What is the effect of crowd support on team performance?

**Using the covid-19 pandemic to analyze the influence of crowd support on team performance**



1. Rijnders (**2063482**)

04-06-2021

MSc Marketing Analytics

Tilburg School of Economics and Management

Tilburg University

Master Thesis Supervisors:

1. van der Plas MSc

Dr. A.I.J.G. van Lin (co-reader)

Management Summary

The notion of home-advantage is a widely studied concept in sports related academic literature. Studies such as those conducted by Schwartz and Barsky (1977) and Pollard(1986) document a persistent advantage for teams playing at home. Crowd support is considered to be one of the major drivers of team performance and home advantage in sports. However, the effect of crowd support on home advantage and team performance is difficult to directly assess. During the covid-19 pandemic all soccer matches in major European soccer leagues had to played without fans. We used this extraordinary situation to directly assess the impact of crowd support on team performance.

To give managers an insight on how they could impact the performance of their team, we extended the analysis by looking into the effect of several moderating variables, namely the average age of players, the number of foreign players in the squad, the absolute crowd size and the stadium occupancy rate. Furthermore, we assessed to what extent the influence of crowd support on team performance is mediated by a referee bias in favor of home teams. We used a structural equation model(SEM) to perform a moderated mediation analysis to assess the influence of crowd support on team performance. We incorporated referee bias as a latent construct comprising of differences in fouls and yellow cards between home and away teams. In goal difference and points difference we used two seperate measures of team performance as dependent variables.

We found that crowd support significantly affects team performance for both home and away teams. Part of this effect is mediated by a referee bias for the home team. Furthermore, we found evidence that rather than absolute crowd size, the stadium occupancy rate plays a major role in the size of the effect of crowd support on team performance. For high levels of crowd occupancy the difference in points and goals was .514 and .737 whereas for low levels of occupancy the gap only decreased by .197 for goals and .339 for points. We found no evidence for a significant influence of team age and team composition on the relationship between crowd support and team performance.

Based on these findings marketing managers know that the importance of a filled stadium lies beyond the extra revenue associated with more fans. Higher occupancy levels can mean the difference between qualifying for European club competitions or not or the difference between staying up in the highest division and relegation. Both outcomes have immense financial consequences. This implies that decisions such as moving to a new stadium should only be taken when you are really certain to be able to fill the stadium as much as the current stadium.

Preface

In front of you lies the thesis “*What are the effects of crowd support on team performance and home advantage”*, which aims to understand the influence of crowd support on team performance and referee decision making. This thesis was written as the final step of completing the master program

Marketing Analytics at Tilburg University, Tilburg School of Economics and Management. The

process of writing this thesis has engaged me from February 2021 to June 2021.

This thesis was written under supervision of Joep van der Plas, who introduced me to this

topic. The process of writing this thesis has been challenging. Fortunately, Joep van der Plas

was always very helpful during our meetings. He provided excellent feedback to help me through the process of writing this thesis.

I would like to thank Joep van der Plas for the excellent guidance. I also would like to thank my parents, friends, and family for their support and encouraging words. This thesis will not be my final chapter at Tilburg University, as I am looking forward towards my participation in the QTEM program, for which I am very thankful to Tilburg University to be given this opportunity. I hope you enjoy reading my master thesis

Alan Rijnders

Tilburg, June 4th, 2021

Table of contents

[1 Introduction 5](#_Toc32222)

[1.1 Problem Indication 5](#_Toc3524)

[1.2 Research Approach 7](#_Toc931)

[1.3 Academic Relevance 7](#_Toc2946)

[1.4 Managerial Relevance 8](#_Toc26845)

[1.5 Structure of the Thesis 9](#_Toc7574)

[2. Theoretical Background 10](#_Toc22013)

[2.1 Literature Review 10](#_Toc12870)

[2.1.1 Crowd support and home advantage 10](#_Toc12640)

[2.1.2 Team composition and home advantage 12](#_Toc19170)

[2.1.3 Covid-19 and home advantage 13](#_Toc474)

[2.2 Conceptual Framework 14](#_Toc15019)

[2.2.1 The impact of crowd support on team performance 14](#_Toc32374)

[2.2.2 The impact of crowd support on referee bias 16](#_Toc4618)

[3. Data and Methodology 18](#_Toc32099)

[3.1 Data collection 18](#_Toc16459)

[3.2 Variable Operationalization 19](#_Toc17194)

[3.3 Descriptive statistics 22](#_Toc2175)

[3.4 Home advantage pre and post covid-19 22](#_Toc31894)

[3.5 Model 28](#_Toc2618)

[4. Analysis and Findings 29](#_Toc16644)

[4.1 Assumptions 29](#_Toc25909)

[4.2 Results overview 30](#_Toc10217)

[4.3 Robustness check 33](#_Toc30502)

[4.4 Moderating effect occupancy 35](#_Toc5074)

[5. Conclusion and Discussion 36](#_Toc27988)

[5.1 Conclusions 36](#_Toc9776)

[5.2 Managerial recommendations 38](#_Toc28579)

[5.3 Limitations and areas for future research 39](#_Toc10919)

[5.4 Final overview 40](#_Toc6449)

[6. Bibliography 40](#_Toc32192)

[7. Appendix 45](#_Toc21184)

# Introduction

## Problem Indication

Soccer is the most popular sport in the world with 265 million of active soccer players around the globe (FIFA, 2007). In other words, roughly 4 percent of total world population regularly plays soccer. The amount of fans is even more staggering, with 3.5 billion people tuning in for the FIFA World Cup Final between France and Croatia in 2018 (FIFA, 2019). Many of these fans were left without their favorite pastime for a considerable time when the covid-19 pandemic struck Europe and other parts of the world in March 2020. Almost all sports games were postponed until further notice due to strict measures and even complete lockdowns in several countries. After a few months of lockdown, the Bundesliga was the first major league to restart on the 16th of May 2020 featuring 6 matches behind closed doors. Other major European leagues such as the Premier League, Serie A and La Liga followed swiftly. Global social distancing measures during the covid-19 pandemic introduced the phenomenon “Ghost Matches”; soccer matches without any spectators attending. Without the fans attendance the matches seemed mere training games. Some surprising results in favor of away teams in the first few weeks following the restart renewed interest in the role of crowd support and team performance.

Thanks to soccer’s worldwide popularity, an abundant base of research on the possible advantage for teams playing at home existst. Goumas (2014) finds that home teams on average win more games, collect more points, and score more goals than away teams. The authors findings resonate with results from earlier studies such as those conducted by Clarke and Norman (1995) and Pollard, (2006). Both studies found a continual advantage for teams playing at home across countries and time. However, the exact source of home advantage and the role of crowd support herein remains debatable. Pollard & Polla[rd (2005](https://shapeamerica.tandfonline.com/doi/full/10.1080/02640410601038576?scroll=top&needAccess=true)) for example negate the role of crowd support in the formation of home advantage whereas Smith (2003) argues that the performance boost for home teams as a consequence of crowd support is the main driver of home advantage. Crowd support could also affect match outcome through referees decisions. (Bokyo, 2007; Neville & Holder, 1999) find that crowds could influence referee decisions subconsciously in favor of the home team. This favoritism is often named the “referee bias”, a term we will use as well in our research from now on. Endrich & Gesche (2020) quantify the referee bias in their paper where they find that away teams on average receive 0.3 cards less and home teams 0.5 cards more per match when there are no spectators.

In general, home advantage seems to be declining over time (Peeters & van Ours, 2021). One of the major factors behind this decline is the greater distance between fans and players. With many players coming from all over the world there may presumably be less of a connection with the fans who are used to supporting their local heroes (Pollard 2006; Smith 2003). This aspect of crowd support and home advantage has been overlooked in the literature. Therefore, we use this unique situation to assess the extent to which potential distance between fans and players impacts crowd support and team performance.

To summarize, opinions on how and to what extent home advantage is shaped by crowd support are very mixed. Considering the uncertainty regarding the effect of crowd support, the current extraordinary circumstances thus provide a special opportunity to increase our understanding of the relevance of crowd support.

## Problem Statement

We summarize the aim of this thesis in the following problem statement.

*What is the effect of crowd support on attitude team performance, mediated by a referee bias towards the home team, and what is the moderating effect of team average age, share of foreign players within the team, crowd size and stadium occupancy rate on the relationship between crowd support and team performance?*

## Research Questions

Our central problem will be answered by addressing the following research questions.

* What is the effect of crowd support on team performance?
* To what extent is the effect of crowd support on team performance mediated by a referee bias?
* To what extent does the average age of teams affect the relationship between crowd support and team performance?
* To what extent does the share of foreigners within a team affect the relationship between crowd support and team performance?
* To what extend does the crowd size affect the relationship between crowd support and team performance?
* To what extend does the stadium occupancy rate affect the relationship between crowd support and team performance?

## Academic Relevance

The batch of matches without fans provides a unique opportunity to delve deeper in the impact of crowds on soccer matches since there is an abundance in new data for matches played without crowds. Past papers often required advanced econometric techniques, relying on various assumptions, to be able to discern the extent to which a crowd influences home team advantage. The difficulties in assessing drivers of home advantage lies in confounding variables effects (Pollard, 2008). This makes it unclear how much of home advantage can be attributed to what factors since most of the factors of home advantage are connected to each other. However, in the current situation, there are suddenly substantial amounts of data on “ghost games” which can be used to directly compare matches with and without spectators. Thus significantly reducing the number of needed assumptions.

A few papers on home advantage in times of covid-19 have been published already, including the work of (Fischer & Haucap, 2020;Deutscher & Winkelmann, 2020; Endrich & Gesche, 2020). However, these papers focused on German leagues in their analysis. We add to the current research by extending the preliminary research already completed on numerous other soccer leagues, Enabling us to obtain a comprehensive overview of the evolution of home advantage during the pandemic. Also compared to preliminary studies on home advantage, in our paper we use the extended availability of data for games played behind closed doors available in the 2020/21 season. Papers from last year mostly use data from the 2019/2020 season which was partly played under normal circumstances and partly played behind closed doors. Incorporating data from 2020/21 into the analysis increases the sample size and decreases confounding effects resulting from possible biased schedules in partial seasons.

Finally, we extend existing research by constructing a comprehensive framework of the influence of crowd support on home advantage by using several mediating and moderating variables to assess causal links between crowd support and home advantage.

## Managerial Relevance

Knowledge on the variables driving home advantage and team performance provides great insight for soccer clubs in how to optimize their clubs environment and team to improve the chances of performing well. Fischer & Haucap (2020) for example see a significant effect of stadium occupancy rate on home team performance by comparing home advantage during covid-19 between the German Bundesliga and the 2nd and 3rd level of German soccer. Similary, Goumas (2013) and Nevill, Newell and Gale (1996) find evidence for increased home advantage for teams playing for larger crowds. Stadium occupancy and crowd size are, to a certain extent, under the control of soccer club executives. Wetzel, Hattula, Hammerschmidt & van Heerde (2018) show, for example, that soccer clubs possessing a stronger brand name can leverage this brand name to increase attendance. An effect which increases over the length of time the brand exists. Creating a stronger brand could increase attendance rates and match day revenues. Other variables that are under control of management and possibly related to team performance differences in home and away matches are the average age of the squad players and the composition of the squad in terms of local and foreign players. Previous research conducted by van de Ven (2016) signal a small effect of age on team performance, with teams with a higher average age performing slightly better away than teams with a lower average age. We aim to examine whether a soccer club branding their club to their supporters as an experienced squad with local players can increase the team’s performance.

By examining the effect of variables such as crowd occupancy and team composition on home performance, we aim to provide marketers direct tools to influence the performance of their clubs. Hypothetically speaking, knowing that improving the occupancy rates for their team increases team performance gives the marketing department a significant task in finding ways to attract more fans to the stadium. Perhaps,a slight reduction in ticket price with the associated lower customer revenue could actually turn out as a smart investment, resulting in better team performance and an upwards positive spiral both on and off the field. We believe the current literature is lacking in this area. Most of the papers available on this topic refrain from applying their findings to managerial recommendations for soccer clubs.

The effect of team performance on the pitch translates to a soccer clubs’ performance off the pitch. Team results influence soccer clubs’ performance outside of the pitch in several ways. Samagaio, Couto & Caiado (2009) find a positive relationship between on pitch results and stock performance. Moreover, increased team performance leads to increased market value for players (Galariotis, Germain & Zopounidis, 2018 ; He, Cachucho & Knobbe, 2015; Müller, Simons & Weinmann, 2017). All studies find positive relationships between revenues and position in the league table as well as between revenues and individual performance. Therefore, estimating the influence covid-19 has on team performance will also help soccer clubs evaluating the effect of the pandemic on their marketing, financial and on-field performance.

## Research Approach

Providing meaningful insights into the role of crowd support requires a combination of theoretical and empirical analysis. Firstly, we examine the current and historical literature to obtain a view on the current knowledge on the relationship between crowd attendance and team performance. This will serve as our basis on which we can conduct the right analytical approach to analyze our data. We decide to include the seasons 2018/19, 2019/20 and 2020/21 into our final analysis. Around 2018, the Video Assitant Referee(VAR) was introduced in most of the major leagues in Europe, possibly changing our results on referee bias significantly. We then combine match data with a dataset from Fivethirtyeight to incorporate team strength and match importance, which we include as control variables into the analysis. The data for our moderator variables, team age, amount of foreigners, crowd size, and stadium occupancy are collected manually from Transfermarkt. This website provides a vast amount of detailed statistics on players, clubs and leagues.

## Structure of the thesis

This thesis is divided into 5 chapters. The first chapter serves as a background chapter for the rest of the thesis in which we outline the concept to be researched in combination with the academic and managerial relevance of the concept. In the second chapter we construct the theoretical framework that will represent the basis of the empirical analysis in the later sections. First, we analyze the current literature to obtain an overview of what is currently known. Thereafter, we define the conceptual framework that serves as overview of the relationships we examine. In chapter 3 we meticulously describe the data set and the variables we use to define the concepts that we want to analyze. We then go on to provide our first model free evidence to examine trends in home and away team performance. Furthermore, based on our variable selection and data structure, we elect the most suitable method of analysis. In the fourth chapter we summarize the analysis and findings of our model to answer the empirical questions. Additionally, we extend our model with a robustness check to ensure that our findings are stable. In our final chapter, we generate conclusions and recommendations based on our findings. We use this chapter to provide soccer club management with deeper insights into the drivers of team performance at home, and to what extent these can be influenced by management. Additionally, we discuss the limitations of this study and provide a guideline for possible future research in this area to solidify the understanding of crowd support in relation to home advantage.

## Theoretical Background

## Literature Review



## 2.1.1 Crowd support and home advantage

Home advantage has been widely studied in the literature. One of the first to formally document the existence of a certain home advantage in sports were Schwartz & Barsky (1977). They find that home advantage exists in varying degrees across different sports. In their research they suggest that the major contributor to home advantage is social support as they find a strong relationship between audience size and home advantage. Nevill and Holder (1999) support this claim as they produce similar results in their analysis of home advantage in English and Scottish soccer matches. Ponzo & Scoppa (2018) argue that a home crowd can be a positive stimulus for home team players and can create an intimidating and hostile environment for the opposition. Home team performance is raised relative to away team performance as their effort and energy is stimulated by the positive support from the crowd. Ponzo & Scoppa (2018) base their conclusions on the analysis of same stadium derbies in Rome and Milan to mitigate other possible factors of home advantage such as traveling and familiarity effects. They find that when controlling for referee decisions and other factors such as team strength, the home team still performs better in the local derby.

