The influence of fans on the performance of the home team in football

Name student: Ben Mantel

Student ID: 544998

Supervisor: Prof. Dr. Dirk Schindler

Second Assessor: Prof. Dr. Aksel Erbahar

Date final version: 26-02-2023

Abstract

This paper aims to investigate whether the fans of the home team positively influence the performance of their team. Data of 538 different matches in the highest Dutch league over two seasons were used to analyse several factors which can possibly influence the outcome of a football match. In the first season, the spectators were still allowed to go to the stadium. In the majority of the second season, the spectators were not allowed to go to the stadium due to the pandemic. I find that not all of the factors that could possibly lead to a home advantage, have a significant and positive effect on the result of the home team. A significant positive effect of home advantage was found on referee decisions, while no significant effect of home advantage on the outcome of the match was found.

Table of Contents:

Chapter 1: Introduction

Thousands of years ago humans started to introduce sports and adding them into their daily activities. Sports became a major and an essential factor in people's lifestyles. Sports have proven to be highly valued in economics and entertainment. Sports are about winning, and this competitiveness attracts people to come and watch the games. Every city has its own sports club, and every club has its own fans. This varies between the best club in the world, and a local club in a tiny village. Sports bring people together and they unite the crowd of supporters. The collective support gives a team the power to push to a win.

So, why am I looking into home advantage and what is affected by the possible removal of home advantage? Football is scheduled in a way that every team plays one game at home and one game away against a certain team. By playing every team both at home and away, the home advantage is cancelled out as much as possible. Due to Covid-19, people were banned from the stadium for a while. Thus, the pandemic gives the unique opportunity to test whether the home advantage changed without any fans in the stadiums.

As mentioned before, due to the pandemic the stadiums were empty. So, clubs miss out on revenues of selling tickets. The economic relevance is bigger than only the missing out on selling tickets. The food and drinks in the stadium during the game and the merchandise sales before the game are some examples. On the other hand, clubs might earn more money on tv rights. The thousands of supporters who are usually watching in the stadium, are now watching at home. Clubs wants to maximize the home advantage when they play at home. By maximizing home advantage, clubs are aiming to get better results. Better results can lead to higher prize money and tv rights and an increase in selling of merchandise. That is one of the reasons why some clubs increase their stadium capacity; to get as much of the described home advantage as possible and to yield higher revenues.

For years, people came to football stadiums with the intention of supporting their favourite club. The so-called home advantage is considered an important factor in the result of the game.
The fans can impact the behaviour of the referees, players and coaches (Pollard, 2008). Every city or village has its own sports team with supporters. Be it the best team in the world or a local team playing in an amateur division, fans have the unique ability to push their home team to a win. We cannot imagine sports without fans anymore.

Around March 2020 this all changed, when the Covid-19 outbreak occurred. Due to the enormous infection rate, many countries went into lockdown. At first, many sports games were cancelled. Later on in the Covid-19 outbreak, the games resumed but fans were not allowed to enter stadiums and watch their favourite team play. The players felt a lack of atmosphere without any fans in the stadium, so they created an artificial crowd with sounds like "cheering" and "booing".

I want to research if the home advantage has disappeared during the Covid periods, when comparing football results pre-Covid and during Covid. The main question in this paper is:

"What is the influence of fans on the performance of the home team in football?"

The lockdowns and the empty stadiums provide a perfect situation to research the pre-Covid and during Covid differences. What I expect before performing any test is a decrease in the home wins and that the referee has less influenceable decisions. This, because the positive effect of the crowd could no longer be enjoyed by players.

This research will focus on the Eredivisie, the top football division in the Netherlands. The data that has been used in this paper is from the 2019/2020 and the 2020/2021 season and partly the most recent season of 2021/2022. The matches played in the 2019/2020 are all played with spectators. In the 2020/2021 almost all matches were played behind closed doors, but a few games were played where the minority of the stadium has been filled. For the most recent season, we only take the games played without spectators. This, to improve the reliability of the research.

The structure of the thesis is as follows. The second chapter will discuss empirical evidence on related literature of previous studies. Hereafter, the data of this research will be discussed including the collected data sources. After this, the methodology section will present the measures of home advantage and the regression formula. Subsequently, the results of the regression will be shown and discussed. Finally, the last chapter will give a conclusion on the thesis and some suggestions for further research.

Chapter 2: Literature review

Home Advantage

Everyone has a certain idea about home advantage, but it is difficult to provide a definition of the phrase. In the study of Nevill and Holder (1999), home advantage can be measured by three factors: Crowd, learning and travel factors. Several studies investigated those factors and the crowd aspect is the most dominant. The crowd factor is how the fans in the stadium affects the results of the home team. Learning factors are being familiar with the local conditions when playing at home. Travel factors only play a role in big countries, because travelling to away games takes more time compared to smaller countries.

According to the study of Barsky and Schwartz (1977), one always performs better whilst receiving support compared to when we have to rely on our own resources. This research also concludes that 'crowd' is the most important factor in home advantage.

