

Referee Home Bias in English Premier League

by

Juan Arboleda

Submitted to the Department of Economics

School of Natural and Social Sciences

In partial fulfillment of the requirements

For the degree of Bachelor of Arts

Purchase College

State University of New York

May 20, 2022

First Reader: Iuliia Chikish

Second Reader: Sanford Ikeda

Abstract

Referee bias is thought to contribute to home advantage, especially in soccer. Recent research has shown that sports where referees have power to influence the outcome of the match tend to experience a greater home advantage. Referees' decisions may be influenced by crowds, but little work has been done to examine how referee behavior would change when crowds are not present in stadiums since it was impossible to isolate crowds from stadiums. To analyze how crowds in stadiums influence the decision making of referees, I examine 1,428 English Premier League (EPL) official matches during the seasons 2018/19 through 2021/22, including ghost matches induced by Covid19. The results show that referees are responsible for some of the home advantage in the EPL, meaning there was a clear advantage for the home team over the away team before Covid. But during the Covid phase, the home team's winning rate dropped by ten percent and referees' pressure to make decisions decreased and awarded more yellow cards to the home team than before Covid. I hypothesize that referees respond differently to factors such as pressure and I suggest further research at referees' behavior within the different top European leagues.

Table of Contents

1. Introduction4
2. Literature Review5
3. COVID-19 Restrictions9
4. Data.....11
5. Results14
6. Conclusion.....15
7. References17

Introduction

In team sports, the term home advantage is used to describe the benefit that home teams gain over visiting teams. The existence of home advantage has been well documented in many sports. Gómez et al. (2011) finds that soccer shows a greater home advantage rate than all other sports, followed by rugby and basketball. While the lowest home advantage values were in baseball, golf, cricket, and American football. The evidence of home advantage can be found across all soccer leagues.

The researchers are trying to identify the reasons behind the home advantage. One of the leading theories – is the referee's bias toward the home team. In other words, when crowds are presented at stadiums, they present a type of threat to the decision making of referees that results in given the advantage to the home team. Sutter et al. (2003) provides a great anecdote: in the season of 2000/1 of the German League (Bundesliga), where in the last game of the season, the 34th game, after the regular 90 minutes Schalke 04 was the leader of the league and they would have been crowned the champions if the referee from the parallel match had not given three minutes of extra time in which Bayern München scored against Hamburger SV, making Bayern München the champion for the third time in a row in the German football league. In these circumstances, those three minutes added to that game lead to Bayern München to become the champion on that season and making it their third time winning the championship in a row.

Another great case of home bias is presented by Buraimo et al. (2009). The authors provided a quote stated by Sir Alex Ferguson as a great example for this so-called bias refereeing. "It was like playing against 12 men" Sir Alex Ferguson stated after the performance of referee Herbert Fandel following Manchester United's loss of 2-1 for the Champions' League Match away to Roma in April 2007. The paper states that fans, players and head coaches worldwide often

complain for what they call an inconsistent application of the rules and an alleged bias against their team by referees.

Sutter *et al.* (2003) & Garicano *et al.* (2005) find the evidence that referees call less fouls against the home team and add more time if the home team is losing the match. Crowds pressure is something that affects referees' decisions, and referees experienced this pressure through the crowd screaming or whistling when they think that they made a "bad call" or a call against the home team, this affects the referee's future calls. In many cases referees are threaten by fans in the middle of a game so this puts a lot of pressure in their shoulder when making decision that involves the home team being in disadvantage.

To investigate how the absence of crowds during the COVID-19 restrictions affected the behavior of referees, I follow up game-level data in the English Premier League (EPL). Also, it was nearly impossible to make these types of studies that tested referee bias before COVID-19, since it was impossible to isolate crowd's noise from stadiums. Now with the presence of a pandemic that has allowed soccer to have ghost matches, where stadiums are empty and isolated from fans

Literature Review

According to Buraimo *et al.* (2009), in professional soccer, referees follow certain rules by the Laws of Association Football, and these are determined by the Fédération Internationale de Football Association (FIFA). Using data from the English Premier League and the German *Bundesliga* investigate the potential bias in awarding of players disciplinary sanction, in the forms of 'yellow' cards that come in the context of soccer are cautions that can turn into a 'red' card that means the dismissal of a player from the pitch. They were able to find that what may

appear as referee favoritism can actually be an excessive or illegal aggressive behavior from players in teams that are behind in score.