An interesting question then arises whether all teams experience a similar boost from their home crowd or that certain team characteristics or crowd characteristics could be associated with higher levels of home advantage. Each home crowd is unique, crowds differ substantially in size, density and also fanaticism. Carron and Agnew (1994) find a significant positive relationship between home advantage and crowd density. In other words, more crowd support leads to a stronger home performance relative to away performance and consequently a higher chance of a home win than an away win. Fischer & Haucap (2020) also find that there seems to be a significant alteration in the strength of home advantage due to differences in crowd occupancy. They found a significant decrease in home advantage in the Bundesliga when crowd support is absent. However, in the 2nd Bundesliga and 3rd Liga home advantage did not change significantly during ghost games. They account this difference to the differences in occupancy rates between these competitions.

However, other researchers questioned whether crowd size and occupancy actually are important. Pollard (1986) negates the importance of crowd size and crowd density. In his argument Pollard (1986) uses the notion of a similar magnitude in home advantage across first and second divisions across Europe. Despite the vast differences in crowd size and crowd density between first and second divisions, the home advantage still persists. Furthermore, Salminen (1993) and Strauss (2002) claim that crowd support in the form of cheering does not affect team performance. In fact, they even find support for the case that teams are motivated by non-supportive audiences and play better in such situations.

Different results seem to occur because of the difficulty in disentangling each of the various forces driving home advantage. Pollard (2008) explains that struggles concerning the unraveling of individual factors effect on home advantage sterns from the phenomenon that the multiple psychological and physiological influences involved all interact with each other and possibly reinforce each others significance. For exactly this reason “ghost games” provide such a unique opportunity to specifically study changes in home advantage as a consequence of crowd support. Therefore we use “ghost games” to examine whether crowd occupancy and crowd size significantly affect team performance.

Apart from directly raising home team performance, crowd support is said to affect team performance through the referee. Multiple studies including: (Nevill, Balmer & Williams, 1999 ; Nevill, Balmer & Williams, 2002; Garicano, Palacios-Huerta & Prendergast, 2005; Unkelbach & Memmert , 2010; Sutter & Kocher, 2004) find consistent evidence of a referee bias in favor of the home team probably due to social pressure from the crowd. Examples of this bias include the issue of more stoppage time at the end of the first and second half when the home team is trailing. In more recent research Endrich & Gesche (2020) find that referees give less cards and fouls to home teams and more cards and fouls to away teams on average, which could be interpreted as a sign of referee bias in favour of the home side. A referee has a large influence on the outcome of the games (Boyko, Boyko & Boyko, 2007). Especially in a sport as soccer which is of a low scoring nature (Decroos, Bransen & Davis, 2019). A decision to award a team a penalty in the 89th minute of the match with the score at 0-0 could completely change the match outcome. Or an early red card significantly alters the course of the match with both teams adjusting their tactics and strategy and as such, influence team performance.

Previous studies found evidence that crowd cheering and noise are the main contributor to referee bias (Endrich & Gesche, 2020; Nevill, Balmer & Williams, 2002). Referees can be heavily influenced in their decision making by the heavy cheering of the crowd favoring the home team (Unkelbach & Memmert , 2010). Experiments conducted by Nevill, Balmer and Williams (2002) show the role of crowd noise by asking participating referees to evaluate fouls. One group of the referees were shown the fouls with sound of the crowd in the background whereas the other group watched the fouls in silence. The referees watching with crowd noise on average gave 15 percent less fouls to the home team compared to referees watching in silence. Referee bias and crowd noise is well documented in the literature. However, there is less empirical research on to what extent crowd size and occupancy influence referee decisions. Research like ours on the incidence of referee bias in ghost games settings could provide useful in discerning whether or not crowd size and occupancy play a role in referee decision making. Furthermore, most of the papers tend to focus solely on the existence of referee bias rather than on the implication of a possible referee bias on team performance. We extend the current literature by incorporating the influence of referee decisions on team performance as a mediating variable.

## 2.1.2 Team composition and home advantage

Another stream of the literature on home advantage focuses on familiarity effects. Pollard (2008) describes familiarity effects as key stadium attributes that could help players locate themselves more precisely on the pitch and consequently make better decision on where and how hard to pass the ball or where to position themselves to get the best shot on goal. Older players who have more experience, especially when that experience is within the same league, will be more familiar with venues and could have similar advantages as home team players in visual cues when playing away. The concept of familiarity can also be extrapolated to the realm of crowd support. This school of thought has not been widely studied yet but studies such as that of van der Ven (2016) report a slightly better away performance for teams with a higher average age, compared to teams with a lower average age. Teams with a higher average age could be more experienced with facing home crowds, which in turn could decrease the effect of these crowds on their performance. Russell (1983) for example finds that older players develop certain coping strategies to deal with the influence of the crowd on their performance.

A different component of team composition that could influence the effect of crowd support on team performance is the amount of foreign players featuring for the teams. In the increasing globalized world, international transfers are increasingly common, leading to an influx of foreign players into squads of soccer clubs (Adcroft, Teckman & Madichie, 2009). These foreign players, with increasingly high salaries are difficult to relate to for local often working class soccer supporters (Petersen-Wagner, 2015; Smith, 2003). This leads to fans and players becoming more and more detached from each other, decreasing the bonding between fan and players. Gutierrez (2019) claims that this bonding process between fans and players is a crucial component for fan engagement and consumption. Increased fan engagement leads to a better atmosphere and louder crowds. Lee, Gipson and Barnhill ( 2017) provide further evidence for the influence of fan identification with their team. They surveyed attendants of basketball and baseball games in the NCAA division. They found that measures of team identification significantly influenced crowd atmosphere through an indirect effect on flow of supporters, with flow being defined following the definition of Csikszentmihalyi (1990): "the state in which people are so involved in an activity that nothing else seems to matter”. Their findings suggest that a reduced identification of supporters with their team decreases atmosphere within the stadium. The difference in atmosphere within the stadium could influence team performance.

## 2.1.3 Covid-19 and home advantage

A few preliminary studies attempted a similar approach to ours by analyzing “ghost games” played between the restart after corona and the end of the season 2019/20. Thilp & Taller (2020) for example find that home advantage has actually turned into a home disadvantage in case of “ghost games”. Fischer & Haucap (2020) also support the notion of a signifcant alteration in the strength of home advantage in the Bundesliga when crowd support is absent. McCarrick, Bilalic, Neave and Wolfson (2020) report similar findings in their study of home advantage across 11 countries. They discovered that across those leagues the number of goals scored and points obtained by home teams was significantly lower in matches played during the covid-19 pandemic. However, apart from McCarrick et al (2020), most of the recent papers only include one single country into their analysis. We believe that this approach is limited because single countries could be an anomaly. Especially when the data is also limited to only the end of the 2019/20 season. Within this smaller sample, a few abnormal results could already influence conclusions. We incorporate multiple countries in our dataset and extend the analysis to all the “ghost games” played up to date to obtain a larger sample and more generalizable results.

The table below gives a short overview of the current literature and our contribution.

## 2.2 Conceptual Framework

In the next section we provide our hypotheses based on the literature. After the hypotheses, we present a schematic overview of our conceptual framework to clarify the concepts and relationships we investigate in this study.

## 2.2.1 The impact of crowd support on team performance

Crowd support can raise home team performance relative to away team performance directly through cheering and booing. Crowd support can inspire home teams to perform to their potential, increasing home team performance. Thus, home team performance relative to away team performance could be lower when in a situation of no or less crowd support. This leads us to generate the following hypothesis regarding the effect of crowd support on team performance.

*H1: Crowd support positively influences Home Team Performance.*

Crowds come in all shapes and sizes and different crowds will have different influences on team performance. Bigger crowds in general make more noise and can be more intimidating than smaller crowds. There is a big difference in playing for large crowds compared to small crowds. Crowd support has a direct effect on team performance and a larger crowd size is associated with larger crowd noise. This larger crowd noise and size could boost confidence of the home team, knowing they have got the backing from so many fans, and thus could lead to larger performance boosts for home teams for teams backed by large crowds compared to teams supported by smaller crowds. This leads to the following hypothesis on the effect of crowd size on team performance.

*H1b: The effect of crowd support on team performance increases when crowd size increases.*

Crowd occupancy is also important for atmosphere within a stadium, and in turn the effect of crowd support on team performance support. Fischer & Haucap (2020) find that teams with higher occupancy rates pre covid-19 experience a greater decrease in home advantage post corona. If you play for 30.000 fans in a stadium where 100.000 fit, the atmosphere seems to be less intense and the stadium can appear to be almost empty. The switch to a completely empty stadium in this case might be less severe than a case where 15.000 very fanatic fans completely fill up a small stadium with stands close to the pitch and a fiery atmosphere. A completely empty stadium then all of a sudden makes a very big change.Therefore we hypothesize the following on the effect of stadium occupancy on team performance.

*H1c: The effect of crowd support on team performance increases when stadium occupancy rate increases.*

The degree to which crowd support will influence team performance will vary per team. Each individual player reacts differently to playing environments. Team composition thus seems to play a role. Vvan de Ven (2016) finds that experienced sides with older players tend to perform better away from home than inexperienced sides. Possible reasons could include familiarity with the away venue and more experience with hostile crowds. Older players who have more experience, especially when that experience is within the same league, will be more familiar with venues and could be more experience with home crowds. Older players can develop coping strategies to decrease the influence of opposition crowds on their performance when playing away (Russell, 1983). This leads to the following hypothesis of the effect of age on the relationship between crowd support and team performance.

*H1d: The effect of crowd support on team performance is weaker for teams with older players.*

Another aspect of team composition that we analyze is the division between local and foreign players for teams. Tilp and Taller (2020)mention an increased global outlook of soccer clubs, both for recruiting fans and players as a factor, which has led to an increased gap between fans and players. Fans and players due to the increased differences in pay and origins live in completely different realities from each other. Fans do not recognize themselves in the extremely rich and foreign players who play for their local team, Lower fan identification with a soccer team decreases the support of those same fans when attending the match. This decreases crowd cheering and thus indirectly decreases the effect of crowd support on team performance. Consequently we hypothesize the following regarding the effect of share of foreigners within a team on team performance.

*H1e: The effect of crowd support on team performance is weakened when the share of foreigners increases.*

## 2.2.2 The impact of crowd support on referee bias

Referees have shown a consistent bias towards home teams when awarding fouls and cards. Referees are subconsciously influenced by crowd noise when making decisions on potential fouls, cards and penalty’s (Nevill & Holder, 1999). Punishing home teams less severe in situations with crowd noise. Potential explanations include the use of visual cues in decision making when the situation is not very clear, with crowds reaction to a foul serving as a potential indicator of the actual situation and referees relying partly on these crowd judgements when making a decision. Additionally, referees could favor home teams in order to avoid potential crowd displeasure aimed at them during the rest of the game and even after the game. In soccer, much more compared to other sports, one action can decide the entire game. A 1-0 win with a single shot on goal is certainly a possibility. Additionally, a red card can change a teams entire game plan, tactics and performance.

*H2: The effect of crowd support on team performance is mediated by Referee Bias*

Similar to the expected moderating effect of crowd occupancy and crowd size on the direct relationship between crowd support and team performance, we expect crowd occupancy and crowd size to influence the relationship between crowd support and referee bias (Nevill, Balmer & Williams, 2002) demonstrated in an experimental setting the significant effect of crowd noise on referee decision making. Referees are more uncertain in their decisions when crowd noise is present compared to situation where there is only silence. Often more favoring the home team in a situation with crowd noise by being more lenient in giving fouls and cards . Therefore, a higher occupancy and a higher crowd size, with more crowd noise, will result in a stronger referee bias towards the home team. This leads to the following two hypotheses.

*H2b: The mediating effect of referee Bias on the relationship between crowd support and team performance increases when Stadium Occupancy increases.*

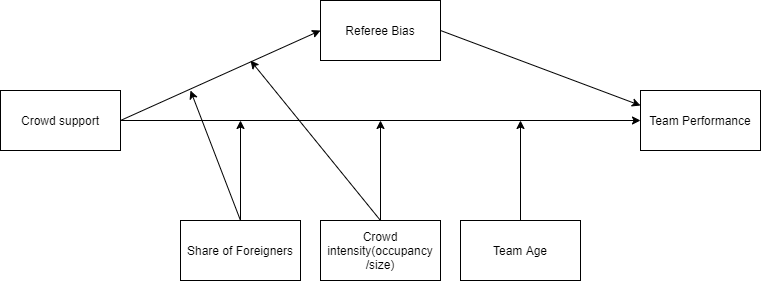
*H2c: The mediating effect of referee Bias on the relationship between crowd support and team performance increases when average Crowd Size increases.*

Similar to the expected moderating effect of share of foreigners within a team on the direct relationship between crowd support and team performance, the share of foreigners within a team influences the relationship between crowd support and referee bias. A team with a high share of foreign players makes it harder for the home crowd to bond with their own team, producing a less intense atmosphere in the match and consequently less crowd noise. Since referee decisions in favor of home teams are influenced by crowd noise and crowd reactions, their decisions will be less favorable for home teams when the crowd noise is lower. Accordingly, we construct the following hypothesis.

*H2d: The mediating effect of referee Bias on the relationship between crowd support and team performance decreases when the share of foreigners increases.*

Figure 1 shows the conceptual model we establish based on the current literature and hypotheses. This conceptual model will be used in later stages to build the correct model to analyze the data.