The research of Pollard (2008) gives the most comprehensive definition. "Home advantage in football has long been established as an important factor in determining the result of a game. Its existence is certain to affect the attitude of players, coaches, referees, fans and the media alike". Home advantage is a worldwide phenomenon and does exist since the start of organized football in the end of the 19th century. Home advantage is caused by the factors in the study of Nevill and Holder, and the consequences of those factors are the aspects in the study of Pollard.

Home advantage has been researched in multiple different sports competitions. Other previous studies have also studied the effect of the home team on performances. They have shown that the home advantage is difficult to measure and consists of multiple factors. One of those studies is the one by Carron and Courneya (1992). This study is followed by the one of Bray, Carron and Loughhead (2005). The factors that they measured in those studies are mainly based on travel, crowd and rules. The main insights in these studies are that some of the factors influences the home team advantage more than other factors. These studies focus more on explaining the term, while this study will research the home advantage on teams playing in the Eredivisie.

In other soccer competitions like the highest English division, the home advantage is already measured. The study of Allen and Jones (2012) claims that the colours of the kit can affect the home advantage.

Teams wearing red are more successful and teams with lower ability have a higher home advantage. The study of Clarke and Norman (1995) have shown that the clubs playing in London have lower home advantage than clubs playing elsewhere. The study of van Ours and Peeters (2021) is one of the most recent studies and continues with the Pollard (1986) and the Clarke and Norman (1995) studies. They measure relative and absolute home advantage.

Pandemic

There are not many studies before the pandemic (2020) on home advantage. It might be harder to measure before the pandemic than during or after the pandemic, because it was much more difficult to measure crowd effects. One of the studies that did measure home advantage in soccer before the pandemic is the one by Van de Ven (2011). He provided evidence of a home advantage existing for teams playing in their own stadium. However, he did find out that it does not makes a difference whether the crowd is present. That is why the pandemic can be seen as a perfect opportunity to compare the stadiums with and without spectators. Home advantage caused by crowd effects can be measured easier and that also resulted in more studies on this topic in the last two years.

One of most recent studies is a study by Matos, Monteiro, Antunes, Mendes, Botas, Clemente and Amaro (2021). They measured home advantage during Covid-19. Matos et al. (2021) compared the matches at the beginning of the season with crowd, with the matches at the end of the season without spectators. They used data of the first division in Portugal. They used two methods to measure home advantage. The first method is by Pollard (1986) and home advantage is measured by taking the total home points won expressed as a percentage of the total points won. The second method is by Amaro, Matos and Pollard (2019) and takes the difference between home and away points and express this in a percentage of total away points. The study by Matos et al. (2021) concluded that there was no significant effect of spectators on home advantage. However, in the study of Bilalic, McCarrick, Neave and Wolfson (2020) they found that playing without spectators negatively affected the home team's performance. One of the reasons was the fact that the home team's attacking opportunities decreased without a crowd. They researched all European leagues that finished their 2019/2020 without spectators.

Stadium familiarity and travelling

Another possible cause of home advantage is crowd presence and stadium familiarity. In the paper of Ponzo and Scoppa (2018), they research the effect of crowd attendance in stadium derbies. Stadium derbies remove the effect of travel fatigue and familiarity with the stadium, so it mainly focusses on the home team having a greater proportion of fans during the game. They have found that crowd support influences the result significantly in favour of the home team. The home team scores on average 0.45 more goals than the away team and the probability of winning for the home team is about 13 percent points higher.

In the study of Pollard (2002), he looks at the effect of familiarity with the stadium for home teams. He concludes that on average, 25% of the home advantage is due to the familiarity with the stadium. He compared the last games in the old stadiums for several clubs, with the first games in the new stadium.

Pollard (2002) managed to do this for different sports and different clubs in sport leagues. He found out that for the first games in the new stadium, the home advantage was significantly lower than for the last games in the old stadium. He concludes that this comes from fans and players whose do not have familiarity with the new stadium yet. The study by Van de Ven (2011) has found that players feel more confident and less anxious before the game when playing at home compared to playing on the road. He finds that in the case of familiarity having a significant impact, stadium derbies would not differ in the pre-game state of players.

That is because the familiar environment does not differ in preparations before the game. Van de Ven (2011) did not manage to research this, so he just took this as an assumption. The study of Pollard (1986) researched the home advantage in the 20th century. This study concluded that the travel factor and the crowd factor are relatively unimportant. It did not mention the effect of referees and the local conditions.

Continuing on (stadium) derbies, Pollard and Seckin (2008) researched the effect on football teams located in Istanbul. They researched the derbies and found that the percentage of the home team winning in derbies, was lower than the percentage of the home team winning in the rest of the season.

This is not only due to stadium familiarity, but also to the tension of the game, the average level of both teams and the pressure from external circumstances.

In the Ponzo and Scoppa (2018) study, as mentioned above, stadium derbies neutralize almost all factors that can play a role in home advantage. Except the influence of seat reservations for home supporters. The effect of home team wins in derbies and home team goals is slightly lower in these derbies than in normal matches. This implies that the role of crowd support is more important in this paper, but the travel fatigue in non-stadium derbies also influences the outcome.