Buraimo et al. (2009) analyzed that the *Bundesliga* stadiums unlike the Premier League stadiums have a running track that separates the crowd from the pitch and the reason for this investigation is to find out if this structural difference of stadiums have any effect on referees' decision making during the game. When a track separates the crowd from the pitch, referees will feel less pressured by crowds and will penalized players by awarding yellow and red cards. In football, referees determine how much time should be added at the end of each half, they do this by stopping their clocks immediately if they think that a player is injured and needs to have treatment in the field or when there's substitutions happening.

An influential paper by Garicano *et al.* (2005) presented evidence from the Spanish top division, called "La Liga," that referees awarded less added time after the 90 minutes of regular time in games where the home team was ahead in the scoreboard and added more time when the home team was behind in the scoreboard.

Dohmen (2008) hypothesizes that crowd pressure on referees will most likely decrease if there's a running track present in the infrastructure of the stadium. Also, the author affirms that referees will feel less influenced when they decide to award a yellow or red cards to the home teams when a running track separates the field from the stands (crowd). The author found that stoppage time added at the end of the game depended on the score difference and on the identity of the team. Which means that the referee would systematically award more stoppage time to the home team in close games. Also, the author was able to find that when the match is decided at the end of the regular 90 minutes, when a team is ahead in the scoreboard by two or more goals, that the amount of stoppage time does not depend on the team being the home team or the away team.

This means that the crowds have a weaker incentive to influence the referee on how much time should be added for stoppage time.

Sutter and Kocher (2004) also find the evidence of a systematic home bias of referees. They studied the behavior of soccer referees in the German Bundesliga by examining their decisions when awarding penalties and extra time at the end of each half of each match. They were able to find that referees award less extra time only if it was clear that the home team was going to win after the 90 minutes. Also, the results indicate that if the home team was ahead by a goal or it was a draw after the 90 minutes, the extra time was significantly shorter (about 30-50 seconds) than if the home team was behind by a goal.

To test the hypothesis that referees' decisions are affected by the crowds pressure one would need to compare the behavior of referees with and without crowds present. It wasn't too easy to conduct this kind of experiment in the pre-Covid time, but there were a couple of attempts. First, Nevil et al. (2002) conducted an experiment in which they showed a video-tape of 47 tackles from the English Premier League to 40 qualified referees, in which they were asked to classify these tackles as regular or irregular. They divided the referees into two groups: group one was shown the videotape with the sound of the crowd and group two was shown the videotape without sound. The results indicate that the group of referees that watched the videos with sound was more opposed to classify tackles for the home teams as irregular by about 15% and this group showed more uncertainty in their decision. Also, the decisions made by the group of referees who heard the noise of the crowd were similar to those made by the original match referees. In the pre-covid era, this is how these types of experiments regarding home bias were done since it was impossible to isolate to crowds from the stadiums. In these experiments you would have treatments and control groups like in the case of this experiment those treatments and control groups were the

factors of sound and no sound and the two groups of referees who watch the videos in the two different ways.

Second, Pettersson-Lidbom and Priks (2010) used a natural experiment in which fans were banned from some Italian stadiums. On February 2, 2007, supporters from two Italian football clubs, Calcio Catania and Palermo Calcio, clashed with each other and the police was involved in acts of hooligan violence. These riots forced the Italian government to banned fans from stadiums and teams played their home games without spectators. The authors took advantage of this situation and did an experiment using data from Serie A and Serie B for the season 2006/2007, they would be evaluating the hypothesis that referees may be biased due to social pressure. The season consists of 842 games and excluding the 21 games that were played without spectators. Their results show that the home team is punished less than the away team in games that have fans. In contracts, the home is punished more than the away team in games without spectators.

The Covid-19 restrictions provided a unique opportunity to test the referee bias hypothesis, because the spectators were limited or banned for some period of time. A couple of studies exploits that natural variation. Endrich et al. (2020) investigated ghost matches that were influence by Covid-19 in the Bundesliga. Considering that in the German's professional football league attracts an average crowd of 41,000 and 20,300 in the top two divisions, to a sudden change where teams are forced to play with empty stadiums for the safety of everyone. Their findings indicate that home teams received on average about 0.6 fouls fewer than visiting teams in the pre-Covid-19 period, this suggests a clear referee bias for home teams. During ghost matches, the number of fouls for the home team increased by about one. Similar pattern is present for the number of yellow cards awarded by referees.