**Fig**%3CmxGraphModel%3E%3Croot%3E%3CmxCell%20id%3D%220%22%2F%3E%3CmxCell%20id%3D%221%22%20parent%3D%220%22%2F%3E%3CmxCell%20id%3D%222%22%20value%3D%22Crowd%20support%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%2220%22%20y%3D%22110%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%223%22%20value%3D%22Referee%20Bias%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22340%22%20y%3D%2220%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%224%22%20value%3D%22Team%20Performance%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22660%22%20y%3D%22110%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%225%22%20value%3D%22Share%20of%20Foreigners%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22200%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%226%22%20value%3D%22Crowd%20intensity(occupancy%26lt%3Bbr%26gt%3B%2Fsize)%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22340%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%227%22%20value%3D%22Team%20Age%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22480%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%228%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BentryX%3D0%3BentryY%3D0.5%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20target%3D%224%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22140%22%20y%3D%22140%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%2270%22%20y%3D%22320%22%20as%3D%22targetPoint%22%2F%3E%3CArray%20as%3D%22points%22%3E%3CmxPoint%20x%3D%22140%22%20y%3D%22140%22%2F%3E%3C%2FArray%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%229%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D1%3BexitY%3D0.5%3BexitDx%3D0%3BexitDy%3D0%3BentryX%3D0%3BentryY%3D0.5%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20source%3D%222%22%20target%3D%223%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22150%22%20y%3D%22160%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22330%22%20y%3D%2250%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2210%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%225%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22260%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2211%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D1%3BexitY%3D0.5%3BexitDx%3D0%3BexitDy%3D0%3BentryX%3D0%3BentryY%3D0.25%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20source%3D%223%22%20target%3D%224%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22650%22%20y%3D%22120%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2212%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3B%22%20edge%3D%221%22%20source%3D%226%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22230%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2213%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%227%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22540%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2214%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%225%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22210%22%20y%3D%22110%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2215%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3B%22%20edge%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22240%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22270%22%20y%3D%2280%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2216%22%20value%3D%22%22%20style%3D%22endArrow%3Dnone%3Bdashed%3D1%3Bhtml%3D1%3B%22%20edge%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22620%22%20y%3D%2250%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22620%22%20y%3D%2250%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3C%2Froot%3E%3C%2FmxGraphModel%3E**ure 1 Conceptual model**

****

%3CmxGraphModel%3E%3Croot%3E%3CmxCell%20id%3D%220%22%2F%3E%3CmxCell%20id%3D%221%22%20parent%3D%220%22%2F%3E%3CmxCell%20id%3D%222%22%20value%3D%22Crowd%20support%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%2220%22%20y%3D%22110%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%223%22%20value%3D%22Referee%20Bias%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22340%22%20y%3D%2220%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%224%22%20value%3D%22Team%20Performance%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22660%22%20y%3D%22110%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%225%22%20value%3D%22Share%20of%20Foreigners%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22200%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%226%22%20value%3D%22Crowd%20intensity(occupancy%26lt%3Bbr%26gt%3B%2Fsize)%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22340%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%227%22%20value%3D%22Team%20Age%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22480%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%228%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BentryX%3D0%3BentryY%3D0.5%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20target%3D%224%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22140%22%20y%3D%22140%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%2270%22%20y%3D%22320%22%20as%3D%22targetPoint%22%2F%3E%3CArray%20as%3D%22points%22%3E%3CmxPoint%20x%3D%22140%22%20y%3D%22140%22%2F%3E%3C%2FArray%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%229%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D1%3BexitY%3D0.5%3BexitDx%3D0%3BexitDy%3D0%3BentryX%3D0%3BentryY%3D0.5%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20source%3D%222%22%20target%3D%223%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22150%22%20y%3D%22160%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22330%22%20y%3D%2250%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2210%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%225%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22260%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2211%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D1%3BexitY%3D0.5%3BexitDx%3D0%3BexitDy%3D0%3BentryX%3D0%3BentryY%3D0.25%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20source%3D%223%22%20target%3D%224%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22650%22%20y%3D%22120%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2212%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3B%22%20edge%3D%221%22%20source%3D%226%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22230%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2213%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%227%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22540%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2214%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%225%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22210%22%20y%3D%22110%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2215%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3B%22%20edge%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22240%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22270%22%20y%3D%2280%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2216%22%20value%3D%22%22%20style%3D%22endArrow%3Dnone%3Bdashed%3D1%3Bhtml%3D1%3B%22%20edge%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22620%22%20y%3D%2250%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22620%22%20y%3D%2250%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3C%2Froot%3E%3C%2FmxGraphModel%3E%3CmxGraphModel%3E%3Croot%3E%3CmxCell%20id%3D%220%22%2F%3E%3CmxCell%20id%3D%221%22%20parent%3D%220%22%2F%3E%3CmxCell%20id%3D%222%22%20value%3D%22Crowd%20support%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%2220%22%20y%3D%22110%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%223%22%20value%3D%22Referee%20Bias%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22340%22%20y%3D%2220%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%224%22%20value%3D%22Team%20Performance%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22660%22%20y%3D%22110%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%225%22%20value%3D%22Share%20of%20Foreigners%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22200%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%226%22%20value%3D%22Crowd%20intensity(occupancy%26lt%3Bbr%26gt%3B%2Fsize)%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22340%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%227%22%20value%3D%22Team%20Age%22%20style%3D%22rounded%3D0%3BwhiteSpace%3Dwrap%3Bhtml%3D1%3B%22%20vertex%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20x%3D%22480%22%20y%3D%22240%22%20width%3D%22120%22%20height%3D%2260%22%20as%3D%22geometry%22%2F%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%228%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BentryX%3D0%3BentryY%3D0.5%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20target%3D%224%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22140%22%20y%3D%22140%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%2270%22%20y%3D%22320%22%20as%3D%22targetPoint%22%2F%3E%3CArray%20as%3D%22points%22%3E%3CmxPoint%20x%3D%22140%22%20y%3D%22140%22%2F%3E%3C%2FArray%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%229%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D1%3BexitY%3D0.5%3BexitDx%3D0%3BexitDy%3D0%3BentryX%3D0%3BentryY%3D0.5%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20source%3D%222%22%20target%3D%223%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22150%22%20y%3D%22160%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22330%22%20y%3D%2250%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2210%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%225%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22260%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2211%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D1%3BexitY%3D0.5%3BexitDx%3D0%3BexitDy%3D0%3BentryX%3D0%3BentryY%3D0.25%3BentryDx%3D0%3BentryDy%3D0%3B%22%20edge%3D%221%22%20source%3D%223%22%20target%3D%224%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22650%22%20y%3D%22120%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2212%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3B%22%20edge%3D%221%22%20source%3D%226%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22230%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2213%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%227%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22540%22%20y%3D%22140%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2214%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3BexitX%3D0.5%3BexitY%3D0%3BexitDx%3D0%3BexitDy%3D0%3B%22%20edge%3D%221%22%20source%3D%225%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%2220%22%20y%3D%22370%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22210%22%20y%3D%22110%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2215%22%20value%3D%22%22%20style%3D%22endArrow%3Dclassic%3Bhtml%3D1%3B%22%20edge%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22400%22%20y%3D%22240%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22270%22%20y%3D%2280%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3CmxCell%20id%3D%2216%22%20value%3D%22%22%20style%3D%22endArrow%3Dnone%3Bdashed%3D1%3Bhtml%3D1%3B%22%20edge%3D%221%22%20parent%3D%221%22%3E%3CmxGeometry%20width%3D%2250%22%20height%3D%2250%22%20relative%3D%221%22%20as%3D%22geometry%22%3E%3CmxPoint%20x%3D%22620%22%20y%3D%2250%22%20as%3D%22sourcePoint%22%2F%3E%3CmxPoint%20x%3D%22620%22%20y%3D%2250%22%20as%3D%22targetPoint%22%2F%3E%3C%2FmxGeometry%3E%3C%2FmxCell%3E%3C%2Froot%3E%3C%2FmxGraphModel%3E

## Data and Methodology

## 3.1 Data collection

The website ([http://www.football-data.co.uk](http://www.football-data.co.uk/data.php)) provides weekly updated datasets for all important leagues around the world. The datasets include statistics on full-time and half-time results, shots, shots on targets, number of fouls, number of yellow and red cards and corners for each team on match level. Our sample includes all the matches played from season 2018/19 onwards. We chose 2018/19 as cutoff as around this time the VAR got introduced most competitions (Farrell, 2019). The VAR has major implications especially in the realm of crowd supports influence on decision making by referees. The referee can be overruled by the video referee, often located outside of the stadium, and thus less influenced by crowd noise. We picked the top 10 leagues in european soccer to be included in our sample. Afterwards we decided to remove the Russian Premier League from our analysis given the fact that data on our moderating variables was missing for the Russian Premier League. The 9 leagues incorporated are: Dutch Eredivisie, German Bundesliga, Portuguese Primeira Liga, The Turkish Super Lig, Belgian Jupiler League, French Ligue 1, English Premier League, Spanish Primera Division and the Italian Serie A. We added a dummy variable equal to 0 for the leagues where the VAR had not yet been introduced yet in a particular season. This was only the case for the 2018/19 Premier League and 2018/19 Primeira liga seasons. For data on our 4 moderators - team average age, percentage of foreigners, crowd occupancy, and crowd size - we consulted the website of transfermarkt.com.

We include two measures of team performance in our model; the difference between points obtained by the home and away team and the difference in goals scored by the home and away team. Points obtained by teams is the primary measure of match outcome in soccer. We also use goal difference because information on goal difference in a match can give insight into the magnitude of a win. A 2-1 win and a 5-0 win both result in the same points difference but represent completely different matches.

We also need to control for potential endogeneity resulting from other variables affecting team performance. Team performance depends heavily on the quality of the team (Lago-Peñas & Lago-Ballesteros, 2011). Thus a measure of team strength should be included in the model as control variable. The most comprehensive measurement of team strength we know is the Soccer Power Index(SPI) used by FiveThirtyEight[[1]](#footnote-0). Their SPI index is constructed by computing an offensive rating and defensive rating. This rating is equal to the number of goals expected to score/concede by the team against an average opponent on neutral ground. The SPI then is the percentage of points that the team will take if the match against an average team on neutral ground is played. FiveThirtyEight’s public github repository provides weekly updated dataset with SPI data. Another variable influencing team performance is the importance of a match. Link & de Lorenzo (2018) discovered that players make more faster runs and more fouls in matches that were influential on final ranking compared to matches which were not. Intuitively it makes sense that a team will attempt to peak for a match that is important and be less focused and motivated when the outcome of the match has no consequences. The SPI dataset also includes a measure of match importance for both the home and away team. The match importance is calculated through expected probabilities of the match outcome that would alter the ranking of the team in the competition.

Since our sample includes multiple leagues, we need to account for potential country specific differences that could confound the relationships within our model. Therefore we add a dummy variable for each league in our regression model to control for these league specific differences. Additionally, for our regression model for referee bias, a confounding variable could be the dominance in a match. Dominant teams with more possession will make less fouls within a match (McCarick et al 2020; Goumas, 2014b). If home teams play more attacking soccer and therefore need to commit less fouls, the referee bias we found might simply be a result of the playing style and not from an actual bias in referee decision making. Therefore we include the difference in shots between home and away teams into our model.

## 3.2 Variable Operationalization

In table 2 we provide an overview of the main variables included in our analysis with an explanation on how the variable is operationalized. Our unit of analysis is a match between two soccer teams, with the team playing at home called “Home Team” and the team playing away “Away Team”. For many of the variables we use the differences between home and away metrics within the match to reduce the number of variables used in our models. For the majority of these differences, we calculate the difference by subtracting away values from home values. For example, goal difference is calculated as number of goals scored by the home team minus number of goals scored by the away team. However, for referee decisions we decided to reverse the calculation to facilitate interpretation of outcomes. When calculated in this way, a positive difference in cards implies higher cards for away teams, which can be seen as a positive bias towards home teams.

**Table 2 Variable operationalization**

## 3.3 Descriptive statistics

Table 3 provides the descriptive statistics for the variables mentioned in section 3.2.

The data set contains match data for 8137 matches played in 9 major leagues of Europe. We observe that home teams on average score .292 goals more and collect .365 more points compared to away teams. Furthermore, the table reveals that home teams on average receive .185 fewer yellow cards and .026 fewer red cards than away teams, while committing .108 fouls less on average. Appendix 1a contains the full summary statistics table.

**Table 3 Summary Statistics**



## 3.4 Home advantage pre and post covid-19

For the entire sample, there is a clear home advantage, with more points, and more goals for home teams over the past 3 years. However, solely based on the entire sample we cannot conclude whether home advantage has remained during the covid-19 pandemic. Therefore we split the dataset into two different datasets. One dataset containing matches only played before the start of the covid-19 pandemic and the other only containing matches played during the covid-19 pandemic. 2,996 of the matches have been played behind closed doors and 5,141 were played with spectators. We proceed with statistical tests to examine whether home advantage has changed significantly during the covid-19 pandemic. Due to the continuous nature of our variables we use a Mann-Whitney U test, which handles our non-normal data better than traditional t-tests. For the percentage of home and away wins, which are coded as categorical variables, we perform a chi square test of comparison. Table 4 contains the mean pre and post covid-19 values for our variables of interest accompanied by p-values for the null hypothesis of equal distributions.

**Table 4 mean level differences pre and post covid**

*****Note*: Statistic is the Mann-Whitney Estimate, tests 0 hypothesis of equal distribution. Significance levels: *p* < .05\* , *p* < .01 \*\*, *p < .001 \*\*\**

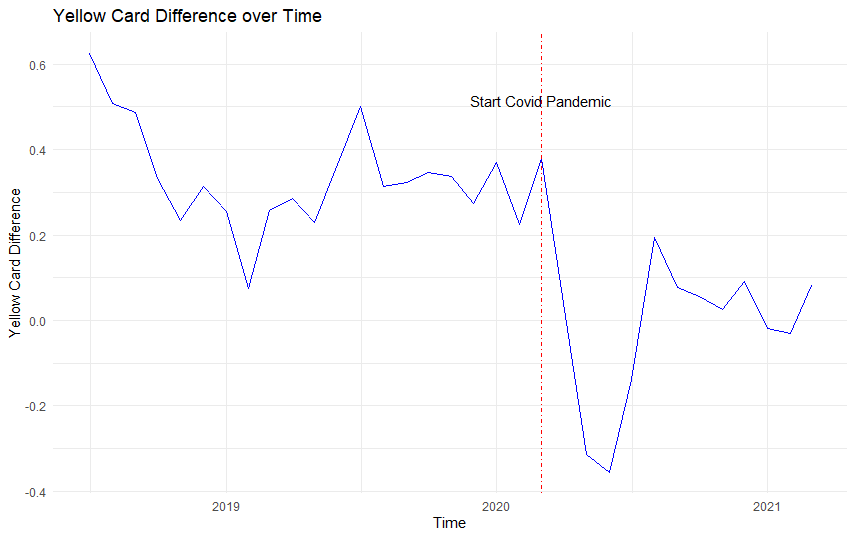
The gap between home and away teams for both expected and actual goals has declined significantly following the disappearance of crowd support. Table 4 proves that with supporters, home teams scored .36 more goals than away teams per match, a number that fell to .17 (*p* < .001) when crowd support was absent. For expected goals, we see a similar drop from .31 to .16 (*p <* .001). Table 4 also shows the differences in referee punishments before and after covid. In matches without crowd support, the difference in yellow cards has been reduced by .3 (*p* < .001), the difference in red cards by.02 (*p* = .010) and the difference in fouls has been reduced by .5 (*p* < .001) relative to away teams compared to matches with crowd support. These results imply that away teams have been punished less severely relative to the home team in matches without crowd support. These differences are depicted in figure 2, clearly highlighting the change in referee decisions during the pandemic.

**Figure 2 Comparison of Foul, Yellow card and Red card gap**



*Note:* for each pair of bars, the left bar represents the level during the covid pandemic, with the right bar representing the levels before the pandemic. From left to right, we compare Fouls, red cards and yellow cards. Error bars represent 95% confidence interval for mean value.

**Figure 3 The difference in Yellow Cards over time**



*Note*: the blue line indicates the average yellow card difference per month. The vertical red dashed line indicates the start of the covid-19 pandemic.

Figure 3 shows the trend in the differences in yellow cards between home and away teams. We observe a significant dip in the few months after the covid-19 pandemic but interestingly towards the end of 2020 we see the difference decreasing again, to remain relatively stable in 2021. This suggests that the effect of missing supporters was the heaviest right after the restart, with referees adjusting to the new situation afterwards. We have to be cautious Appendix 3b shows a similar but less extreme drops in foul difference immediately after the start of the pandemic. For red card difference however, as we can observe from appendix 3a, there is no clear pattern visible. This is probably caused by the low number of red cards per game.