To continue with travel distance, Pollard (2008) found some evidence for its effect. The home advantage in European football matches is greater than in national matches. This is caused by the larger distance between the two clubs. He also mentioned the significant lower percentage in home team wins in derbies, compared to regular league games. An older study of Pollard (1986) could not find any evidence of travel distance yet. Clarke and Stefani (1992) did find evidence on the insights of Pollard. They have found that home teams have a significant higher chance of winning in European games compared to national league games. This is mainly caused by the shorter travel distance abroad.

Referees

Another possible cause of home advantage can be due to referee biases. The study by Bryson et al. (2021) focussed on refereeing decisions on performances during Covid-19. The study used data of more than 20 different professional football competitions and used the ordinary least squares (OLS) model for the match outcomes. The study has shown that the absence of the crowd did not influence the outcome of the matches. However, the absence of a crowd caused a significant reduction in the average yellow cards per game for the away team compared to the home team. He concluded that the reduction in yellow cards is an explanation for the reduction in crowd pressure towards the referee, when stadiums are empty compared to the presence of a crowd.

A study by Endrich and Gesche (2020) also researched the referee decisions in empty stadiums. They used the ordinary least squares model to estimate the absence of crowd on yellow cards and the amount of given fouls. They have researched that the away team is punished less when the crowd is absent, compared to stadiums filled with people. The paper has shown that home team is awarded more fouls in empty stadiums.

Moreover, the number of yellow cards given to the home team increased significantly with an absence of the crowd. They concluded the same as Bryson (2021): The referees are affected by the supporters. So, they perform under social pressure.

More papers researched the referee biased based on yellow and red cards. A study by Dawson, Dobson, Goddard and Wilson (2007) found out that the away playing team receives more yellow and red cards compared to the home team. Their data is from several years in the English top league. They conclude that the difference in receiving cards is due to the referees favouring the home teams. The study by Memmert and Unkelbach (2010) researched the correlation between crowd noises and referee decisions. The main finding was that loud audience reactions led to a higher probability of receiving yellow cards. The correlation with home advantage is that the majority of spectators in the stadium are supporters of the home team. So, they indirectly affect the referee when booing or whistling after a foul of the opposite team for example.

The paper of R.H. Boyko, A.R. Boyko and M.G. Boyko (2007) also researches the effect of home advantage on referee decisions. More experienced referees are less affected by the social pressure in stadiums in comparison with less experienced referees. This phenomenon is caused by a psychological mechanism in referees. The crowd size correlates positively with the referees' decisions in favour of the home teams.

Research gap

Football teams were always used to play in front of their own fans. The restrictions during the pandemic could give a different view on home advantage. Previous papers have researched home advantage on different sports competitions, but home advantage has not been researched yet in the absence of a crowd. The study of Pollard (2008) researched the crowd effects on home advantage, while the study of Dawson et al. (2007) researched the effect of crowd on referee decisions solely based on cards awarded by the away team. This paper will show if there is a significant effect of crowd on home advantage by looking for a change in certain variables in the absence of a crowd. Previous studies only used data in the presence of a crowd, therefore this study can possibly give some new insights.

This paper aims to contribute to the currently existing knowledge. As indicated, some topics surrounding the research to home advantage in football games are still untouched.

By providing statistics and information on the relationships between home-crowds, players and referees, the relevance of this study can be felt in both economic and social aspects. This study hopes to provide clarity on social relations between parties involved in football matches. Next to theoretical relevance, this paper will have practical implications. Based on the outcomes of this study, I can draw practical conclusions on future behaviour for professional football organizations. For example, the effects of home crowds on outcomes of games will indicate whether investments by these organizations should be made towards their fanbase or perhaps more towards technical staff and players. In turn, those decisions can lead to favourable athletic and economic outcomes for clubs.

Chapter 3: Data

3.1: Data set

For this research, data is used of the Dutch league (Eredivisie) from games during the 2019/2020 and the 2020/2021 season. The data sources to retrieve this data are: *Transfermarkt.nl* and *football-data.co.uk.* Both websites are designed for obtaining data of football results in the past and the net values of players. The data set contains 540 matches played over two seasons in the Eredivisie.

In the 2019/2020 season, all matches were played in stadiums full of people. In the following season, 2020/2021, almost every match was played behind closed doors. Only a few games were played with a restricted crowd. The teams in both seasons are exactly the same, due to rules of the Dutch Football Association. So, no relegation or promotion to or from the second league.

The 540 matches are played by 18 different teams and are judged by a total of 24 referees. Logically, every match contains two teams, a home team and an away team. The home team plays in their own stadium in front of their own fans, and the opposite for the away team. Approximately 50% of the matches (272) were played behind closed doors. All of them were in the 2020/2021 season. The difference in measurement before and during the Covid break is the dummy variable *stadium.* Stadium will become 1 if there are people allowed in the stadiums and 0 if the stadiums are not filled with people. So, before Covid 1 and during Covid a value of 0.

