)	(0.0496)	(0.0558)	(0.0443)
passer_rating_diff	0.00350	0.00205	0.000101	-0.000352
	(0.00447)	(0.00250)	(0.00337)	(0.00328)
turnover_margin	-0.0240	0.0950	0.0539	0.00778
	(0.0796)	(0.0595)	(0.0695)	(0.0534)
average_salary	4.80e-08	5.69e-08	1.87e-08	1.56e-07
	(8.51e-08)	(7.67e-08)	(1.27e-07)	(1.49e-07)
penalties_diff	0.0207	0.00750	0.0318*	-0.0166
	(0.0280)	(0.0218)	(0.0171)	(0.0189)
2017	-0.00585	0.101	-0.0249	-0.0527
	(0.0729)	(0.0651)	(0.0653)	(0.0644)
2018	-0.00908	0.0702	0.0275	-0.00583
	(0.0757)	(0.0589)	(0.0656)	(0.0596)

Appendix B: Continue

2019	-0.00163	0.0572	0.0394	0.0149
	(0.0731)	(0.0610)	(0.0731)	(0.0676)
Constant	0.546***	0.372***	0.454***	0.517***
	(0.0644)	(0.0521)	(0.0530)	(0.0531)
Observations	32	32	32	32
R-squared	0.803	0.890	0.791	0.816

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix C: NFL NFC divisional home field advantage

Win_Percent	NFC East	NFC North	NFC South	NFC West
home	0.0591	0.0130	0.0231	0.0229
	(0.0539)	(0.0633)	(0.0582)	(0.0484)
yards_diff	-0.000173	0.00173	-0.00217	-0.000350
	(0.000874)	(0.00116)	(0.00150)	(0.000780)
score_diff	0.0245**	0.00460	0.0468***	0.0160
	(0.00960)	(0.0180)	(0.00964)	(0.00974)
redzone_diff	0.00982	0.0986*	0.0377	0.0348
	(0.0527)	(0.0534)	(0.0483)	(0.0558)
passer_rating	0.00203	-0.00201	0.00113	0.00160
	(0.00276)	(0.00345)	(0.00222)	(0.00248)
turnover_margin	0.00191	0.151	-0.120	0.0881*
	(0.0658)	(0.0933)	(0.0973)	(0.0484)
average_salary	1.14e-07	-4.74e-08	1.38e-08	1.34e-07
	(1.04e-07)	(1.35e-07)	(5.73e-08)	(1.04e-07)
penalties_diff	-0.00908	0.0175	0.00294	-0.0183
	(0.0216)	(0.0311)	(0.0272)	(0.0182)
2017	-0.0514	0.0657	0.0470	0.0312
	(0.0667)	(0.0760)	(0.0567)	(0.0636)
2018	-0.00839	-0.0202	0.00346	0.0655
	(0.0668)	(0.0758)	(0.0601)	(0.0612)
2019	-0.0763	0.102	0.0704	0.0774
	(0.0710)	(0.0845)	(0.0729)	(0.0622)
Constant	0.509***	0.470***	0.457***	0.448***
	(0.0551)	(0.0643)	(0.0456)	(0.0530)
Observations	32	32	32	32
R-squared	0.817	0.754	0.821	0.860

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix D: UEFA win percentage regression results

win_percent	Bivariate	Multivariate	Fixed effect (One-Way)	Fixed effect (Two-Way)
home	0.164***	0.0168***	0.0168***	0.0168***
	(0.0126)	(0.00634)	(0.00634)	(0.00633)
goal_diff		0.206***	0.206***	0.206***
		(0.00564)	(0.00564)	(0.00563)
shots_diff		-0.00329	-0.00331	-0.00330
		(0.00245)	(0.00245)	(0.00245)
fouls_diff		0.00325	0.00347	0.00334
		(0.00406)	(0.00406)	(0.00405)
average_salary		7.63e-09***	7.64e-09***	7.63e-09***
		(2.53e-09)	(2.53e-09)	(2.53e-09)
EPL			0.00408	0.00408
			(0.00792)	(0.00791)
La Liga			-0.00117	-0.00116
			(0.00792)	(0.00791)
Ligue 1			-0.0115	-0.0115