Our statistics on differences show that there has been a significant reduction in differences between home and away teams on various metrics. However, it does not show whether these differences stem from reduced home team performance or increased away team performance. Or in the case of referee decisions, whether the differences come from a reduced punishment for away teams or increased punishment for home teams. We delve deeper into home and away team data to uncover these patterns. Table 5 presents the results for referee decisions. The results in the table are revealing in several ways. Firstly, it seems that rather than punishing home teams more severely, the gap in cards has mainly been reduced by a more lenient attitude towards away teams. They receive significantly lower numbers of yellow and red cards since the start of the pandemic, while not making significantly more fouls. Secondly, somewhat contradictory to the first finding is that despite the significant increase in fouls made by the home team, the number of cards the home team received remained relatively stable. This could be interpreted as less severe punishment for fouls, or perhaps that home teams made more minor fouls that were not serious enough to be a bookable offense.

**Table 5 mean levels referee metrics pre and post covid**

*****Note:* Statistic is the Mann-Whitney Estimate, tests 0 hypothesis of equal distribution. Significance levels: *p* < .05\* , *p* < .01 \*\*, *p < .001 \*\*\**

To distinguish between differences in home and away performance, we provide the same table for different metrics of match outcome and team performance from both a home team and away team perspective. As shown in table 6, the performance of home and away teams has changed significantly after the start of the covid-19 pandemic. The percentage of wins at home has declined 5 percent points from 45 percent pre covid to 40 percent post covid (*p <*.001*)* whereas the percentage away wins have rose from 30 to 34 percent (*p <*.001*)*. The chi square proportion for difference in proportions home and away wins gives a chi square value of 74.252 (*p <*.001) which suggests that home advantage indeed has significantly decreased following the exclusion of supporters.

Also number of goals and points for home teams declined substantially. Away teams fare better in games behind closed doors compared to games with fans. With crowd support, away teams on average collected 1.15 points per game, scoring an average of 1.22 goals per game in the process. When crowd support is not present however, away teams have increased their points per game to 1.27 (*p <*.001) also increasing the amount of goals scored on average by 0.08 to 1.30 goals per game(*p <*.001). There is clear evidence of a decrease in home advantage following the exclusion of fans.

**Table 6 mean levels performance metrics pre and post covid**

*Note: ^ = chi square statistic, others = Mann-Whitney estimate. ,* Significance levels: *p* < .05\* , *p* < .01 \*\*, *p < .001 \*\*\**

We also examine differences in expected goals. Metrics such as goals and points provide information on team performance but can be dependent on luck in finishing. Expected goals is a metric that calculates the quality of chances created by a team over the entire match and therefore is very suitable to evaluate team performance. The decrease in home advantage seems to come from both an increase in away team performance and a decrease in home team performance. With crowd support, home team’s actual goals slightly outperformed their expected goals, scoring 1.58 goals per match where 1.57 would be expected, whereas the away team’s actual goals where slightly below their expected goals value (1.22 to 1.25).

Both goal values have converged closer to their expected goals value after the lockdown, which has lead to a significant reduction in home advantage. Also, the decrease in expected goals for home teams is bigger, with an expected goals drop of over .1(*p* <.001) goals per match, compared to a .05 increase in expected goals for away teams(*p* = .044). This suggests that decreased home advantage in a situation of no crowd support is to a greater extent caused by a drop in home team performance than to an improved away performance.The differences in goals and expected goals are displayed in figure 4.

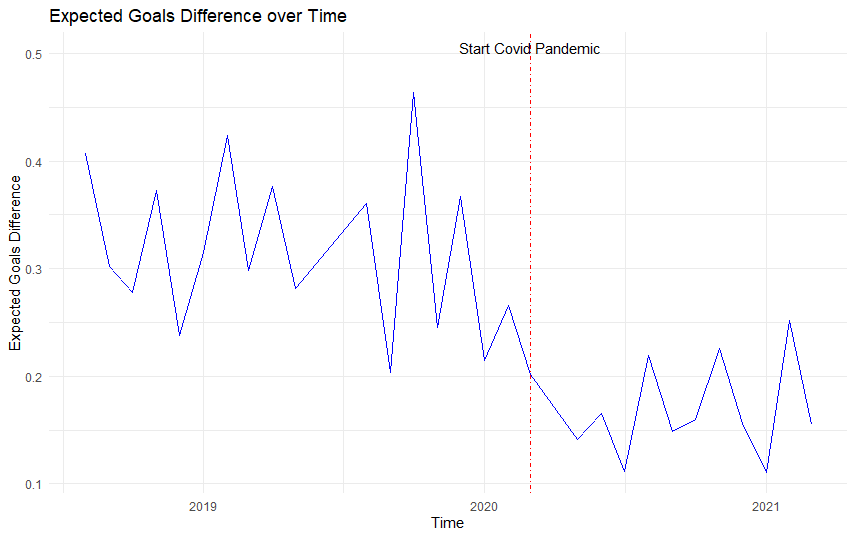
**Figure 4 comparison of goal and expected goals gap**



*Note:* for each pair of bars, the left bar represents the level before covid and the right bar the levels after covid. From left to right we compare Expected goals and Goals. Error bars represent 95% confidence interval for mean value.

On average, we observe a significant decline in the expected goals difference and goal difference in games without crowd support. In the next graph we examine the development of expected goals difference over time. From figure 5 we observe a sharp drop immediately after the start of the pandemic. However, contrary to referee decisions, for expected goals we do not see a movement back towards pre-covid levels as expected goals difference remain quite low. For points and goal difference however, as shown in appendix 3c and 3d we see a sizeable increase after the initial drop.

**Figure 5 Expected goals difference**

****

*Note:* the blue line indicates the mean difference in expected goals per month. The red dashed line indicates the start of the covid pandemic.

## 3.5 Model

Our model free evidence indicates a significant reduction in home advantage in matches played without crowd support. However, based on these numbers we cannot make conclusions about causal mechanisms. In the next chapter we examine whether the absence of crowd support caused the drop in home team performance relative to away team performance. In our model, we scrutinize the possible mediating role of referee decision making in this process. Moreover, we check for the effects of multiple potential moderating variables: occupancy, crowd size, team age and the ratio of foreigners playing for a team. Finally, we add the control variables for team strength, match importance and VAR availability. Our first equation involves the path between our independent variable crowd support and our mediating variable, referee bias. Referee bias will enter the model as a latent construct measured from foul and yellow card decisions by the referee since the Cronbach’s Alpha for a latent construct comprising of the observed variables yellow cards and fouls is considerably higher(.32 vs .25) than for a measure including red cards as well. Appendix .. shows the scale reliability analysis for our latent construct. Despite our value for alpha not reaching the minimum value of .7 introduced by Nunnally (1978), we decided to keep the latent construct within our model. We deem it important that an abstract concept such as referee bias is actually measured through combining multiple different variables togethe We perform a robustness check for our model by deploying the same model with only the observed variable yellow cards as measure of referee bias.

**Table 7**



Below are the regressions equations for our model..

**Equation 1:** Referee Bias: *β0 + β1Covid + β2OccupancyRate + β3ForeignersShareDifference + β4Crowdsize + β5Covid\*OccupancyRate + β6Covid \* ForeignersShareDifference + β7Covid \* Crowdsize + β8RatingDifference + β9ImportanceDifference + β10VAR +β11ShotsDifference + + ɛ*

The second equation within our model is the path between our independent variable crowd support and our dependent variable team performance. We use 2 different measures of team performance, goal difference and points difference to increase the robustness of our results.

**Equation 2:** Team Performance *= β0 + β1Covid+ β2OccupancyRate + β3ForeignersShareDifference + β4Crowdsize + β5AgeDifference + β6 Covid\*OccupancyRate + β7 Covid \* ForeignersShareDifference + β8 Covid \* Crowdsize + β9 Covid \* AgeDifference + β10RatingDifference + β11ImportanceDifference + β12VAR +β13Referee bias + + ɛ*

Since we want to assess multiple relationships and also include a latent construct within our model we deem a SEM model most suitable for our analysis. Hair, Black, Babin, Anderson and Tatham (2014) state the main advantage of SEM models vis-à-vis other powerful techniques such as multiple regression and factor analysis is that SEM allows for the examination of multiple relationships together simultaneously. Furthermore, SEM allows both latent and observed constructs in the model. We use the Lavaan package in R to conduct our analysis.

## Analysis and Findings

In this chapter we deploy our previously defined model on our dataset in order to examine the causal relationships within our conceptual framework. We estimate a Structural Equation model, and therefore we start by checking the assumptions. We use the method proposed by Pena & Slate (2006) who introduce a global test for the major assumptions for linear models.

## 4.1 Assumptions

**Normality:** The Anderson-Darling tests show that our variables are not normally distributed, the results of these tests are shown in appendix 2a. To assess multivariate normality we use the globalized gvlma test, which tests multiple assumptions of linear regression in combination to eachother. Pena & Slate (2006) mention, assumptions are often interrelated with eachother, which makes the combined test powerful as it takes these interactions into account. The test result in appendix 4a, 4b and 4c show that similar to univariate normality, the assumption of multivariate normality is violated. However, the central limit theorem states that for large samples with sample size over 30, these deviations from normality are not problematic (Brosamler, 1985). With our sample size of over 8,000, we can safely ignore the normality assumption and proceed further.

**Multicollinearity**: When interpreting the correlation table in appendix 7 we see a few high correlations. Notably between our control variable Rating difference and our dependent outcomes, as well as correlation between Rating Difference and Importance Difference. Furthermore, there exists a correlation of 0.75 between occupancy rate and average attendance, which makes sense since these variables are closely related. We calculate the variance inflation factors for our model variables, the output of which is shown in appendix 8a and 8b. A frequently used benchmark for VIF values is 10 (Vittinghof, Glidden, Shiboski and McCulloch , 2011). James, Witten, Hastie and Tibshirani(2013) on the other hand pose a value of VIF over 5 to be problematic. The highest VIF factor within our model is 3.09 , therefore we can safely assume that multicollinearity is not a problem within our model.

**Homoscedasticity:** The assumption of homoscedasticity is important for the ability to interpret standard errors of our model. The gvlma test also tests for homoscedasticity within our sample, non-constant error variance test to test for homoscedasticity. As we observe from appendix 4a and 4b, for both our equations, the test statistic is insignificant (*p* = .95 and *p* = .94) and as such we conclude that the assumption of homoscedasticity is met in our model.

We mean centered all our moderating variables to ease interpretation of the moderating effect on the relationship between crowd support and team performance (Iacobucci, Schneider, Popovich & Bakamitsos, 2017). Additionally, since some of the moderating variables within our model are measured on different scales, we standardize them as well.

## 4.2 Results overview

Table 8 shows the results of our moderated mediation analysis

**Table 8 Model estimates**

****

The main goal of our study was to examine whether crowd support has a significant effect on team performance. In our first hypothesis we formulated the expectation that crowd support significantly impacts team performance. In our model we find evidence to support this hypothesis, the coefficient of -.284 (*p* < .001) implies that the disappearance of crowd support decreased the points gap between home and away teams by .284 points per match, when controlling for various factors such as match importance and team rating.

In hypothesis 1b we defined the moderating effect of crowd size, with larger crowds magnifying the effect of crowd support on team performance. Supporting the findings of Fischer & Haucap (2020) but contrary to the results of among other (Agnew & Carron, 1994) we find no evidence for the moderating effect of crowd size on the relationship between crowd support and team performance.(*β* = .041, *p* = .536*).* Therefore we reject the hypothesis.Apparently a larger crowd size does not result in a larger boost from crowd support in team performance. The presence of a home crowd increases home team performance but whether this is a crowd of 500 or 50,000 might be not important for soccer players. This could explain the fact that home advantage still exists at amateur levels, with very small crowd sizes generally.

Contrary to crowd size, crowd occupancy do seems to matter for crowd supports effect on team performance. We find that for high levels of crowd occupancy, the loss of performance following the absence of crowds is as high as .73 For low levels of occupancy however, this effect is only .34. Teams with higher occupancy rates got hit harder by the fact that crowd support disappeared during the covid-19 outbreak. This also implies that in matches where crowd support is present, they have a bigger home advantage all else equal, and thus perform better compared to teams with low occupancy rates.

Contrary to our expectations we conclude that player age is not a significant moderator of the relationship between crowd support and team performance (*β* = -.074, *p* = .178*).*  In other words, there is no significant difference in the influence of crowd support on team performance for younger or older teams. However, since previous finding such as those by van der Ven(2016) were relatively small, the insignificance of this moderator is not entirely unexpected. Moreover, the average age of the teams were quite close to each other, with the difference between 25th and 75th percentile being 1.9 years. Perhaps due to the nature of the data we used for team age, there was low variation in team performance, since a majority of the teams would have both old and young players.

We also do not find evidence to support our hypothesis that the share of foreigners moderates the relationship between crowd support and team performance (*β* = -.011, *p = .847).* Apparently, the share of foreigners does not affect stadium atmosphere significantly enough to warrant larger effects on referee bias. Should crowd atmosphere really be affected by the share of foreigners, the total change in stadium atmosphere is apparently not large enough to significantly influence team performance.

Our 2nd group of hypotheses concerned the mediating effect of referee bias on the relationship between crowd support and team performance. The coefficient for the indirect effect of crowd support on team performance with a value of -.019(*p* = .006) signals that referee bias indeed mediates the relationship between crowd support and team performance. In the post-pandemic situation with no crowd support, the actual referee bias decreased and therefore reduced the point gap between home and away teams. However the percentage of the total effect of crowd support on team performance that is mediated by referee bias with 6.69% Is quite low. Crowd occupancy and crowd size do not seem to matter for referee bias however. Apparently, the sole fact that a home crowd is present leads to an increased referee bias in favour of home teams, however, the number of home fans, the share of foreigners and the degree to which the stadium capacity is filled all do not influence the bias of referees.

The control variables importance difference and rating difference have face-valid effects. All else equal, a higher rating(*β* = .059, *p <* .001) for the home team and a higher match importance for the home team(*β* = .003, *p <* .001) increase the points difference between home and away teams. This make sense, teams that are better will collect more points on average, and a team for whom the match is more important will be more motivated to play well and win the match. The league in which the match was played turns out to be not a significant predictor of points difference. This implies that the effects of crowd support on team performance are consistent across competitions.

**Table 9 Fit indices**



Table 9 provides an overview of the fit indices for our moderated mediation model. Our chi square statistic of 206.452 (*p* <.001). suggests a discrepancy between the actual and predicted observations.. However, Kyriazos (2018) mentions the sensitivity of the chi square statisticin relation to sample size. Since our sample of 8,054 is quite large we decided to ignore the chi square statistic measure of fit. To benchmark the goodness of fit of our model we use the standard that Dion(2008) provides in his paper on preferred values for the major fit measures of SEM. He proposes that the TLI measure should approach 1 preferably, our TLI has a value of .903 which is not great but not terrible either. Secondly, the CFI should have a minimum value of .90 for a model to be considered a “good fit”, Our CFI measure has a value of .964, which is above the desired cut-off value, we conclude that for this second metric our model seems to be acceptable. Thirdly, Root mean square error of approximation(RMSEA) is a measure of the difference in the sample data with what would be expected in the situation of a correct model. Thus for RMSEA the lower value the better, with .05 the generally accepted maximum value. Our value of .027 falls well within the acceptable range. The final measure we use is SRMR, which is the standardized difference between predicted and observed correlations (Taasoobshirazi & Wang, 2016). Similar to RMSEA, the cut-off value is .05 , with a value of .009 our model appears as good fit for the data. Overall our model seems to fit reasonably well, passing 3 out of 4 fit statistic cut-offs.