Unlike other competitions abroad, the 2019/2020 season ended suddenly and did not restart. In other countries the competitions restarted a few months later without a crowd, but the Eredivisie stopped due to Covid and was never allowed to resume the last matches of the season. The average attendance of the games played in the 2019/2020 is approximately 16.000 attendees per match.

3.2. Descriptive Statistics

Table 1: Descriptive match statistics in the Eredivisie seasons 2019-2020 and 2020-2021.

	With spectators	Without spectators	Difference
Home team win %	50,9%	40,2%	26,2%
Away team win %	27,2%	33,3%	-18,3%
Mean home points	1,73	1,46	0,27
Mean away points	1,03	1,26	-0,23
Mean attendance*	16,087	0	16,087
Mean market value home**	35,8	32,4	3,4
Mean market value away	34,1	32,6	1,5
Matches	268	272	4***
Teams	18	18	0****
Home yellow	408	347	61
Away yellow	485	341	144
Home red	20	22	2
Away red	35	29	6
Home goals	400	345	55
Away goals	317	390	73

*Notes: Column 1 is with spectators and is before Covid-19. Column 2 is without spectators and is during Covid-19. Column 3 are the differences between column 1 and column 2. This can be in relative and absolute numbers. * Attendance in 1000s (people). ** Market value in millions (euros). *** Difference in absolute numbers. **** Exactly the same teams in both seasons.*

Table 1 shows some descriptive statistics before and during Covid-19, and the differences between those values. In the first row, one can see that the win percentage of home teams decreases when the spectators are not allowed to go to the stadium. The home team won around half of the matches before Covid-19, and this number decreased with circa ten percent point during Covid-19. The away team win percentage moves the opposite way. Less influence of the spectators of their opponent, gives the away team on average better results (more points). The teams are exactly the same in both seasons and the matches played have a very small difference. Table 1 shows a decrease in goals scored by the home team during Covid-19 and a significant increase in goals scored by the away team. Moreover, the awarded yellow and red cards for the away team decreased and the awarded red cards for the home team increased. This suggests a reduction in home advantage when the matches are played behind closed doors.

Table 2: Differences in sample means between matches played with crowd and behind closed doors in the Eredivisie seasons 2019-2020 and 2020-2021.

	Mean Difference
HG	0.271*
AG	-0.196*
HY	0.201*
AY	0.491**
HR	-0.012
AR	0.005
HP	0.273**
AP	-0.229*
GD	0.468**
HW	0.262**

*Notes: The tests are probit tests and the significance is categorized as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The data description is given in Table 1.*

Table 2 shows the difference in sample means between the two Eredivisie seasons, with crowd and played behind closed doors. The matches played with fans were 26,2% more likely to end in a win for the home team. In those matches, the away team scored more goals and gained more points when the games were played without a crowd.

The home team were awarded more red cards during the pandemic, without the fans. The points gained at home and the goal difference are significant higher when there was a crowd.

3.3 Methodology

To research the influence of fans on the performance of the home team, I will perform a probit regression. The dependent variable will be home advantage. The definition or the dependent (**Y**) variable consists of several other variables.

A variable that will be included in the dependent variable is the referee variable. This includes the yellow and red cards awarded for the home and the away team, before and during Covid-19. The research of Clarke and Norman (1995) also mentioned the referee bias. They looked at the home advantage and if the referees are biased due to this home advantage. The hypothesis we have in this research is:

H1: Yellow and red cards awarded to the away team are decreasing significantly during Covid-19 and vice versa for the home team, due to home advantage.

The variables of home goals (**HG**) and away goals (**AG**) are measured in absolute numbers. The majority of the teams have played two times, both away and at home against the other teams. A single team has played a minimum of three times against any certain other team and a maximum of four times. The assumption is that the results at home are better than the results away, due to the home advantage. The assumption that the home team always wins does not hold. In terms of all goals cumulated, the difference between home goals and away goals should be positive. The hypothesis is:

H2: Due to the home advantage, the home team scores less goals during the pandemic and the away team scores more than before the pandemic.

The variable of home and away points is in line with the variable above. There is a difference between goals scored and points gained, but if the difference between home goals and away goals is positive, one may assume that the average of home points (**PH**) gained is higher compared to away goals (**PA**). I will measure the home and away points gained with and without the crowd. The assumption is that the mean average in absolute numbers for home points with crowd will decrease when the fans are removed. Consequently, this means that the mean average for away teams with crowd will increase when the fans are not allowed in

the stadium and vice versa. On top of that, I also take the goal difference in total into account.

The dummy variable covid (**Covid**) is an independent variable and it indicates whether a match is played behind closed doors or with an audience. The attendance (**ATT**) will measure the crowd attendance in thousands of people.

The first control variable I will include is the market value per match (**MVH/MVA)**. Every match contains two teams and twenty-two players, so every team has eleven players. In this paper, I am measuring the market value of the starting eleven players per team, per match. All the market values of the teams are coming from www.transfermarkt.nl. Having a higher market value as a team compared to your opponent does not guarantee a win, but it increases the probability of winning. This makes the variable more interesting to see if the lower market value teams have better results at home relatively compared to the higher market value teams, taking away games also into account. I assume that the advantage of having a higher value in a team's starting eleven, increases the probability of winning. Moreover, I know there are a lot of other factors which can play a role, especially crowd. When I remove the spectators, one of the important factors is removed and we expect market value of a team to weigh heavier in the probability of winning. The market values will be measured in millions of euros.