Appendix D: Conutinue

			(0.00792)	(0.00791)
Serie A			0.00248	0.00248
			(0.00792)	(0.00791)
2017				0.0138**
				(0.00695)
2018				0.00649
				(0.00695)
2019				0.000211
				(0.00695)
Constant	0.294***	0.368***	0.369***	0.364***
	(0.00892)	(0.00402)	(0.00658)	(0.00784)
Observations	784	784	784	784
R-squared	0.177	0.875	0.876	0.877

Note. Standard errors in parentheses

***Statistical significance at 99% confidence

**Statistical significance at 95% confidence

*Statistical significance at 90% confidence

Appendix E: Bivariate relationship of attendance on winning percentage for the NBA and NFL and percentage of average points received each game for UEFA

	NBA	NFL	UEFA
Attendance	0.0323***	0.00297	0.00889***
	(0.00815)	(0.00190)	(0.000837)
Constant	0.00713	0.366***	0.163***
	(0.146)	(0.129)	(0.0293)
Observations	120	128	392
R-squared	0.118	0.019	0.225

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix F. NBA divisional home field advantage

Win_Percent	Southwest	Pacific	Northwest	Southeast	Central	Atlantic
home	0.0170	-0.0336*	0.0714**	0.0312	0.0630***	-0.0335
	(0.0237)	(0.0173)	(0.0293)	(0.0229)	(0.0177)	(0.0226)
score_diff	0.0313***	0.0286***	0.0298***	0.0327***	0.0271***	0.0312***
	(0.00284)	(0.00419)	(0.00330)	(0.00271)	(0.00384)	(0.00300)
rebounds_diff	0.00647	-0.00375	0.000977	-0.00636	-0.00574	0.00188
	(0.00505)	(0.00406)	(0.00477)	(0.00460)	(0.00349)	(0.00426)
assists_diff	-0.00478	0.00251	-0.00754	-0.00483	0.00186	0.00397
	(0.00659)	(0.00510)	(0.00541)	(0.00526)	(0.00722)	(0.00393)
turnover_diff	-0.0116	-0.00527	0.0210*	0.0108	0.0110	-0.0121
	(0.0121)	(0.0101)	(0.0108)	(0.00969)	(0.00695)	(0.00895)
steals_diff	-0.00370	-0.0208*	0.0263	-0.0244	-0.000277	-0.00923
	(0.0165)	(0.0115)	(0.0162)	(0.0149)	(0.0117)	(0.0119)
fouls_diff	0.00159	0.00320	-0.0122	-0.00161	0.00705	0.0184**
	(0.00727)	(0.00571)	(0.00832)	(0.00451)	(0.00663)	(0.00808)
average_salary	0.00246*	0.00106	0.000340	0.00111	0.000499	-0.00198
	(0.00141)	(0.00123)	(0.00111)	(0.00102)	(0.000967)	(0.00184)
2017	-0.00247	-0.00608	0.0372	0.00522	-0.0117	0.0354
	(0.0250)	(0.0206)	(0.0231)	(0.0226)	(0.0212)	(0.0243)
2018	-0.000384	-0.00954	0.0508*	-0.0400	0.00901	0.0282
	(0.0261)	(0.0215)	(0.0292)	(0.0244)	(0.0235)	(0.0262)

Appendix F: Continue

2019	-0.00767	0.00808	0.0518*	-0.0175	-0.0265	0.0211
	(0.0263)	(0.0208)	(0.0300)	(0.0229)	(0.0229)	(0.0256)
Constant	0.487***	0.517***	0.415***	0.498***	0.495***	0.507***
	(0.0207)	(0.0185)	(0.0263)	(0.0186)	(0.0167)	(0.0219)
Observations	40	40	40	40	40	40
R-squared	0.943	0.969	0.935	0.920	0.947	0.958