## 4.3 Robustness check

To ensure the robustness of our results, we deployed our model on a second measure of team performance. Points difference between home and away teams is the primary outcome metric of a soccer match, We also are interested in goal difference because not all wins are the same. A narrow win by just one goal and a thumping 4 goal victory both have the same points result but a very different match process. Perhaps while home and away wins are evenly divided, every away win could be a scrappy 1-0 victory whereas all home wins are convincing 3-0 victories. In this case there is still a home advantage. Table 10 shows the results of the same model now regressed with goal difference as dependent variable. Some coefficients slightly change in estimation but all the signs and signficance levels remain similar, such that our findings not only hold for point differences but also for goal differences. This strengthens the validity of our model findings. Furthermore, we check the robustness of our results for different measures of referee bias. As mentioned before, the internal scale reliability of our latent construct is quite low, therefore we check our model results with yellow card difference as mediating variable. The results are shown in appendix .. for both points and goal difference. Some coefficients change slightly but again we see no major differences. Additionally, appendix … shows the results for our model when using a latent construct which also includes red card decisions. This model scores considerably lower on model fit(appendix … ) which strengthens our decision to remove red cards from the model.

**Table 10 Goal difference**



Table 11 depicts the fit measures for our 2nd model. Similar to our first model we decided to ignore the significant chi square statistic as a consequence of our large sample size. For the other 4 fit measures, our model fits well to the data according to 3 out of 4. Therefore we deem our model to fit the data sufficiently.

**Table 11 Fit measure**



## 4.4 Moderating effect occupancy

Having ensured the robustness of our results, we scrutinize the effect of our only significant moderator crowd occupancy on the relationship between crowd support and team performance. We use the simple slope analysis, introduced by O’Connor (1998) to compare the differences between high and low levels of crowd occupancy. The standard levels used in the literature for high and low levels of a variable are one standard deviation below and above the mean (O’Connor, 1998). Table 12 shows the differences in total effect of crowd support on team performance for high, low and medium levels of crowd support.

**Table 12 High versus Low Occupancy rates**

****

For both our measures of team performance, the effect of crowd support on team performance is significant for high and mean levels of Occupancy. However, for low levels of crowd occupancy the effect is not significant. This implies that teams with low stadium occupancy rates did not see a significant alteration in their performance during the covid-19 pandemic. For mean levels of occupancy rates the decrease in the points gap during the pandemic is -.284. .For high levels of crowd occupancy, this effect is almost twice as large with a decrease of around -.481 Thus, teams with high occupancy rates pre covid suffered more from the absence of crowd support than teams with low occupancy rates. Conversely, teams with higher occupancy rates have a bigger home advantage when crowd support is actually present. In practice, this could mean that over the course of a season of 38 games, a team with a high occupancy rate compared to a team with a lower occupancy rate could collect around 4.56 (19 home games \* ((.481-0)/2)) points more in home games purely because of the higher occupancy rate . With differences between teams close to eachother in the league table often being very marginal, these 4/5 points could be the difference between qualifying for the Champions League rather than the Europa League or between relegation and staying up. Each of the outcomes have huge financial consequences. The Champions League guarantees clubs an income of €15.25 million for participation alone with the option of millions in bonuses for winning matches and reaching further into the tournament. The Europa League on the other hand offers a mere €2.92 million starting fee and considerably lower bonuses.[[2]](#footnote-1) Relegation has similar consequences, with Sky sports estimating the estimated loss in tv revenues for premier league clubs relegating to the championship to be at least £50 million.[[3]](#footnote-2) A huge number that does not even account for potential losses in commercial deals.

For Goals, we see that for low levels of crowd occupancy the gap between home and away goals is again not significantly different during the covid-19 pandemic. For teams with high level of crowd occupancy the gap in goals decreases by .359 per match. Again, we see that teams with high occupany rates have been more heavily affected by the exclusion of home supporters.

In the next chapter we summarize our main findings and provide a schematic overview of the answers to our hypothesis. Subsequently we use our results to generate managerial recommendations for managers working at soccer clubs. Finally we provide the limitations of this study plus some possible areas for future research.

## Conclusion and Discussion

## 5.1 Conclusions

We document the effect of crowd support on team performance. We used the extraordinary opportunity of the covid-pandemic to systematically scrutinize a large number of matches played without fans in order to directly assess the impact of crowd support on team performance. We first use several descriptive metrics of referee decisions and team performance to uncover whether they were significantly different in matches with crowd support compared to matches without crowd support. We found that home advantage has decreased significantly but still exists without crowd support. Both home and away team performance has been affected following the absence of crowd support. Home teams score less, create less chances and take less points per game. Away teams, seem to have increased their effectivity, scoring significantly more goals and having a significantly higher epxected goals tally without having more shots or more shots on target.

Referees give less cards to away teams in matches without supporters while cautioning home teams similarly in matches with and without crowd support. The number of fouls for away teams remained stable in matches with and without crowd support while the number of fouls for home teams increased following the exclusion of away supporters.

Subsequently we proceeded to assess causality links between crowd support, referee decisions and team performance while including several marketing related moderating variables. Given the complexity of our model we used a structural equation modelling approach to effectively model the relationships between our numerous observable variables and the latent construct referee bias.

Table 13 provides a summary of our hypotheses generated in chapter 2 and the evidence we find for the hypotheses in our analysis.

**Table 13 results summary**



To summarize our findings, we find that crowd support significantly affects team performance, with the points gap reduced by .284 between home and away teams whereas the goals gap reduced by .202 following the exclusion of supporters. We further propose that crowd occupancy but not absolute crowd size affects the degree to which crowd support influences team performance. With higher occupancy rates the points gap was reduced by.481 during the covid-19 pandemic whereas the goal gap was reduced by.359.. This in comparison to lower occupancy rates for which the points and goal gap are not significantly different before and during the covid-19 pandemic.

We also find that crowd support not only directly influences team performance but also affects team performance indirectly through referee decisions. We do not find evidence that crowd size and crowd occupancy influence the relationship between crowd support and referee decisions. Furthermore, we do not find evidence to support the claims that the share of foreigners and the average age of teams influences the effect of crowd support on team performance, nor the effect of crowd support on referee decisions.

## 5.2 Managerial recommendations

The results of our study illustrate the significant effect of crowd support on referee decisions and team performance. Future researchers can take the notion of the influence of crowd support on team performance in future research. Having directly estimated the relationship between crowd support and team performance by using the current circumstances in soccer, we add to similar studies performed to strengthen the call that crowd support indeed affects team performance. Marketing managers can use our insights to incorporate performance effects of their decisions into their decision making process. Our research shows that not crowd size but crowd occupancy matters for home advantage. The most important task for a marketing manager is to deploy the right marketing strategies to fill the stadium as much as possible. Marketing managers who can effectively bring fans to the stadium can positively impact team performance, resulting in a few extra points per season. These few points could have a significant difference however on final outcome. In the league tables of 2018/19, for all of the leagues in our sample, four extra points would have helped at least one and often more of the relegated teams to stay up in the first division. Similarly, for all of the leagues apart from the Dutch, Belgian and Portuguese ones the teams without European spots or with Europa League spots would have secured a Champions League spot instead. As we mentioned before, reaching the Champions League or staying up can have tremendous financial consequences.

Furthermore, a lot of clubs spend millions on the development of larger stadiums to increase attendances and consequently ticket revenues. Given the impact of occupancy rates, rather than trying to increase revenues by moving to bigger stadiums. Perhaps a more viable strategy would be to remain in the same stadium to increase or ensure high occupancy rates. The remaining fans can be catered through attractive online engagement strategies and match coverage on TV. Future developments such as streaming services for broadcasting soccer matches could provides opportunities for marketing managers to increase customer revenue while optimizing stadium occupancy rates.

Another interesting finding is the lack of influence of foreign players on crowd atmosphere. As a consequence global marketing strategies to increase the global presence and increase associated revenue streams through for example social media engagement can then be used to generate more revenue. As such, using foreign players from exotic countries to increase fan engagement could be a viable option for soccer clubs. Examples nowadays include Ajax increasing their presence in Brazil through the purchase of local talents David Neres and Anthony. The welcome song that Ajax posted was well received and gained global attention, increasing the engagement with fans around the world. Since our results show that the share of foreigners does not significantly alter the influence of crowd support on team performance, soccer clubs can buy these foreign players without having to worry about eventual negative side effects on crowd atmosphere and team performance.

## 5.3 Limitations and areas for future research

In the context of the widely used expression “no research is perfect research”, our study imposes several limitations, and subsequently, areas for future research.

Firstly a few of our moderators were limited in data availability. For our occupancy rates and average attendances, we used season average attendance data to compile these statistics while these statistics obviously differ per match. Some matches attract high attendance where other matches attract lower attenances over the course of a season. Secondly, our metric for the share of foreigners is calculated as the total playing time over the course of the season for foreign players divided by the total playing time over the course of the season for local players. This metric can be different per match as well. In similar fashion, average age is calculated as the average age of all players used over the course of the season instead of the average age of all players used within a particular match. We decided to use these season-wide measures due to the availability of the data and time concerns. Per match data would have been very cumbersome to accumulate. Interesting extensions to improve the reliability of the effects of the previously mentioned variables could be web scrapers scraping match specific data for more precise data on these variables, to then conduct a similar study to this one.

/

A second area for future research could be a factor analysis to measure team performance. We used 2 seperate measures of team performance with goal difference and points difference. However, there are multiple other measures of team performance. Outcome level measures such as we investigated and deeper lying performance metrics such as shots, possession and expected goals. An interesting path for future research could be to investigate whether a latent construct of team performance can be constructed from all these different measures of team performance.

Thirdly, a possible extension of this study could be to delve deeper in referee characteristics that influence referee bias. Since it is established that crowd support influences referee behavior, kowing which type of referees are influenced disproportionally by crowd noise could provide useful in preparing referees for the influence of the crowd, as well as selecting the right candidates as potential referees.

Fourthly, in our study we established the importance of occupancy rates on team performance and subsequently briefly touched upon how marketing managers play a role in attendance rates. Future studies could include more extensive research on the factors that drive fan behaviour and stadium attendance. A deeper knowledge on what types of fans come to stadiums can help managers improve customer targeting and improve the marketing efforts for soccer clubs. Improving the ability to attract customers to the stadium as well as an increased fan engagement and consumption.

Fiftly, our cronbach’s alpha for our latent construct is quite low. We decided to keep the measure since we deem a latent construct more appropriate to measure a concept such as referee bias. Our decision is strengthened by the robustness of our results when checking our model for an observed measure of referee bias by only using yellow cards. However, it still remains a limitation of our study and thus an area where future research is useful. Incorporating extra data on referee decisions such as extra time and penalty’s could prove to be a useful extension of our current data.

Finally, in our study we use share of foreigners to scrutinize whether increased globalization affects the influence of crowd support on team performance. However, a second way through which this increased globalization of soccer clubs could influence atmosphere within the stadium is through the composition of the crowd itself. “football tourism” is a well known term which signifies the influx of global supporters at renowned European clubs at home matches (Graakjær & Grøn, 2020). These supporters their main objective is the experience of visiting a match and not necessarily rooting for the home team to win. An interesting future area for research could be to examine the impact of football tourism on crowd atmosphere and team performance.