Ramchandani and Millar (2021) proposed in their paper to examine whether playing matches behind closed doors had an effect on football clubs competing throughout Europe to be benefitted from a quantifiable home advantage when crowds weren't present in stadiums. The study incorporates five of the top football leagues in Europe, and about 506 fixtures were played with spectators across these leagues during the season of 2019/20. The authors conducted analysis of inter-season comparing home advantage between the 2019/20 season and the 2018/19 season. Their main findings were that the Italian Serie A and the German Bundesliga were the only two leagues that show any evidence of declining in the inter-season home advantage. But overall, they argued that there isn't enough evidence to argue that the absence of fans in stadiums affects home advantage in football.

To assess the importance of referee bias in determining how referees are influenced by crowds, this paper contributes to the home advantage literature by examining how referees react to the variation of environment during an official match, where stadiums can have thousands of people to having an empty stadium. Before, it was nearly impossible to isolate referees from crowds. COVID-19 restrictions allowed for the natural experiment which I am going to use in this study.

COVID-19 Restrictions

At the beginning of the COVID-19 outbreak in 2020, the Premier League football was severely affected to the point where the tournament had to be suspended. According to Wikipedia¹, on March 13th of 2020, The Football Association (FA), the English Football League and the FA

¹ https://en.wikipedia.org/wiki/2019-20_Premier_League#Effects_of_the_COVID-19_pandemic

Women's Super League, on a meeting took an unanimously decision to suspend the professional football in England until April 4th of 2020. But this suspension was extended, on March 19th of 2020, to at least April 30th of 2020.

The Premier League started a project called "Project Restart" in the late April, which aimed to resume playing the 92 remaining matches over a period of only six weeks in neutral stadiums. On May 17th and 18th around 748 players and staff members of the clubs were tested for COVID-19. Stage 1 of this project was the "Return to Training Protocol", which involved small group players training starting May 19th, 2020². These groups were to be no more than five players in each training session and the sessions were no longer than 75 minutes for each player. Stage 2 of the project began on May 27th, allowing clubs in the Premier League to start contact training.

On May 28th, many clubs agreed that the season should resume on June 17th, with the two games in hand of Manchester City vs Arsenal and Aston Villa vs Sheffield United. Then to continue with the first full round of matches since March to start on the weekend of June 19-21, but to be played behind closed doors. Further it was announced that the remaining of the 92 matches will be televised and to improve the accessibility of domestic broadcasts. All matches were televised for free through British Broadcast Corporation (BBC), Pick (a British free-to-air television channel), Amazon Prime, and Twitch.

The pandemic brought many changes to Premier League, one of those changes were that the EPL announced that teams would be allowed to have a nine men roster as substitutes per match, instead of having the usual of seven men, also teams were allowed to use up to five substitutes per match instead of three.

² <https://www.premierleague.com/news/1682374>

According to Deloitte, the suspension of the EPL for about three months estimated that clubs in the league would face a £1 billion loss of revenue for the 2019-20 season and £500 million that would be accounted for broadcasters and matchday revenue. Before restarting the games, the Premier League put together a set of guidelines to be followed at all games. Other than having the matches to be behind closed doors, they wanted to keep a maximum of 300 people inside the stadium at the time of the game. Stadiums were divided into three zones that consist of: red zone which includes pitch and technical areas, yellow zone which are the stands and green zone which is the stand concourses, and there were restrictions on who was allowed to be in each zone.

Data

The data collected comes from the website FBREF.com; which has football stats, history statistics, scores and history for men's and women's club and national team competitions. I collected the data from the English Premier League. This data is organized by team name and opponent team (home and away team), scores, date, attendance, main referee name, probability of goals per team, stadium, yellow and red cards. At the time of the collection of the dataset was from season 2018/19 to almost the end of season 2021/22, in which there were 1,428 official matches. Then, I had to expand data to the point that it gave me a closer look to the referee decisions (such as yellow and red card awarded to each team in the dataset). This would leave me with a total of 2,856 observations.