Note. Standard errors in parentheses

***Statistical significance at 99% confidence

**Statistical significance at 95% confidence

*Statistical significance at 90% confidence

Appendix G: UEFA covid tobit model

ln_avg_ppg	Bivariate	Multivariate	Fixed Effect (One-way)	Fixed Effects (Two-way)
home	0.0857	-0.00232	-0.00109	-0.000412
	(0.0850)	(0.0536)	(0.0534)	(0.0534)
goal_diff		0.429***	0.429***	0.428***
		(0.0224)	(0.0224)	(0.0226)
shots_diff		-0.00292	-0.00351	-0.00387
		(0.00504)	(0.00504)	(0.00505)
fouls_diff		0.0182	0.0174	0.0174
		(0.0234)	(0.0233)	(0.0233)
average_salary		1.54e-09	1.75e-09	2.18e-09
		(1.42e-08)	(1.42e-08)	(1.42e-08)
EPL			0.0551	0.102
			(0.0725)	(0.0979)
La Liga			0.0771	0.120
			(0.0702)	(0.0948)
Serie A			0.113	0.161
			(0.0721)	(0.0979)
June				-0.0318
				(0.103)
July				-0.0821
				(0.118)
var(e.ln_avg_ppg)	0.468***	0.164***	0.162***	0.162***
	(0.0411)	(0.0144)	(0.0143)	(0.0142)
Constant	0.252***	0.151***	0.0881	0.103
	(0.0604)	(0.0370)	(0.0577)	(0.0742)
Observations	259	259	259	259

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix H: UEFA tobit model

ln_avg_ppg	Bivariate	Multivariate	Fixed effect	Two way fixed effect
home	0.416***	0.0486***	0.0486***	0.0472***
	(0.0293)	(0.0166)	(0.0166)	(0.0167)
goal_diff		0.521***	0.521***	0.522***
		(0.0148)	(0.0148)	(0.0148)
shots_diff		-0.0126*	-0.0126*	-0.0122*
		(0.00645)	(0.00645)	(0.00645)
fouls_diff		0.0135	0.0134	0.0134
		(0.0107)	(0.0107)	(0.0107)
average_salary		-2.35e-08***	-2.35e-08***	-2.48e-08***
		(6.44e-09)	(6.44e-09)	(6.60e-09)
EPL			-0.00304	-0.00304
			(0.0209)	(0.0209)
La Liga			-0.00819	-0.00819
			(0.0209)	(0.0209)

Appendix H: Continue

Ligue 1			0.00404	0.00404
			(0.0209)	(0.0209)
Serie A			-0.00313	-0.00313
			(0.0209)	(0.0209)
2017				-0.0241
				(0.0183)
2018				-0.000772
				(0.0185)
2019				0.00250
				(0.0187)
var(e.ln_avg_ppg)	0.168***	0.0331***	0.0331***	0.0330***
	(0.00851)	(0.00167)	(0.00167)	(0.00167)
Constant	0.0172	0.201***	0.203***	0.209***
	(0.0207)	(0.0106)	(0.0173)	(0.0205)
Observations	784	784	784	784
R-squared	0.204	0.843	0.844	0.844