The market values are expected to have a greater impact during the pandemic, because the power of the fans is removed. The home advantage effect due to crowd is removed, so the other variables are expected to have a greater impact on the result of the matches.

The other control variable I will use is another referee variable. A total of 22 different referees in the pre-Covid season and 24 different referees during the Covid-19 season, whistled all the matches (separately). Although, the referees have been assisted by the linesmen and the fourth official, I only include the referees. The referee variable is the average career penalties awarded by the referee (**P**). The paper of Schwarz (2011) has already mentioned the penalties awarded by referees and even concluded that a referee who gives a penalty during the game, is more likely to give another penalty during that same match. The statistic about penalties awarded measures, or is the closest approximation, to the ability of a single referee to handle pressure and tension when the home advantages of the crowd increases, as covered in the paper of Page and Page (2010).

The formula is as follows:

$$Y_{hijk} = \beta 1 ATT_{hijk} + \beta 2 Covid_{hijk} + MVH_h + MVA_i + P_j$$

Chapter 4: Results

The form of regression I use is the Probit model. By using the Probit model, the fitted dependent variables can be interpreted as fitted probabilities. I take the logarithm of all the variables except from the dummy variables. In the first table, I take attendance as our main independent variable. I want to estimate the influence of spectators on the outcome of the game, home advantage. I take the logarithm of attendance and in this first table we do not add any control variables. I also take the logarithm of all the dependent variables except for Home wins.

Table 3A: Probit regression taking the logarithm of the dependent variables and the logarithm of attendance as independent variable.

	HG	AG	HY	AY	HR	AR	GD	HP	AP	HW
Attendance	0.254**	-0.164	-0.103	0.042	0.891	0.139	0.066	0.155	-0.143	0.234**
	(0.116)	(0.120)	(0.109)	(0.107)	(1.012)	(0.456)	(0.142)	(0.125)	(0.149)	(0.101)
Constant	-2.029*	1.484	1.243	-0.049	-10.166	-2.846	-0.180	-0.908	1.500	-2.164**
	(1.096)	(1.133)	(1.026)	(1.012)	(9.885)	(4.277)	(1.351)	(1.188)	(1.391)	(0.950)
Obs	218	176	210	225	20	33	138	194	129	267
Pseudo R^2	0.0171	0.0077	0.0032	0.0005	0.1297	0.0061	0.013	0.0066	0.0053	0.0149

Notes: The stars indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The columns exists of dependent variables. The first row is the main independent variable, attendance. The second row is the constant term of the different probit regressions. The third row is the number of observations in regressions and the fourth row is the R-squared.

As mentioned above, all the dependent variables are in the logarithm form. However, I did not take the logarithm of the Home wins variable. The home win variable is a dummy, where the value is "1" when the home team wins the match and the value is "0" when the home team loses the match or the result ends in a draw. Dummies are binary variables, so I do not take the logarithm of the variable. I can see a significance in the home goals variable and the home wins variable. The coefficient for home goals in the first column indicates that when the attendance increases by 1%, the goals scored by the home team increases with 0.254%. These values are percentual changes, rather than changes in absolute values, since I am working with logarithms. The value of the constant term is negative, so the increase in goals scored by the home team starts from a certain number of spectators.

The value of home wins in the sixth column increases with 0.234, when the attendance increases by 1%. The same applies for the constant term as I could see for the goals scored. The term is negative, so the increase in home wins starts from a certain number of attendances. The value of the increase in home wins is not in percentage, since I did not take the logarithm of the variable.

The coefficients of away goals, away points and home yellow cards have a negative sign. The negative sign implies that when the attendance increases with 1%, the variables in question decrease with a certain percentage. However, those variables are not significant. Although I can see in several papers, for example in the Pettersson-Lidbom and Priks (2010) one that the number of spectators does influence the referee decisions and especially the reduced yellow cards awarded by the home team.

To further investigate the influence of spectators on home advantage, I will use two different tables. I will take all the different variables into account for the probit regression. The difference between the two tables is the two variables attendance and Covid. The attendance during Covid was always 0, so I will use one table with the variable attendance and one table with the variable Covid.

Table 3B: Probit regression taking the logarithm of the dependent variables and independent variables. The independent Covid dummy is left out.