Table 1: Summary Statistics					
Variable	Observations	Mean	Std. Dev.	Min	Max
<i>Attendance</i>	2,856	26,177	22467.57	0	81,332
<i>Close</i>	2,856	0.303	0.459	0	1
<i>Yellow</i>	2,856	1.662	1.256	0	7
<i>Red</i>	2,856	0.058	0.242	0	2
<i>Home Won</i>	2,856	0.3901	0.488	0	1

Table 1 shows the summary statistics of all the dataset collected. There are five variables in this summary statistics, in which there are 2,856 observations. I was able to find that the average attendance was 26,177 during almost four years of data gathered. The most attendance recorded in a single match was of 81,332. Yellow cards are an important factor to determine home advantage and with this summary statistics I found that yellow cards being awarded had an average of 1.66 and having seven yellow cards awarded in one game. In the case of red cards, they were few that were given to have reached an average of 0.05. Home team winning had an average of 0.39, meaning that the home team won about 39% of games.

Table 2 breaks down observations on the home team by Covid indicator which takes value equal to zero before the Covid restrictions and takes value one during. Before covid there were 668 matches that were recorded, in which the home team had an average of 46% winning rate. Now when comparing the 760 matches during covid, the home team only had a 40% winning rate. Meaning that the rate at which the home team was winning decrease by 6%. In the case of attendance, there's a huge gap where before Covid the average of people attending the stadiums were of 38,670. Rather than just 15,197 during the Covid period. Yellow cards also slightly decreased for the home team by around 0.02.

Table 2: Summary Statistics by Groups							
-> Home = 1,	Covid	=	Before Covid				
Variable			Observations	Mean	Std. Dev.	Min	Max
<i>Attendance</i>			668	38,670	16457.38	9,980	81,332
<i>Close</i>			668	0.311	0.463	0	1
<i>Yellow</i>			668	1.609	1.282	0	7
<i>Red</i>			668	0.052	0.222	0	1
<i>Home Won</i>			668	0.464	0.499	0	1
-> Home = 1,	Covid	=	During Covid				
Variable			Observations	Mean	Std. Dev.	Min	Max
<i>Attendance</i>			760	15,197	21290.71	0	73,458
<i>Close</i>			760	0.295	0.456	0	1
<i>Yellow</i>			760	1.584	1.212	0	6
<i>Red</i>			760	0.0513	0.227	0	2
<i>Home Won</i>			760	0.4026	0.491	0	1

Table 3 breaks down observations for the visiting team by Covid indicator. Before Covid the away team had a winning rate of 32%. Meanwhile, during the Covid restrictions their winning rate increase by 5 percent. Meaning that they had a winning rate of 37 percent during Covid. Looking at the average of yellow cards for the visiting team before and during Covid, it is noticeable that there is a bigger impact than the average of yellow cards awarded to the home team before and during covid. Before Covid, the away team were awarded an average of 1.83 cards per game. During Covid, they were awarded on average 1.63 cards per game. Meaning that yellow cards given to the away team dropped by 0.20.

Table 3: Summary Statistics by Groups							
Away = 0,	Covid	=	Before Covid				
Variable			Observations	Mean	Std. Dev.	Min	Max
<i>Attendance</i>			668	38,670	16457.38	9,980	81,332
<i>Close</i>			668	0.311	0.463	0	1
<i>Yellow</i>			668	1.838	1.265	0	6
<i>Red</i>			668	0.064	0.252	0	2
<i>Away Won</i>			668	0.321	0.468	0	1
Away = 0,	Covid	=	During Covid				
Variable			Observations	Mean	Std. Dev.	Min	Max
<i>Attendance</i>			760	15,197	21290.71	0	73,458
<i>Close</i>			760	0.295	0.456	0	1
<i>Yellow</i>			760	1.634	1.257	0	7
<i>Red</i>			760	0.066	0.264	0	2
<i>Away Won</i>			760	0.372	0.484	0	1

Results

The results from this research are presented in Table 4. The first column shows the estimated coefficients for the win indicator as the dependent variable. The coefficient on the Home variable is positive and significant. It indicates that home team wins on average 14.06% more often than the visiting team, which constitutes the evidence of the home advantage. The negative and significant coefficient on the interaction term suggests that during Covid restrictions the home advantage decreases by 10.7% with respect to matches won. The plausible explanation for this effect is that crowds had a positive effect on the home team.