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix I: NBA tobit model

Win Percent	Bivariate	Multivariate	Fixed Effect	Two Way Fixed Effect
home	0.172***	0.0150*	0.0126	0.0127
	(0.0202)	(0.00809)	(0.00810)	(0.00810)
score_diff		0.0309***	0.0308***	0.0308***
		(0.00113)	(0.00113)	(0.00113)
rebounds_diff		-0.00154	-0.00118	-0.00119
		(0.00149)	(0.00160)	(0.00160)
assists_diff		-0.00144	-0.000841	-0.000841
		(0.00176)	(0.00176)	(0.00176)
turnover_diff		0.00419	0.00387	0.00387
		(0.00349)	(0.00349)	(0.00349)
steals_diff		0.00291	0.00262	0.00262
		(0.00467)	(0.00469)	(0.00469)
fouls_diff		0.00106	0.00134	0.00134
		(0.00229)	(0.00231)	(0.00231)
average_salary		0.000599	0.000507	0.000507
		(0.000388)	(0.000385)	(0.000385)
Atlatnic			0.0107	0.0107
			(0.0117)	(0.0117)
Central			0.0255**	0.0255**
			(0.0114)	(0.0114)
Southeast			0.0108	0.0108
			(0.0117)	(0.0117)
Pacific			0.0182	0.0182
			(0.0118)	(0.0118)
Southwest			0.00676	0.00676
			(0.0121)	(0.0121)
2017				0.000111
				(0.00893)
2018				-0.000154
				(0.00894)
2019				7.88e-05
				(0.00893)
var(e.win_percent)	0.0244***	0.00246***	0.00239***	0.00239***
	(0.00222)	(0.000224)	(0.000218)	(0.000218)
Constant	0.414***	0.492***	0.482***	0.482***
	(0.0143)	(0.00518)	(0.00885)	(0.0104)
Observations	240	240	240	240
R-squared	0.232	0.923	0.925	0.925

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix J: NFL tobit model

Win percent	Bivariate	Multivariate	Fixed effect	Two way fixed effect
Home	0.133***	0.0320**	0.0299*	0.0299*
	(0.0278)	(0.0159)	(0.0158)	(0.0158)
yards_diff		-0.000187	-0.000150	-0.000150
		(0.000274)	(0.000272)	(0.000272)
score_diff		0.0235***	0.0235***	0.0235***
		(0.00301)	(0.00299)	(0.00299)
redzone_diff		0.0103	0.00958	0.00959
		(0.0137)	(0.0136)	(0.0136)
passer_rating_diff		0.000965	0.000959	0.000960
		(0.000842)	(0.000838)	(0.000838)
turnover_margin		0.0376**	0.0369**	0.0369**
		(0.0177)	(0.0176)	(0.0176)
average_salary		1.44e-05	4.71e-08*	4.71e-08*
		(0.00579)	(2.45e-08)	(2.45e-08)
penalties_diff		3.67e-08	0.00159	0.00159
		(2.37e-08)	(0.00591)	(0.00591)
AFC North			-0.0647**	-0.0647**
			(0.0285)	(0.0285)
AFC South			-0.0476	-0.0476
			(0.0290)	(0.0290)
AFC West			-0.0458	-0.0458
			(0.0287)	(0.0287)
NFC East			-0.0432	-0.0432
			(0.0287)	(0.0286)
NFC North			-0.0515*	-0.0515*
			(0.0293)	(0.0293)
NFC South			-0.0528*	-0.0528*
			(0.0296)	(0.0296)
NFC West			-0.0343	-0.0343
			(0.0287)	(0.0287)
2017				0.000485
				(0.0200)
2018				0.00292
				(0.0200)
2019				0.00219
				(0.0200)
var(e.win_percent)	0.0496***	0.0131***	0.0128***	0.0127***
	(0.00439)	(0.00115)	(0.00113)	(0.00113)
Constant	0.434***	0.490***	0.533***	0.532***
	(0.0197)	(0.0107)	(0.0223)	(0.0254)
Observations	256	256	256	256
R-squared	0.081	0.758	0.764	0.764

Note. Standard errors in parentheses
***Statistical significance at 99% confidence
**Statistical significance at 95% confidence
*Statistical significance at 90% confidence

Appendix K: UEFA random/fixed effect

ln_avg_ppg	Random	Fixed
home	0.049***	0.047***
	(0.017)	(0.017)
goal_diff	0.521***	0.522***
	(0.015)	(0.015)
shots_diff	-0.013*	-0.012*
	(0.006)	(0.006)
fouls_diff	0.013	0.013
	(0.011)	(0.011)