# 6. Bibliography

1. Adcroft, A., Teckman, J., & Madichie, N. (2009). Management implications of foreign players in the English Premiership League football. *Management Decision*.
2. Agnew, G. A., & Carron, A. V. (1994). Crowd effects and the home advantage. International Journal of Sport Psychology
3. Allen, M. S., & Jones, M. V. (2014). The home advantage over the first 20 seasons of the English Premier League: Effects of shirt colour, team ability and time trends. *International Journal of Sport and Exercise Psychology*, *12*(1), 10-18.
4. Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. Journal of sports sciences, 25(11), 1185-1194.
5. Brechot, M., & Flepp, R. (2020). Dealing with randomness in match outcomes: how to rethink performance evaluation in european club football using expected goals. Journal of Sports Economics
6. Brosamler, G. A. (1988, November). An almost everywhere central limit theorem. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 104, No. 3, pp. 561-574). Cambridge University Press.
7. Carmichael, F., & Thomas, D. (2005). Home-field effect and team performance: evidence from English premiership football. Journal of sports economics, 6(3), 264-281.
8. Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational research methods*, *11*(2), 296-325.
9. 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.
10. Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York, NY: Harper & Row.
11. Decroos, T., Bransen, L., Van Haaren, J., & Davis, J. (2019, July). Actions speak louder than goals: Valuing player actions in soccer. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1851-1861).
12. Dion, P. A. (2008). Interpreting structural equation modeling results: A reply to Martin and Cullen. *Journal of business ethics*, *83*(3), 365-368.
13. Endrich, M. & Gesche, T. (2020). Home-bias in referee decisions: Evidence from “Ghost Matches” during the Covid19-Pandemic. Economics Letters, 197, 109621.
14. Farrell, M. (2019). A Brief History (And Defense) of VAR. Retrieved from https://sites.duke.edu/wcwp/2019/04/01/a-brief-history-and-defense-of-var/
15. Fédération Internationale de Football Association. (2007). FIFA Big Count 2006: 270 million people active in football. Zurich: Fédération Internationale de Football Association.
16. Fédération Internationale de Football Association.(2019), FIFA activity report 2018. Retrieved from https://www.fifa.com/who-we-are/official-documents/
17. Fischer, K., & Haucap, J. (2020). Does crowd support drive the home advantage in professional soccer? Evidence from German ghost games during the COVID-19 pandemic.
18. Galariotis, E., Germain, C., & Zopounidis, C. (2018). A combined methodology for the concurrent evaluation of the business, financial and sports performance of football clubs: the case of France. Annals of Operations Research, 266(1), 589-612.
19. Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism under social pressure. Review of Economics and Statistics, 87(2), 208-216.
20. Goddard, J. (2006). Who wins the football?. Significance, 3(1), 16-19.
21. Goumas, C. (2013). Home Advantage and Crowd Size in Soccer: A Worldwide Study. *Journal of Sport Behavior*, *36*(4).
22. Goumas, C. (2014). Home advantage in Australian soccer. *Journal of Science and Medicine in Sport*, *17*(1), 119-123.
23. Goumas, C. (2014b). Home advantage and referee bias in European football. *European journal of sport science*, *14*(sup1), S243-S249.
24. Graakjær, N. J., & Grøn, R. (2020). Football tourism and the sounds of televised matches. In *The Routledge Companion to Media and Tourism* (pp. 83-91). Routledge.
25. Gutierrez, D. (2019). *Impact of Special Events and Fan-Player Bonding on Identified Fan Consumption-A Study of Professional Soccer in the United States* (Doctoral dissertation, Creighton University)
26. He, M., Cachucho, R., & Knobbe, A. J. (2015, September). Football Player's Performance and Market Value. In Mlsa@ pkdd/ecml (pp. 87-95).
27. Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The stata journal*, *7*(3), 281-312.
28. Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2017). Mean centering, multicollinearity, and moderators in multiple regression: The reconciliation redux. *Behavior research methods*, *49*(1), 403-404.
29. Inan, T. (2020). The Effect of Crowd Support on Home-Field Advantage: Evidence from European Football. Annals of Applied Sport Science,
30. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.
31. Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(08), 2207.
32. Lago-Peñas, C., & Lago-Ballesteros, J. (2011). Game location and team quality effects on performance profiles in professional soccer. *Journal of sports science & medicine*, *10*(3), 465.
33. Lee, H., Gipson, C., & Barnhill, C. (2017). Experience of spectator flow and perceived stadium atmosphere: Moderating role of team identification.*Sport Marketing Quarterly, 26*(2), 87-98
34. Magen, Z. (1980). Encounter group effects on soccer team performance. *Small Group Behavior*, *11*(3), 339-344.
35. McSharry, P. E. (2007). Altitude and athletic performance: statistical analysis using football results. Bmj, 335(7633), 1278-1281.
36. Müller, O., Simons, A., & Weinmann, M. (2017). Beyond crowd judgments: Data-driven estimation of market value in association football. European Journal of Operational Research, 263(2), 611-624.
37. Neave N, Wolfson S. Testosterone, territoriality, and the 'home advantage'. Physiol Behav 2003; 78: 269-75.
38. Neave N, Wolfson S. The home advantage: psychological and physiological factors in soccer. In: Lavalle D, Thatcher J, Jones MV, Eds. Coping and Emotion in Sport. New York, Nova Science Publishers 2004; 127-44.
39. 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.
40. Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport. Sports Medicine, 28(4), 221-236.
41. Nevill, A., Balmer, N., & Williams, M. (1999). Crowd influence on decisions in association football. The Lancet, 353(9162), 1416.
42. Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. Psychology of Sport and Exercise, 3(4), 261-272.
43. Nunnally, J. C. (1978). *Psychometric theory*: New York : McGraw-Hill, c1978. 2d ed.
44. Peeters, T., & van Ours, J. C. (2021). Seasonal Home Advantage in English Professional Football; 1974–2018. *De Economist*, *169*(1), 107-126.
45. Pena, E. A., & Slate, E. H. (2006). Global validation of linear model assumptions. *Journal of the American Statistical Association*, *101*(473), 341-354.
46. Pinnuck, M., & Potter, B. (2006). Impact of on‐field football success on the off‐field financial performance of AFL football clubs. Accounting & Finance, 46(3), 499-517.
47. Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. Journal of sports sciences, 4(3), 237-248.
48. Pollard, R. (2008). Home advantage in football: A current review of an unsolved puzzle. The open sports sciences journal, 1(1)
49. Pollard, R., & Pollard, G. (2005). Long-term trends in home advantage in professional team sports in North America and England (1876–2003). Journal of sports sciences, 23(4), 337-350.
50. Pollard, Richard,(2006), Home Advantage in Soccer: Variations in its Magnitude and a Literature Review of the Inter-related Factors Associated with its Existence , Journal of Sport Behavior, Vol. 29, Iss. 2, 169-189.
51. Ponzo, M., & Scoppa, V. (2018). Does the home advantage depend on crowd support? Evidence from same-stadium derbies. Journal of Sports Economics, 19(4), 562-582.
52. Poulter, D. R. (2009). Home advantage and player nationality in international club football. *Journal of sports sciences*, *27*(8), 797-805.
53. Rodríguez, P., Humphreys, B. R., & Simmons, R. (Eds.). (2017). The Economics of Sports Betting. Edward Elgar Publishing.
54. Salminen, S. (1993). The effect of the audience on the home advantage. *Perceptual and motor skills*, *76*(3\_suppl), 1123-1128.
55. Samagaio, A., Couto, E., & Caiado, J. (2009). Sporting, financial and stock market performance in English football: an empirical analysis of structural relationships. Journal Department of Management, ISEG/School of Economics and Management, Techincal University of Lisbon, Portugal.
56. Schwartz, B., & Barsky, S. F. (1977). The home advantage. Social forces, 55(3), 641-661.
57. Smith, D. R. (2003). The home advantage revisited: Winning and crowd support in an era of national publics. *Journal of Sport and Social Issues*, *27*(4), 346-371.
58. Strauss, B. (2002). Social facilitation in motor tasks: a review of research and theory. *Psychology of sport and exercise*, *3*(3), 237-256.
59. Sutter, M., & Kocher, M. G. (2004). Favoritism of agents–the case of referees' home bias. Journal of Economic Psychology, 25(4), 461-469.
60. Taasoobshirazi, G., & Wang, S. (2016). The performance of the SRMR, RMSEA, CFI, and TLI: An examination of sample size, path size, and degrees of freedom. *Journal of Applied Quantitative Methods*, *11*(3), 31-39.
61. Theodorakis Y, Goudas M, Papaioannou A, Eds. Book of long papers, 12th European Congress of Sport Psychology; 2007: Halkidiki, Greece: FEPSAC 2007; pp. 57-60.
62. Tilp, M., & Thaller, S. (2020). Covid-19 has turned home-advantage into home-disadvantage in the German Soccer Bundesliga. Frontiers in Sports and Active Living, 2, 165
63. Tucker, W., Mellalieu, D. S., James, N., & Taylor, B. J. (2005). Game location effects in professional soccer: A case study. International Journal of Performance Analysis in Sport, 5(2), 23-35.
64. Unkelbach, C., & Memmert, D. (2010). Crowd noise as a cue in referee decisions contributes to the home advantage. Journal of Sport and Exercise Psychology, 32(4), 483-498.
65. van de Ven, N. (2016). Does Player Age Affect the Home Advantage? Coaches’ Perceptions and Data from Professional Football.
66. Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2011). *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. Springer Science & Business Media.
67. Wetzel, H. A., Hattula, S., Hammerschmidt, M., & van Heerde, H. J. (2018). Building and leveraging sports brands: evidence from 50 years of German professional soccer. *Journal of the Academy of Marketing Science*, *46*(4), 591-611.

# 7. Appendix

**Appendix 1a:** Appendix 1a shows the summary statistics of the total samples for all the variables within our dataset. Home teams overall perform better than away teams, collecting more points, taking more shots and more shots on target and scoring more goals and expected goals.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Summary Statistics total sample** | | | | | | | | |
|  | **N** | **Mean** | **SD** | **Min** | **Pctl(25)** | **Median** | **Pctl(75)** | **Max** |
| Goal Difference | 8,137 | 0.292 | 1.846 | -13 | -1 | 0 | -1 | 10 |
| Points Difference | 8,137 | 0.365 | 2.571 | -3 | -3 | 0 | 3 | 3 |
| Foreigners Share Difference | 8,137 | 0.001 | 0.213 | -0.752 | -0.139 | -0.001 | 0.139 | 0.752 |
| Age Difference | 8,137 | 0.001 | 1.573 | -5 | -1.100 | 0 | 1.10 | 5 |
| Occupancy Rate | 8,137 | 0.713 | 0.227 | 0 | 0.558 | 0.763 | 0.911 | 1 |
| Average Attendance(1000’s) | 8,137 | 24.18 | 18.03 | 0 | 10.51 | 19.23 | 35.19 | 81.17 |
| Foul Difference | 8,137 | 0.108 | 5.243 | -18 | -3 | 0 | 4 | 24 |
| Yellow Card Difference | 8,137 | 0.185 | 1.747 | -7 | -1 | 0 | 1 | 7 |
| Red Card Difference | 8,137 | 0.026 | 0.456 | -3 | 0 | 0 | 0 | 3 |
| Rating Difference | 8,137 | 0.101 | 15.63 | -58.31 | -9.64 | 0.24 | 9.80 | 62.27 |
| Importance Difference | 8,054 | 1.095 | 32.10 | -100 | -16.78 | 0.10 | 19.48 | 100 |
| Home goals | 8,137 | 1.542 | 1.303 | 0 | 1 | 1 | 2 | 10 |
| Away goals | 8,137 | 1.250 | 1.179 | 0 | 0 | 1 | 2 | 13 |
| Home shots | 8,137 | 12.946 | 5.165 | 0 | 9 | 12 | 16 | 38 |
| Away shots | 8,137 | 10.835 | 4.668 | 0 | 7 | 10 | 14 | 45 |
| Home shots on target | 8,137 | 4.899 | 2.640 | 0 | 3 | 5 | 6 | 18 |
| Away shots on target | 8,137 | 4.097 | 2.398 | 0 | 2 | 4 | 6 | 23 |
| Home fouls | 8,137 | 12.970 | 4.184 | 0 | 10 | 13 | 16 | 31 |
| Away fouls | 8,137 | 13.078 | 4.277 | 0 | 10 | 13 | 16 | 34 |
| Home corners | 8,137 | 5.405 | 2.972 | 0 | 3 | 5 | 7 | 23 |
| Away corners | 8,137 | 4.526 | 2.650 | 0 | 3 | 4 | 6 | 17 |
| Home yellow cards | 8,137 | 1.974 | 1.393 | 0 | 1 | 2 | 3 | 8 |
| Away yellow cards | 8,137 | 2.159 | 1.414 | 0 | 1 | 2 | 3 | 8 |
| Home red cards | 8,137 | 0.094 | 0.314 | 0 | 0 | 0 | 0 | 3 |
| Away red cards | 8,137 | 0.120 | 0.351 | 0 | 0 | 0 | 0 | 3 |
| Average age home team | 8,137 | 26.244 | 1.457 | 22.5 | 25.3 | 26.3 | 27.2 | 30.4 |
| Share of foreigners home team | 8,137 | 0.570 | 0.175 | 0.010 | 0.452 | 0.595 | 0.701 | 0.940 |
| Average age away team | 8,137 | 26.244 | 1.458 | 22.5 | 25.300 | 26.3 | 27.2 | 30.4 |
| Share of foreigners away team | 8,137 | 0.570 | 0.175 | 0.010 | 0.452 | 0.595 | 0.702 | 0.940 |
| Spi rating home team | 8,137 | 60.012 | 16.006 | 21.450 | 47.000 | 61.270 | 71.430 | 95.750 |
| Spi rating away team | 8,137 | 59.912 | 15.994 | 21.150 | 46.780 | 61.110 | 71.250 | 95.470 |
| Importance rating home team | 8,054 | 33.614 | 24.951 | 0.000 | 13.325 | 30.800 | 48.900 | 100.000 |
| Importance rating away team | 8,054 | 32.519 | 24.641 | 0.000 | 12.600 | 29.200 | 48.000 | 100.000 |
| Expected goals home team | 5,999 | 1.528 | 0.884 | 0.000 | 0.875 | 1.380 | 2.020 | 7.070 |
| Expected goals away team | 5,999 | 1.272 | 0.803 | 0.000 | 0.670 | 1.130 | 1.720 | 8.270 |
| Expected goals difference | 5,999 | 0.256 | 1.299 | -8.040 | -0.560 | 0.240 | 1.060 | 6.790 |
| Home win percentage | 8,137 | 0.436 | 0.496 | 0 | 0 | 0 | 1 | 1 |
| Away win percentage | 8,137 | 0.314 | 0.464 | 0 | 0 | 0 | 1 | 1 |
| Draw percentage | 8,137 | 0.251 | 0.433 | 0 | 0 | 0 | 1 | 1 |
| Points home team | 8,137 | 1.557 | 1.321 | 0 | 0 | 1 | 3 | 3 |
| Points away team | 8,137 | 1.192 | 1.286 | 0 | 0 | 1 | 3 | 3 |
| Percentage of points home | 8,137 | 0.561 | 0.429 | 0 | 0 | 0.5 | 1 | 1 |
| Percentage of points away | 8,137 | 0.439 | 0.429 | 0 | 0 | 0.5 | 1 | 1 |
| Corner difference | 8,137 | 0.879 | 4.465 | -16 | -2 | 1 | 4 | 20 |
| Shots difference | 8,137 | 8.420 | 6.494 | -14 | 4 | 8 | 12 | 36 |
| Shots on target difference | 8,137 | 0.802 | 3.793 | -23 | -2 | 1 | 3 | 17 |
| Percentage of points difference | 8,137 | 0.122 | 0.857 | -1 | -1 | 0 | 1 | 1 |