Columns 2 and 3 of Table 4 show the results for yellow and red cards as dependent variables. The negative and significant coefficient on Home variable shows that home team on average gets 0.23 yellow cards less per game than visiting team, which is the evidence of home advantage. The negative and significant coefficient on Covid variable indicates that referees awarded 0.47 less

yellow cards per game during the Covid restrictions. The positive and significant coefficient on the interaction term implies that referees awarded more yellow cards to the home team during the Covid restrictions, most likely because of the absence of crowd pressure. This just shows how crowd pressure can affect the decisions made by referees. All estimated coefficients for the red cards as dependents variable are insignificant, probably due to the rarity of the event.

Table 4: Regression Results			
	Wins	Yellow Cards	Red Cards
<i>Home</i>	0.1406***	-0.2306***	-0.0120
	(0.0243)	(0.0652)	(0.0132)
<i>Covid</i>	0.0679*	-0.4747***	-0.0093
	(0.0412)	(0.1106)	(0.0223)
<i>Home*Covid</i>	-0.1070***	0.1818**	-0.0030
	(0.0334)	(0.0893)	(0.0180)
<i>Observations</i>	2,856	2,856	2,856
<i>R-squared</i>	0.1932	0.1272	0.0418

Conclusion

In this paper I have addressed the different types of forms in which home advantage can take place during a soccer match. The term home advantage can be identified as a benefit that the home teams gain over the visiting teams. This benefit is an advantage that is due to the psychological impact that fans have on athletes or referees. One of those is when we investigate how much can crowds in stadiums influence the decision making of referees, especially because referees can see this as a threat to their decisions that results in giving an advantage to the home team. Another great example of this bias is when referees add more time at the end of regular time when the home

team is losing. So under the circumstances how much can crowd in stadiums influence the decision making of referees.

The analysis of data from the seasons 2018/19 to 2021/22 of the English Premier League have shown that referees are much likely to bias the home team. As shown in the paper before Covid, it was clear that the home teams had a bigger winning percentage rate over the visiting teams. This was due to crowds being in stadiums and putting pressure on referees. However, when matches were forced to be played behind closed doors due to Covid, meaning stadiums were separately from crowds, the home team started to win less. This meant that referees did not have as much pressure and were able to make decisions more freely. It is important to notice that during Covid compared to before Covid the home team started to win less and were awarded more yellow cards than before Covid, this is due to the important role that the crowds play in the soccer match. Additionally, this means that the visiting team started to win more and were awarded fewer yellow cards during Covid than before Covid.

It can be concluded that crowds seem to play an important role in referee decisions and the assumptions about home advantage are reasonable. I believe that my findings of home advantage and bias decisions made by referees are important indicator of how the performance of the home team will develop. For future inquiry on this topic, there would be the need to investigate more leagues, not just one league. It would be good to use the top five European leagues to keep studying the role that crowds play in stadiums and how they are able to influence the referees.

References

Buraimo, Babatunde, David Forrest, and Robert Simmons. "The 12th man?: refereeing bias in English and German soccer." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173.2 (2010): 431-449.

Dohmen, Thomas J. "The influence of social forces: Evidence from the behavior of football referees." *Economic inquiry* 46.3 (2008): 411-424.

Endrich, Marek, and Tobias Gesche. "Home-bias in referee decisions: Evidence from "Ghost Matches" during the Covid19-Pandemic." *Economics Letters* 197 (2020): 109621.

Garicano, Luis, Ignacio Palacios-Huerta, and Canice Prendergast. "Favoritism under social pressure." *Review of Economics and Statistics* 87.2 (2005): 208-216.

Gómez, Miguel A., Richard Pollard, and Juan-Carlos Luis-Pascual. "Comparison of the home advantage in nine different professional team sports in Spain." *Perceptual and motor skills* 113.1 (2011): 150-156.

Nevill, Alan M., Nigel J. Balmer, and A. Mark Williams. "The influence of crowd noise and experience upon refereeing decisions in football." *Psychology of sport and exercise* 3.4 (2002): 261-272.

Pettersson-Lidbom, Per, and Mikael Priks. "Behavior under social pressure: Empty Italian stadiums and referee bias." *Economics Letters* 108.2 (2010): 212-214.

Ramchandani, Girish, and Robbie Millar. "Investigating the "Twelfth Man" Effect in Five European Domestic Football Leagues: A COVID-19 Induced Natural Experiment." *Journal of Global Sport Management* (2021): 1-15.

Sutter, Matthias, and Martin G. Kocher. "Favoritism of agents—the case of referees' home bias." *Journal of Economic Psychology* 25.4 (2004): 461-469.