Appendix K: Continue

average_salary	-0.000***	-0.000***
	(0.000)	(0.000)
Constant	0.201***	0.202***
	(0.011)	(0.011)
Observations	784	784
R-squared		0.844
Number of year	4	4

Note. Standard errors in parentheses

***Statistical significance at 99% confidence

**Statistical significance at 95% confidence

*Statistical significance at 90% confidence

Appendix L: NBA random effects and fixed

Win_percent	Random	Fixed
home	0.015*	0.015*
	(0.008)	(0.008)
score_diff	0.031***	0.031***
	(0.001)	(0.001)
rebounds_diff	-0.002	-0.002
	(0.002)	(0.002)
assists_diff	-0.001	-0.001
	(0.002)	(0.002)
turnover_diff	0.004	0.004
	(0.004)	(0.004)
steals_diff	0.003	0.003
	(0.005)	(0.005)
fouls_diff	0.001	0.001
	(0.002)	(0.002)
average_salary	0.001	0.001
	(0.000)	(0.000)
Constant	0.492***	0.492***
	(0.005)	(0.005)
Observations	240	240
R-squared		0.923
Number of year	4	4

Note. Standard errors in parentheses ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix M: Covid-19 UEFA Random Effect and Fixed Effect

ln_avg_ppg	Random	Fixed
home	-0.002	-0.001
	(0.054)	(0.054)
goal_diff	0.429***	0.427***
	(0.023)	(0.023)
shots_diff	-0.003	-0.003
	(0.005)	(0.005)
fouls_diff	0.018	0.018
	(0.024)	(0.024)
average_salary	0.000	0.000
	(0.000)	(0.000)
Constant	0.151***	0.151***
	(0.037)	(0.038)
Observations	259	259
R-squared		0.643
Number of month	3	3

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix N: NFL random and fixed effects

Win_percent	Random	Fixed
home	0.032**	0.032*
	(0.016)	(0.016)
yards_diff	-0.000	-0.000
	(0.000)	(0.000)
score_diff	0.023***	0.023***
	(0.003)	(0.003)
redzone_diff	0.010	0.010
	(0.014)	(0.014)
passer_rating_diff	0.001	0.001
	(0.001)	(0.001)
turnover_margin	0.038**	0.038**
	(0.018)	(0.018)
penalties_diff	0.000	0.000
	(0.006)	(0.006)
average_salary	0.000	0.000
	(0.000)	(0.000)
Constant	0.490***	0.490***
	(0.011)	(0.011)
Observations	256	256
R-squared		0.758
Number of year	4	4

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix O: UEFA Covid: No Goal Differential

ln_avg_ppg	Bivariate	Multivariate	Fixed effect	Two way fixed effect
home	0.086	-0.002	-0.001	-0.000
	(0.085)	(0.054)	(0.054)	(0.055)
shots_diff		-0.003	-0.004	-0.004
		(0.005)	(0.005)	(0.005)
fouls_diff		0.018	0.017	0.017
		(0.024)	(0.024)	(0.024)
average_salary		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
2.league			0.055	0.102
			(0.074)	(0.100)
3.league			0.077	0.120
			(0.071)	(0.097)
5.league			0.113	0.161
			(0.073)	(0.100)
2.month				-0.032
				(0.106)
3.month				-0.082
				(0.121)
Constant	0.252***	0.151***	0.088	0.103
	(0.061)	(0.037)	(0.059)	(0.076)
Observations	259	259	259	259
R-squared	0.004	0.650	0.654	0.655

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix P: UEDA No Goal Differential

ln_avg_ppg	Bivariate	Multivariate	Fixed Effect	Two way fixed effect
home	0.416***	0.049***	0.049***	0.047***
	(0.029)	(0.017)	(0.017)	(0.017)
shots_diff		-0.013*	-0.013*	-0.012*
		(0.006)	(0.006)	(0.007)
fouls_diff		0.013	0.013	0.013
		(0.011)	(0.011)	(0.011)
average_salary		-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)
2.league			-0.003	-0.003
			(0.021)	(0.021)