**Appendix 1b**

Appendix 1b shows the summary statistics for all the variables for all matches in our total sample played before the start of the covid-19 pandemic. 5,141 out of 8,137 matches in our sample were played with supporters. The is a clear gap between home and away teams in terms of points collected, goals scored, expected goals and shots.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Summary Statistics pre-covid matches** | | | | | | | | |
|  | **N** | **Mean** | **SD** | **Min** | **Pctl(25)** | **Median** | **Pctl(75)** | **Max** |
| Home goals | 5,141 | 1.584 | 1.321 | 0 | 1 | 1 | 2 | 10 |
| Away goals | 5,141 | 1.222 | 1.176 | 0 | 0 | 1 | 2 | 9 |
| Home shots | 5,141 | 13.344 | 5.217 | 0 | 10 | 13 | 16 | 38 |
| Away shots | 5,141 | 10.783 | 4.623 | 0 | 7 | 10 | 14 | 34 |
| Home shots on target | 5,141 | 5.058 | 2.666 | 0 | 3 | 5 | 7 | 18 |
| Away shots on target | 5,141 | 4.096 | 2.386 | 0 | 2 | 4 | 6 | 22 |
| Home fouls | 5,141 | 12.873 | 4.237 | 0 | 10 | 13 | 16 | 31 |
| Away fouls | 5,141 | 13.166 | 4.349 | 1 | 10 | 13 | 16 | 34 |
| Home corners | 5,141 | 5.605 | 3.012 | 0 | 3 | 5 | 7 | 23 |
| Away corners | 5,141 | 4.533 | 2.627 | 0 | 3 | 4 | 6 | 16 |
| Home yellow cards | 5,141 | 1.961 | 1.401 | 0 | 1 | 2 | 3 | 8 |
| Away yellow cards | 5,141 | 2.262 | 1.432 | 0 | 1 | 2 | 3 | 8 |
| Home red cards | 5,141 | 0.093 | 0.313 | 0 | 0 | 0 | 0 | 3 |
| Away red cards | 5,141 | 0.127 | 0.357 | 0 | 0 | 0 | 0 | 2 |
| Average age home | 5,141 | 26.201 | 1.453 | 22.5 | 25.2 | 26.3 | 27.2 | 30.4 |
| Foreigners share home | 5,141 | 0.567 | 0.176 | 0.010 | 0.438 | 0.593 | 0.701 | 0.940 |
| Average attendance | 5,141 | 24.416 | 17.852 | 1.969 | 10.897 | 19.369 | 35.191 | 81.171 |
| Occupancy rate | 5,141 | 0.724 | 0.219 | 0.077 | 0.570 | 0.783 | 0.916 | 1 |
| Average age away | 5,141 | 26.202 | 1.452 | 22.5 | 25.2 | 26.3 | 27.2 | 30.4 |
| Foreigners share away | 5,141 | 0.566 | 0.177 | 0.010 | 0.438 | 0.593 | 0.702 | 0.940 |
| Spi rating home | 5,141 | 60.391 | 15.698 | 21.450 | 48.090 | 61.620 | 71.370 | 95.750 |
| Spi rating away | 5,141 | 60.272 | 15.666 | 21.150 | 47.670 | 61.510 | 71.150 | 95.470 |
| Importance rating home | 5,100 | 33.891 | 24.583 | 0.000 | 14.775 | 31.200 | 48.700 | 100.000 |
| Importance rating away | 5,100 | 32.883 | 24.330 | 0.000 | 13.700 | 29.800 | 47.700 | 100.000 |
| Expected goals home | 3,665 | 1.569 | 0.877 | 0.000 | 0.930 | 1.420 | 2.060 | 7.070 |
| Expected goals away | 3,665 | 1.254 | 0.789 | 0.000 | 0.660 | 1.100 | 1.690 | 5.900 |
| Yellow card difference | 5,141 | 0.301 | 1.755 | -7 | -1 | 0 | 1 | 7 |
| Rating difference | 5,141 | 0.120 | 15.664 | -58.310 | -9.280 | 0.170 | 9.670 | 62.270 |
| Expected goals difference | 3,665 | 0.314 | 1.273 | -5.340 | -0.480 | 0.300 | 1.080 | 6.790 |
| Age difference | 5,141 | -0.001 | 1.574 | -5.000 | -1.100 | 0.000 | 1.100 | 5.000 |
| Red card difference | 5,141 | 0.034 | 0.461 | -3 | 0 | 0 | 0 | 2 |
| Importance difference | 5,100 | 1.008 | 31.979 | -100.000 | -16.900 | 0.100 | 19.025 | 100.000 |
| Percentage home wins | 5,141 | 0.454 | 0.498 | 0 | 0 | 0 | 1 | 1 |
| Percentage away wins | 5,141 | 0.299 | 0.458 | 0 | 0 | 0 | 1 | 1 |
| Percentage draws | 5,141 | 0.247 | 0.431 | 0 | 0 | 0 | 0 | 1 |
| Home points | 5,141 | 1.609 | 1.321 | 0 | 0 | 1 | 3 | 3 |
| Away points | 5,141 | 1.145 | 1.277 | 0 | 0 | 1 | 3 | 3 |
| Goal difference | 5,141 | 0.362 | 1.858 | -9 | -1 | 0 | 1 | 10 |
| Points difference | 5,141 | 0.464 | 2.562 | -3 | -3 | 0 | 3 | 3 |
| Foul difference | 5,141 | 0.293 | 5.255 | -18 | -3 | 0 | 4 | 24 |
| Foreigners share difference | 5,141 | 0.0002 | 0.214 | -0.720 | -0.139 | 0.000 | 0.144 | 0.752 |
| Percentage points home | 5,141 | 0.577 | 0.427 | 0 | 0 | 0.5 | 1 | 1 |
| Percentage points away | 5,141 | 0.423 | 0.427 | 0 | 0 | 0.5 | 1 | 1 |
| Corner difference | 5,141 | 1.072 | 4.481 | -14 | -2 | 1 | 4 | 20 |
| Shots difference | 5,141 | 8.812 | 6.531 | -14 | 4 | 8 | 13 | 34 |
| Shots on target difference | 5,141 | 0.961 | 3.804 | -20 | -1 | 1 | 3 | 17 |
| Percentage points difference | 5,141 | 0.155 | 0.854 | -1 | -1 | 0 | 1 | 1 |

**Appendix 1c**

Appendix 1c shows the summary statistics for all matches played after the start of the covid-19 pandemic. For the matches without fans, Just like for matches played with fans, home teams score more goals, collect more points and have more shots than away teams. However, the gap has clearly become more narrow following the exclusion of supporters.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Summary Statistics covid matches** | | | | | | | | |
|  | **N** | **Mean** | **SD** | **Min** | **Pctl(25)** | **Median** | **Pctl(75)** | **Max** |
| Home goals | 2,996 | 1.469 | 1.269 | 0 | 1 | 1 | 2 | 9 |
| Away goals | 2,996 | 1.299 | 1.185 | 0 | 0 | 1 | 2 | 13 |
| Home shots | 2,996 | 12.262 | 5.002 | 1 | 9 | 12 | 15 | 38 |
| Away shots | 2,996 | 10.924 | 4.743 | 0 | 8 | 10 | 14 | 45 |
| Home shots on target | 2,996 | 4.627 | 2.572 | 0 | 3 | 4 | 6 | 16 |
| Away shots on target | 2,996 | 4.098 | 2.419 | 0 | 2 | 4 | 5.2 | 23 |
| Home fouls | 2,996 | 13.137 | 4.087 | 0 | 10 | 13 | 16 | 31 |
| Away fouls | 2,996 | 12.928 | 4.148 | 0 | 10 | 13 | 16 | 30 |
| Home corners | 2,996 | 5.062 | 2.870 | 0 | 3 | 5 | 7 | 19 |
| Away corners | 2,996 | 4.514 | 2.690 | 0 | 3 | 4 | 6 | 17 |
| Home yellow cards | 2,996 | 1.996 | 1.380 | 0 | 1 | 2 | 3 | 8 |
| Away yellow cards | 2,996 | 1.983 | 1.363 | 0 | 1 | 2 | 3 | 8 |
| Home red cards | 2,996 | 0.097 | 0.316 | 0 | 0 | 0 | 0 | 3 |
| Away red cards | 2,996 | 0.108 | 0.339 | 0 | 0 | 0 | 0 | 3 |
| Average age home | 2,996 | 26.319 | 1.460 | 22.700 | 25.400 | 26.300 | 27.200 | 30.300 |
| Foreigners share home | 2,996 | 0.576 | 0.171 | 0.010 | 0.465 | 0.606 | 0.702 | 0.902 |
| Average attendance | 2,996 | 23.777 | 18.313 | 0.000 | 9.825 | 18.301 | 35.191 | 81.171 |
| Occupancy rate | 2,996 | 0.694 | 0.239 | 0.000 | 0.536 | 0.732 | 0.909 | 1.000 |
| Average age away | 2,996 | 26.316 | 1.465 | 22.700 | 25.400 | 26.300 | 27.200 | 30.300 |
| Foreigners share away | 2,996 | 0.577 | 0.171 | 0.010 | 0.465 | 0.605 | 0.706 | 0.902 |
| Spi rating home | 2,996 | 59.362 | 16.504 | 24.210 | 45.130 | 60.425 | 71.692 | 95.610 |
| Spi rating away | 2,996 | 59.294 | 16.526 | 24.240 | 44.870 | 60.405 | 71.630 | 95.320 |
| Importance rating home | 2,954 | 33.136 | 25.572 | 0.000 | 11.100 | 30.200 | 49.600 | 100.000 |
| Importance rating away | 2,954 | 31.892 | 25.161 | 0.000 | 10.600 | 28.150 | 48.400 | 100.000 |
| Expected goals home | 2,334 | 1.464 | 0.891 | 0.000 | 0.800 | 1.300 | 1.960 | 6.190 |
| Expected goals away | 2,334 | 1.300 | 0.824 | 0.000 | 0.680 | 1.170 | 1.760 | 8.270 |
| Yellow card difference | 2,996 | -0.013 | 1.714 | -6 | -1 | 0 | 1 | 6 |
| Rating difference | 2,996 | 0.069 | 15.561 | -53.260 | -10.233 | 0.330 | 10.060 | 57.420 |
| Expected goals difference | 2,334 | 0.164 | 1.334 | -8.040 | -0.710 | 0.125 | 0.990 | 5.570 |
| Age difference | 2,996 | 0.002 | 1.573 | -4.800 | -1.100 | 0.000 | 1.000 | 4.500 |
| Red card difference | 2,996 | 0.011 | 0.447 | -3 | 0 | 0 | 0 | 3 |
| Importance difference | 2,954 | 1.245 | 32.330 | -100.000 | -16.400 | 0.100 | 20.000 | 100.000 |
| Percentage home wins | 2,996 | 0.404 | 0.491 | 0 | 0 | 0 | 1 | 1 |
| Percentage away wins | 2,996 | 0.338 | 0.473 | 0 | 0 | 0 | 1 | 1 |
| Percentage draws | 2,996 | 0.258 | 0.437 | 0 | 0 | 0 | 1 | 1 |
| Home points | 2,996 | 1.469 | 1.317 | 0 | 0 | 1 | 3 | 3 |
| Away points | 2,996 | 1.273 | 1.298 | 0 | 0 | 1 | 3 | 3 |
| Goal difference | 2,996 | 0.170 | 1.821 | -13 | -1 | 0 | 1 | 9 |
| Points difference | 2,996 | 0.196 | 2.578 | -3 | -3 | 0 | 3 | 3 |
| Foul difference | 2,996 | -0.210 | 5.207 | -18 | -4 | 0 | 3 | 19 |
| Foreigners share difference | 2,996 | -0.001 | 0.210 | -0.752 | -0.139 | -0.001 | 0.135 | 0.717 |
| Percentage points home | 2,996 | 0.533 | 0.430 | 0.000 | 0.000 | 0.500 | 1.000 | 1.000 |
| Percentage points away | 2,996 | 0.467 | 0.430 | 0.000 | 0.000 | 0.500 | 1.000 | 1.000 |
| Corner difference | 2,996 | 0.547 | 4.419 | -16 | -2 | 1 | 3 | 17 |
| Shots difference | 2,996 | 7.748 | 6.375 | -13 | 3 | 8 | 12 | 36 |
| Shots on target difference | 2,996 | 0.529 | 3.760 | -23 | -2 | 1 | 3 | 15 |
| Percentage points difference | 2,996 | 0.065 | 0.859 | -1 | -1 | 0 | 1 | 1 |

**Appendix 2a:** Normality test data

Appendix 2a displays the univariate normality tests for all of our variables. Since univariate normality is not guaranteed, we used a Mann-Whitney test rather than t-tests to tests statistical differences between matches played before and after the start of the covid-19 pandemic

|  |  |  |  |
| --- | --- | --- | --- |
| **Normality test** | | | |
|  | **Test Statistic** | **P value** | **Alternative Hypothesis** |
| Percentage points Home | 794.3 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Points Home | 886.1 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Home Goals | 306.3 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Percentage home Wins | 1494 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Expected goals Home | 71.81 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Home Shots | 52.31 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Home Shots on Target | 97.15 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Home Fouls | 31.5 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Home Red | 2660 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Home Yellow | 217.6 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Percentage points Away | 794.3 | *p*  <.001 \* \* \* | *Variable not normally distributed* |
| Points Away | 914.8 | *p*  <.001 \* \* \* | *Variable not normally distributed* |
| Away Goals | 389.3 | *p*  <.001 \* \* \* | *Variable not normally distributed* |
| Percentage Away wins | 1734 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Expected goals Away | 85.07 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Away Shots | 57.83 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Away Shots on Target | 116.9 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Away Yellow | 200.9 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Away Red | 2514 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Away Fouls | 33.29 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Goal difference | 110.8 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Foul difference | 12.45 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Red card difference | 1751 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Yellow card difference | 111.7 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Points difference | 794.3 | *p* <.001 \* \* \* | *Variable not normally distributed* |
| Expected goals difference | 5.64 | *p* <.001 \* \* \* | *Variable not normally distributed* |

Note: test statistic and p-values are calculated using the Anderson darling test for normality, Significance levels: *p* < .05\* , *p* < .01 \*\*, *p < .001 \*\*\**

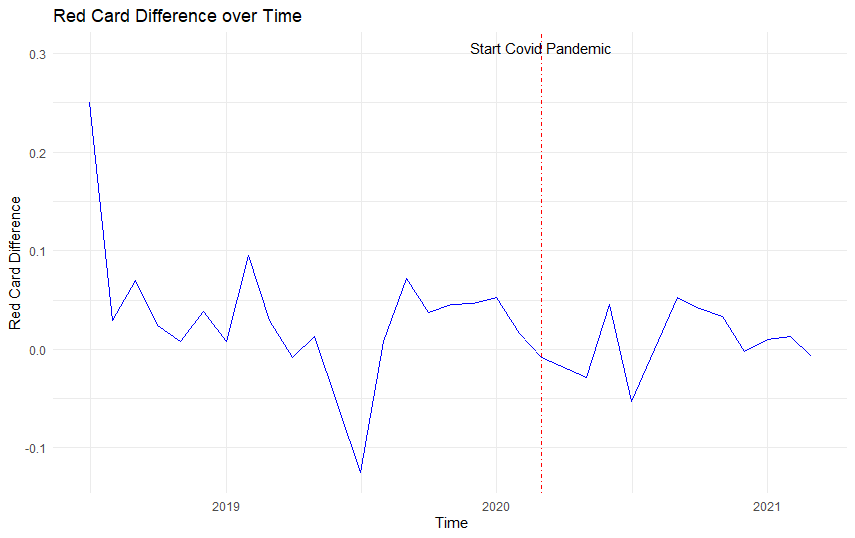
**Appendix 2b:** Significance tests pre and post covid:

Appendix 2b contains the test result for the Mann-Whitney test for all of the variables to test significant differences for these metrics for matches played before and after the start of the covid-19 pandemic. Virtually all metrics have changed significantly following the covid-19 pandemic. Only red and yellow cards for the home team and shots on target and fouls for the away team have not changed significantly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Significance test** | | | | |
|  | **Test Statistic** | **Tie factor** | **P value** | **Alternative hypothesis** |
| Percentage points Home | .5282 | .8707 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Points Home | .5282 | .8707 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Home Goals | .5244 | .9367 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Percentage home Wins | 74.2521 | - | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Expected goals Home | .5406 | 1 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Home Shots | .5605 | .9962 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Home Shots on Target | .5462 | .9854 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Home Fouls | .4802 | .9946 | *p* = .0028 \* \* | *Two distributions are not equal* |
| Home Red | .4977 | .2412 | *p* = .4763 | *Two distributions are not equal* |
| Home Yellow | .4915 | .9508 | *p* = .1898 | *Two distributions are not equal* |
| Percentage points Away | .4718 | .8707 | *p*  < .001 \* \* \* | *Two distributions are not equal* |
| Points Away | .4718 | .8707 | *p*  < .001 \* \* \* | *Two distributions are not equal* |
| Away Goals | .4791 | .9191 | *p*  < .001 \* \* \* | *Two distributions are not equal* |
| Percentage Away wins | 74.2521 | - | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Expected goals Away | .4846 | 1 | *p* = .04 \* | *Two distributions are not equal* |
| Away Shots | .4918 | .9954 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Away Shots on Target | .5016 | .9819 | *p* = .8079 | *Two distributions are not equal* |
| Away Yellow | .5542 | .9534 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Away Red | .5098 | .298 | *p* = ..006 \* \* | *Two distributions are not equal* |
| Away Fouls | .5128 | .9948 | *p* = .053 | *Two distributions are not equal* |
| Goal difference | .53 | .9683 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Foul difference | .5258 | .9966 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Red card difference | .5113 | .4351 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Yellow card difference | .5518 | .9679 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Points difference | .5282 | .8707 | *p* < .001 \* \* \* | *Two distributions are not equal* |
| Expected Goals Difference | .5364 | 1 | *p* < .001 \* \* \* | *Two distributions are not equal* |

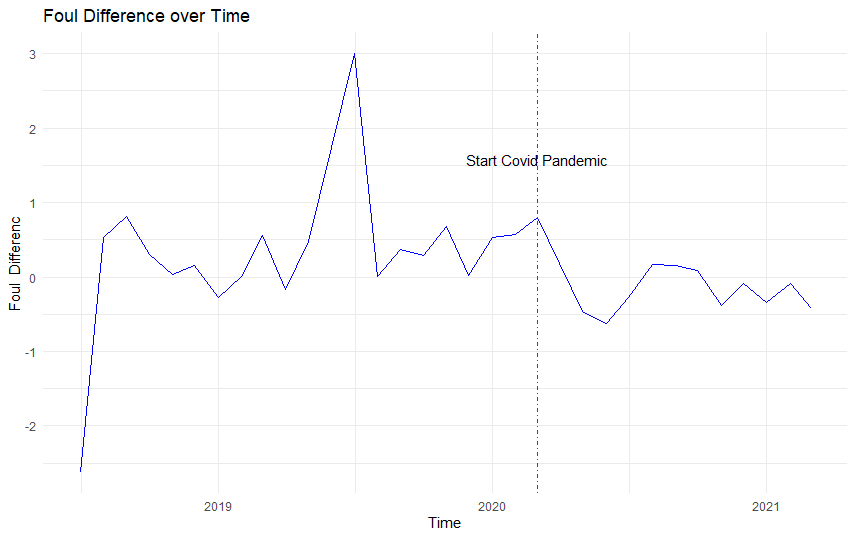
Note: test statistics and p value are calculated with Mann-Whitney test. 1: chi square proportion test, Significance levels: *p* < .05\* , *p* < .01 \*\*, *p< .001 \*\*\**

**Appendix 3a:** This appendix shows the variation in the red card difference between away and home teams over time. There is no clear difference between before and after the start of the covid-19 pandemic. This is perhaps due to the fact that the number of red cards is low.