	HG	AG	HY	AY	AR	GD	HP	AP	HW
Attendance	0.219	0.002	0.072	0.166	-0.377	0.013	-0.034	-0.111	0.042
	(0.148)	(0.162)	(0.144)	(0.134)	(0.695)	(0.184)	(0.164)	(0.184)	(0.134)
P	0.088	-0.662	1.262	1.062	7.378	1.069	-0.376	-1.499	0.462
	(1.155)	(1.284)	(1.118)	(1.088)	(8.299)	(1.405)	(1.208)	(1.492)	(1.034)
MvH	0.119	-0.222*	-0.222**	-0.137	0.394	0.132	0.306***	-0.101	0.348***
	(0.106)	(0.132)	(0.113)	(0.099)	(0.469)	(0.121)	(0.118)	(0.161)	(0.100)
MvA	-0.320***	0.212**	0.032	-0.077	0.356	-0.299**	-0.339***	0.215**	-0.457***
	(0.085)	(0.089)	(0.083)	(0.082)	(0.318)	(0.133)	(0.103)	(0.103)	(0.084)
Constant	1.586	0.236	2.360	2.025	-12.977	2.646	1.450	-0.351	1.308
	(2.020)	(2.341)	(2.046)	(1.926)	(9.140)	(2.532)	(2.252)	(2.770)	(1.848)
Obs	218	176	210	225	33	138	194	129	267
Pseudo R^2	0.0762	0.0508	0.0233	0.0126	0.1864	0.0389	0.0864	0.0380	0.1452

Notes: The stars indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All coefficients are given with the standard errors given in parentheses.

I left out the dummy variable Covid, whereas the variable attendance and Covid are the opposites. If the variable attendance has a value above zero, the variable Covid is zero and vice versa. Attendance was expected to have a positive coefficient for the variables that could positively influence the match result for the home team. This was only true for home wins, home goals, awarded yellow cards for the away team and goal differences. However, the numbers for attendance were not statistically significant. The market value of the home team was expected to have a positive sign for the wins, goals and points scored variables. The logarithm of home points and the value of home wins are statistically significant.

The variables away goals and home yellow cards decrease significantly when the market value increases. The referees could be biased when a more expensive and a probably better team is playing at home. As expected, the signs of the variables reverse when I move from market value of the home team to market value of the away team. Except for the referee decisions, all the numbers are statistically significant. The market value variables do influence the match result. The numbers of observation differ due to the different statistics. The number of yellow cards differ from the amount of goal differences for example. The value of pseudo R^2 cannot be compared with the "regular" R^2. The paper of Veall and Zimmermann (1994) researched the pseudo R^2 and concluded that if the r-squared is ranging between 0.2 and 0.4, this value will indicate a very good model fit.

Table 3C: Probit regression taking the logarithm of the dependent variables and independent variables. The independent variable attendance is left out.

	HG	AG	HY	AY	HR	AR	GD	HP	AP	HW
Covid	-0.247**	0.196	-0.218*	-	-0.278	0.063	-0.239	-0.320**	0.009	-
	(0.126)	(0.133)	(0.127)	0.486***	(0.809)	(0.508)	(0.170)	(0.140)	(0.151)	0.317***
				(0.124)						(0.116)
P	-0.299	0.474	1.002	0.411	-5.803	0.753	-0.330	0.652	0.216	0.590
	(0.663)	(0.728)	(0.653)	(0.649)	(5.755)	(2.478)	(0.897)	(0.740)	(0.783)	(0.597)
MvH	0.219***	-	-	-0.032	0.331	0.051	0.275***	0.303***	-	0.398***
	(0.060)	0.187***	0.158**	(0.584)	(0.392)	(0.245)	(0.075)	(0.068)	0.229***	(0,058)
		(0.066)	(0.063)						(0.089)	
MvA	-	0.255***	-0.029	-0.036	-0.549	0.153	-	-0.333	0.213***	-
	0.246***	(0.061)	(0.060)	(0.059)	(0.807)	(0.240)	0.268***	(0.075)	(0.067)	0.433***
	(0.061)						(0.104)			0.060
Constant	0.892	-1.389	3.065**	1.348	3.390	-5.143	0.280	0.796	0.228	0.445
	(1.426)	(1.524)	(1.461)	(1.398)	(14.371)	(5.787)	(1.982)	(1.599)	(1.789)	1.325
Obs	431	375	403	423	41	60	246	373	292	538
Pseudo R²	0.0621	0.0585	0.0226	0.0286	0.2379	0.0204	0.0674	0.0926	0.0411	0.1554

Notes: The stars indicate the significance level: * p < 0.10, ** p < 0.05, *** p < 0.01. All coefficients are given with the standard errors given in parentheses.

Whereas attendance was expected to have a positive coefficient for the variables that could possibly favour the home team, the Covid variable is the contrary. This is true for the goals scored, points won and even the yellow cards awarded for the away team. Furthermore, the decrease in home goals, the decrease in awarded yellow cards for the away team and both the decrease in home wins and goal differences, are statistically significant for the Covid dummy.

The value of home points won has a significant decrease when there were no people allowed in the stadium. The value of away points won has a increase during Covid, but the number was not statistically significant. Moreover, the number of home wins decreased during Covid, but no significant relationship was found with attendance.

The coefficient of home goals scored and home points won when the market value of the home team increased, increased when the spectators were not allowed in the stadiums anymore. This does not correspond with what I expected. However, the variables that favour a better result for the away team are expected and actually realised a positive change during Covid. The referee decisions did also correspond with what I expected, the amount of yellow cards awarded increased with no spectators allowed.