Appendix P: Continue

3.league			-0.008	-0.008
			(0.021)	(0.021)
4.league			0.004	0.004
			(0.021)	(0.021)
5.league			-0.003	-0.003
			(0.021)	(0.021)
2017.year				-0.024
				(0.019)
2018.year				-0.001
				(0.019)
2019.year				0.003
				(0.019)
Constant	0.017	0.201***	0.203***	0.209***
	(0.021)	(0.011)	(0.017)	(0.021)
Observations	784	784	784	784
R-squared	0.204	0.844	0.844	0.844

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Q: NFL no score differential

Win percent	Bivariate	Multivariate	Fixed effect	Two way fixed effect
home	0.133***	0.032**	0.030*	0.030*
	(0.028)	(0.016)	(0.016)	(0.016)
yards_diff		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
redzone_diff		0.010	0.010	0.010
		(0.014)	(0.014)	(0.014)
passer_rating_diff		0.001	0.001	0.001
		(0.001)	(0.001)	(0.001)
turnover_margin		0.038**	0.037**	0.037**
		(0.018)	(0.018)	(0.018)
average_salary		0.000	0.000*	0.000*
		(0.000)	(0.000)	(0.000)
penalties_diff		0.000	0.002	0.002
		(0.006)	(0.006)	(0.006)
2.division			-0.065**	-0.065**
			(0.029)	(0.030)
3.division			-0.048	-0.048
			(0.030)	(0.030)
4.division			-0.046	-0.046
			(0.030)	(0.030)
5.division			-0.043	-0.043
			(0.030)	(0.030)
6.division			-0.052*	-0.052*
			(0.030)	(0.030)
7.division			-0.053*	-0.053*
			(0.031)	(0.031)
8.division			-0.034	-0.034
			(0.030)	(0.030)
2017.year				0.000
				(0.021)
2018.year				0.003
				(0.021)
2019.year				0.002
				(0.021)
Constant	0.434***	0.490***	0.533***	0.532***
	(0.020)	(0.011)	(0.023)	(0.026)
Observations	256	256	256	256
R-squared	0.081	0.758	0.764	0.764

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence

Appendix R: NBA no score differential

Win percent	Bivariate	Multivariate	Fixed effect	Two way fixed effect
home	0.172***	0.015*	0.013	0.013
	(0.020)	(0.008)	(0.008)	(0.008)
rebounds_diff		-0.002	-0.001	-0.001
		(0.002)	(0.002)	(0.002)
assists_diff		-0.001	-0.001	-0.001
		(0.002)	(0.002)	(0.002)
turnover_diff		0.004	0.004	0.004
		(0.004)	(0.004)	(0.004)
steals_diff		0.003	0.003	0.003
		(0.005)	(0.005)	(0.005)
fouls_diff		0.001	0.001	0.001
		(0.002)	(0.002)	(0.002)
average_salary		0.001	0.001	0.001
		(0.000)	(0.000)	(0.000)
1.division			0.011	0.011
			(0.012)	(0.012)
2.division			0.026**	0.026**
			(0.012)	(0.012)
3.division			0.011	0.011
			(0.012)	(0.012)
5.division			0.018	0.018
			(0.012)	(0.012)
6.division			0.007	0.007
			(0.012)	(0.013)
2017.year				0.000
				(0.009)
2018.year				-0.000
				(0.009)
2019.year				0.000
				(0.009)
Constant	0.414***	0.492***	0.482***	0.482***
	(0.014)	(0.005)	(0.009)	(0.011)
Observations	240	240	240	240
R-squared	0.232	0.923	0.925	0.925

Note. Standard errors in parentheses. ***Statistical significance at 99% confidence. **Statistical significance at 95% confidence. *Statistical significance at 90% confidence