The market value of the away team seems to have the same outcomes, whereas the variables that favour the result for their team realised a positive change during Covid and the variables that could possibly influence the outcome positively for the home team did not correspond with what I expected. The coefficient of away goals scored significantly increased during Covid and the number of yellow and red cards awarded to the away decreased, but the numbers were not statistically significant.

The difference in values for home and away goals before Covid and during the pandemic meets the expectations. The value of home goals statistically decreases, whilst the value of away goals increases. However, the number of away goals did not find a significant relationship with spectators.

Neither the table with attendance nor the one with Covid shows any significance in the number of penalties awarded by the different referees. The coefficients of penalties awarded decreased when the matches were played behind closed doors, but it did not find any significant result.

Chapter 5: Discussion and conclusion

5.1: Discussion

Results of the research done in this paper show that there is a positive significant relationship between both market value for the home and the away team and referee decisions. The positive effect comes from the change between the games played with fans in the stadium before Covid and no crowd during Covid. This is in accordance with what the referee hypothesis H1 assumes and the expectation that the market value would have a greater impact without the fans in the stadium.

Whereas expected that the relationship between the market values and the goals scored and points won would be positive, the opposite is true for some other variables. The number of yellow cards given to both teams significantly decreased when matches are played behind closed doors. The opposite is true in the paper of Unkelbach and Memmert (2010), where the number of yellow cards given to the home team would increase in the absence of a crowd. This can be explained by the tension of the game. With spectators in the stadium, the people will make noise and give some extra tense to the game. The game could become a bit more heated and players will react to this with making more fouls.

The value of the coefficients of home points won decreased during Covid, whilst the value of the away points won increased during Covid. The home points won has a significant relationship with attendance, but no significant relationship was found between away points won and attendance.

The number of goals scored by the home team met the expectations during Covid. The decrease in goals scored decreased significantly. The value of away goals had a positive change, but there was no statistically significant relationship found between attendance and away goals scored.

Even though this paper found some significant results, it still suffered from some limitations. At first, the first season was not completed. The last eight rounds were removed from the season due to the pandemic. This season could have had other outcomes, especially because in the end of the season the stakes of some teams are higher when promotion and relegation draw nearer.

Secondly, the variable red cards awarded by the home team did not gave valid results in Stata. Stata gave 19 failures and the log likelihood was zero. The results were too difficult to implement and are left out of this paper.

5.2 Conclusion

This paper investigated whether the fans of a certain home team could positively influence the result of their team. The expectations were not completely met by the results. I expected an increase in away goals and points, and a decrease in home points and home goals during Covid. From our results, we could not conclude that the goals scored and points won significantly increased or decreased when the matches were played behind closed doors.

The findings of the paper do show that the decisions made by the referees are significantly affected by the number of fans in the stadium and the market value of a single team. This result is partially in accordance with research done by Bryson et al. (2021). They found that the home yellow cards increased significantly during Covid. However, the finding that the away yellow cards significantly decreased behind closed doors, is in line with the paper of Bryson et al. (2021).

The paper of Boyko, Boyko and Boyko (2007) found that referees have a significant bias for the home team due to the home advantage phenomenon. Their results demonstrated that referees could favour the home team due to crowd noises and this in accordance with the paper of Bryson et al., who found a similar significant relationship between referee decisions and home advantage.

So, this research shows that the absence of fans has no significant effect on the match result, the goals scored by both teams, the penalties awarded per referee and the total points won at home. However, the market value of both the home and away team did have a significant greater impact in the absence of a crowd. Furthermore, the absence of a crowd did have a significant effect on referee decisions and cards given to both the home and away team.

This paper has shown a relationship between the decisions of a referee and the number of attendees, so this aspect of the paper contributes to the economic and social relevance. The increase in attendance did not show significant effects on the result for the home team, therefore it would not necessarily be interesting to invest in a bigger stadium to increase

capacity. On contrary, further research could look into investing opportunities for football clubs. For example, an increasingly number of clubs have made deals with "Cryptocurrency companies" and aim to innovate to keep up with the times.

Suggestions for further research

In this part, I will discuss aspects that are missing out and can be helpful for further research. At first, collecting more data would help getting better insights. Results would be more accurate if the number of matches played, increased. By adding more seasons of football games played to the data, the power of individual fixed effects will reduce.

Secondly, it would be valuable if one will expand this research to different competitions abroad and internationally. This will give more reliable and significant results on home advantage. Adding several different leagues from abroad and including the UEFA Champions League and UEFA Europa League, would give a more precise outcome of this research.

As mentioned before, it would be interesting to add travel distance and travel fatigue. Pollard (2008) found evidence for travel distance on home advantage. The distance between the two teams is negatively correlated with the result in the perspective of the away team. Clarke and Stefani (1992) concluded that the home team have a higher chance of winning in European games compared to national games. So, to add travel distance would be valuable on the influence of home advantage or the disadvantage for away teams.

Besides this, the number of injuries per team could also be helpful. Adding a monthly fixed effect on a player's sickness and number of positive tests on the coronavirus. In general, the number of positive Covid-19 tests and ill players are higher in the winter compared to the warmer and sunnier days.

Furthermore, research on the level of individuals involved would be interesting. This could be on characteristics of referees in different countries or performances of a single player in away games for example. The most important person to look at in behaviour and his performances is the captain of the team. Does one act different in derbies compared to other away games by being more aggressive or is he affected by the noise of the crowd of the opponent?

Lastly, it would be interesting to look at economic opportunities for clubs to grow financially and improve their results. I mentioned that investing in a bigger stadium with more capacity would not be profitable, but this was solely based on the fact that an increase in attendance did not show a significant effect on the match results of the home team. So, further researches could look into a profitable way on how to improve results for different clubs.

References

Allen, M.S., and Jones, M.V. (2012). The home advantage over the first 20 seasons of the English Premier League: Effects of shirts colour, team ability and time trends. *International journal of sport and exercise psychology, 12(1),* 10-18.

Amaro, N., Matos, R.M., and Pollard, R. (2019). How best to quantify home advantage in team sports: an investigation involving male senior handball leagues in Portugal and Spain. *Revista internacional de ciencias del deporte, 59(16),* 12-23.

Barsky, S.F., and Schwartz, B. (1977). The home advantage. *Social forces, 55(3),* 641-661.

Boyko, R.H., Boyko, A.R., and Boyko M.G., (2007). Referee bias contributes to the home advantage in English Premiership football. *Journal of sports sciences, 25(11),* 1185-1194.

Bryson, A., Dolton, P., Reade, J.J., Schreyer, D., and Singleton, C., (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. *Economic letters, 198.*

Carron, A.V., Loughhead, T.M., and Bray, S.R., (2005). The home advantage in sport competitions: Courneya and Carron's (1992) conceptual framework a decade later. *Journal of sports sciences, 23(4),* 395-407.

Clarke, S.R., and Norman, J.M. (1995). Home ground advantage of individual clubs in English soccer. *Journal of the royal statistical society, 44(4),* 509-521.

Courneya, K.S., and Carron, A.V., (1992). The home advantage in sport competitions: a literature review. *Journal of sport and exercise psychology, 14(1),* 13-27.

Dawson, P., Dobson, S., Goddard, J., and Wilson, J., (2007). Are football referees really bias and inconsistent?: Evidence on the incidence of disciplinary sanction in the English Premier League. *Journal of the royal statistical society: Series A, 170(1),* 231-250.

Endrich, M., and Gesche, T., (2020). Home-bias in referee decisions: Evidence from "Ghost Matches" during the Covid-19 Pandemic. *Economic letters, 197.*

Matos, R., Monteiro, D., Antunes, R., Mendes, D., Botas, J., Clemente, J., and Amaro, N. (2021). Home Advantage during COVID-19: An analysis in Portuguese Football League. *International journal of environmental research and public health, 18,* 3761.

McCarrick, D., Bilalic, M., Neave, N., and Wolfson, S. (2020). Home advantage during the Covid-19 pandemic: analyses of European soccer leagues. https://doi.org/10.31234/osf.io/2gkht

Nevill, A.M., and Holder, R.L. (1999). Home advantage in sport. *Sports medicine, 28(4),* 221-236.

Page, L., and Page, K. (2010). Alone against the crowd: Individual differences in referees' ability to cope under pressure. *Journal of economic psychology, 31.* 192-199.

Pettersson-Lidbom, P., and Priks, M. (2010). Behavior under social pressure: Empty Italian stadiums and referee bias. *Economic letters, 108(2),* 212-214.

Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. *Journal of sports sciences, 4,* 237-248.

Pollard, R. (2002). Evidence of a reduced home advantage when a team moves to a new stadium. *Journal of sports sciences, 20,* 969-973.

Pollard, R. (2008). Home advantage in football: A current review of an unsolved puzzle. *The open sports sciences journal, 1,* 12-14.

Ponzo, M., and Scoppa, V. (2018). Does the home advantage depends on crowd support? Evidence from same-stadium derbies. *Journal of sports economics, 19(4),* 455-472.

Schwarz, W., (2011). Compensating tendencies in penalty kick decisions of referees in professional football: Evidence from the German Bundesliga 1963-2006. *Journal of sports sciences, 29(5),* 441-447.

Seckin, A., and Pollard, R. (2008). Home advantage in Turkish professional soccer. *Perceptual and motor skills, 107(1),* 3-13.

Stefani, R., and Clarke, S., (1992). Predictions and home advantage for Australian rules football. *Journal of applied statistics, 19(2),* 251-261.

Unkelbach, C., and Memmert, D., (2010). Crowd noise as a cue in referee decisions contributes to the home advantage. *Journal of sport and exercise psychology, 32,* 483-498.

Van Ours, J.C., and Peeters, T. (2020). Seasonal home advantage in English professional football. *De economist, 169,* 107-126.

Van de Ven, N. (2011). Supporters are not necessary for the home advantage: Evidence from same-stadium derbies and games without an audience. *Journal of applied social psychology, 41(12),* 2785-2792.

Veall, M.R., and Zimmermann, K.F. (1994). Evaluating Pseudo R^2's for binary probit models. *Quality and quantity, 28,* 151-164.