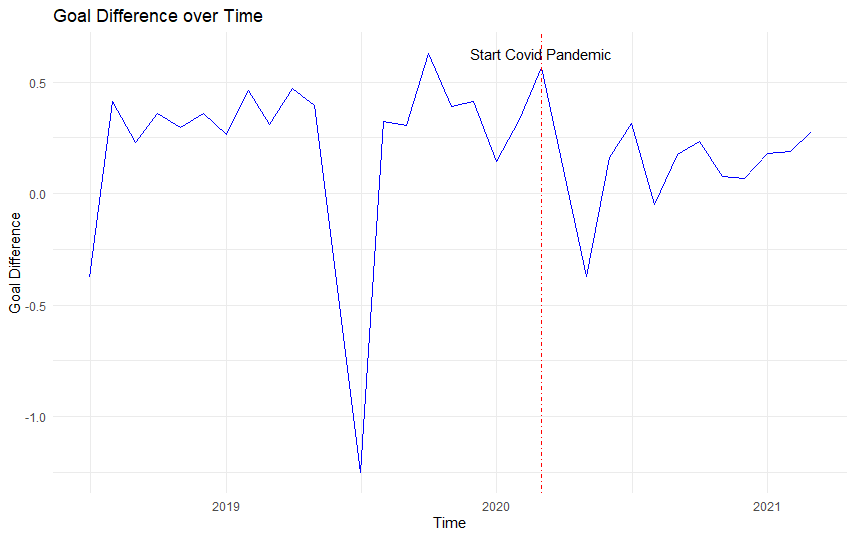
**Appendix 3b:**

This graph shows the evolution of the foul difference between home and away teams over the past 3 years. We see quite a sizeable drop after the starts of the covid pandemic where the foul difference dipped into negative territory after having been positive for the marjority of the 2 years before the covid-19 pandemic.



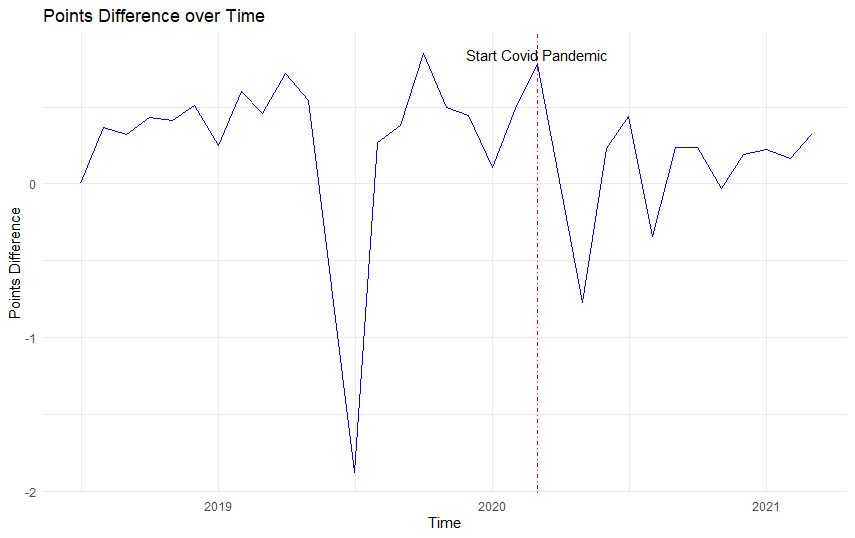
**Appendix 3c:**

Appendix 3c shows the goal difference over time. We see a very sharp reduction in the goal difference between home and away teams right after the start of the covid-19 pandemic. However, afterwards the goal difference gap increases again but still remains a bit lower than before.



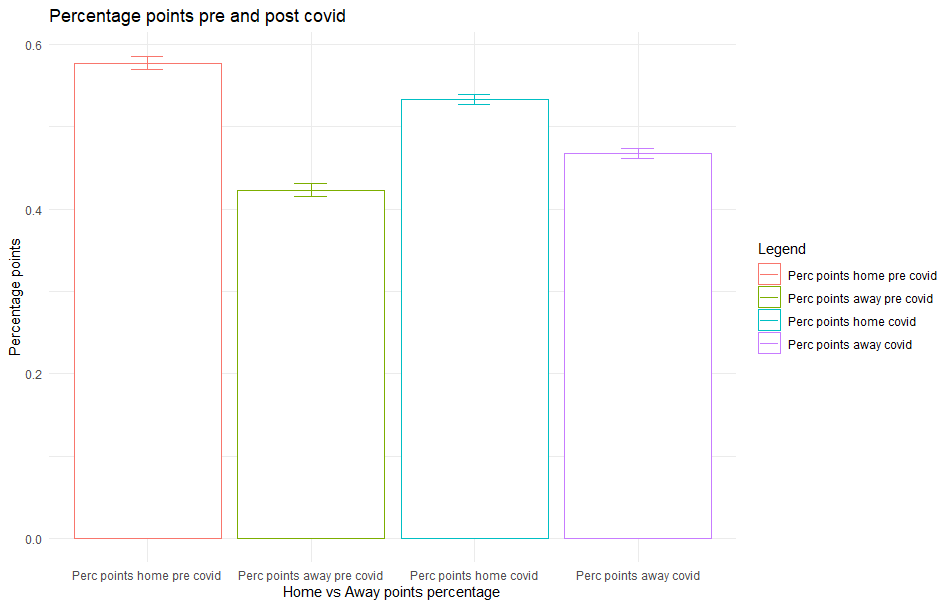
**Appendix 3d:**

Appendix 3d shows the points difference across the past 3 years. Similar to the previous graphs, we see a sharp drop right after the start of the covid-19 pandemic. Subsequently, the difference in points tends to fluctuate quite much in the months after.



**Appendix 3e: Percentage points pre and post covid**

This graph shows that there is still a home advantage in the case of “ghost games”, since the error bars of the percentage of poitns for home and away teams do not touch in the right half. The gap between home and away teams however has significantly decreased.



**Appendix 4a: Assumptions test indirect path**

Value p-value Decision

Global Stat 7.646553 0.10542 Assumptions acceptable.

Skewness 2.452351 0.11735 Assumptions acceptable.

Kurtosis 4.250997 0.03923 Assumptions NOT satisfied!

Link Function 0.940023 0.33227 Assumptions acceptable.

Heteroscedasticity 0.003182 0.95502 Assumptions acceptable

**Appendix 4b: assumption test direct path Points Difference**

Value p-value Decision

Global Stat 442.7895 0.000e+00 Assumptions NOT satisfied!

Skewness 43.0247 5.405e-11 Assumptions NOT satisfied!

Kurtosis 396.1679 0.000e+00 Assumptions NOT satisfied!

Link Function 3.0012 8.320e-02 Assumptions acceptable.

Heteroscedasticity 0.5957 4.402e-01 Assumptions acceptable.

**Appendix 4c: assumption test direct path Goal Difference**

Value p-value Decision

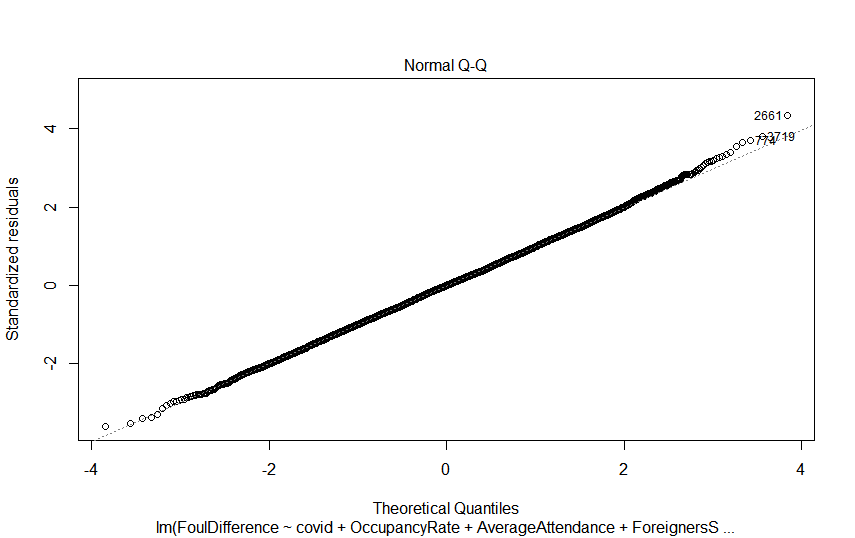
Global Stat 253.05579 0.00000 Assumptions NOT satisfied!

Skewness 0.75124 0.38608 Assumptions acceptable.

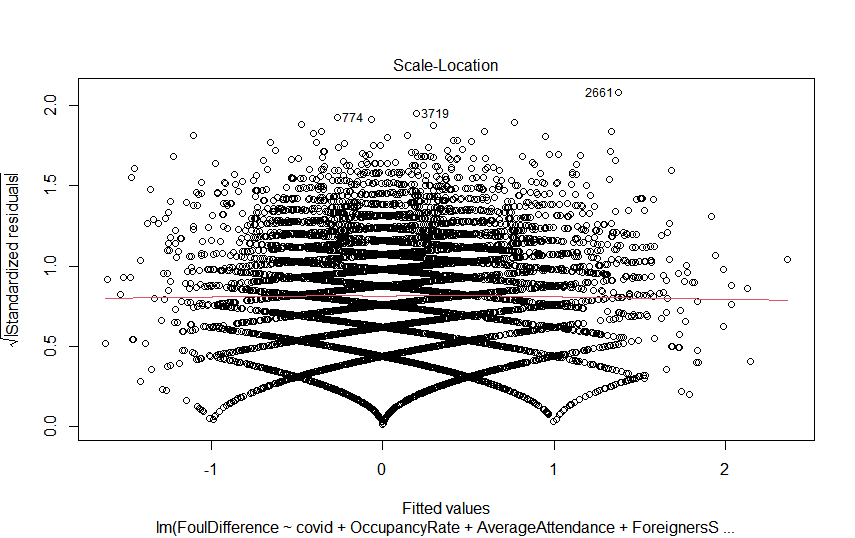
Kurtosis 248.47293 0.00000 Assumptions NOT satisfied!

Link Function 3.82617 0.05046 Assumptions acceptable.

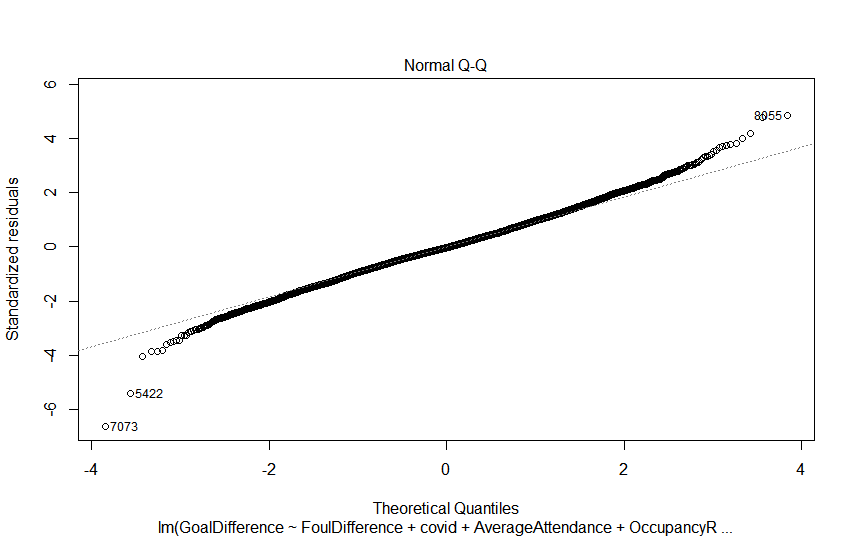
Heteroscedasticity 0.00545 0.94115 Assumptions acceptable.

**Appendix 5a: Normality plot indirect path** 

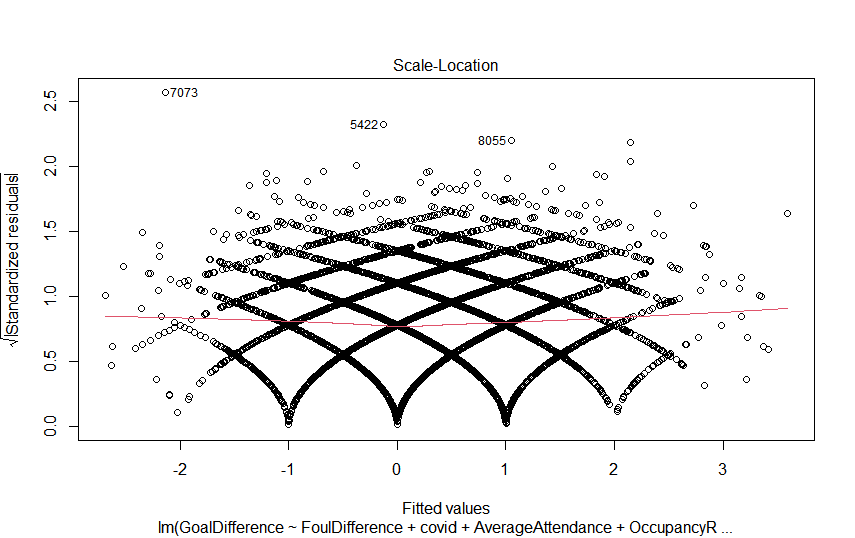
**Appendix 5b: Homoscedasticity indirect path**



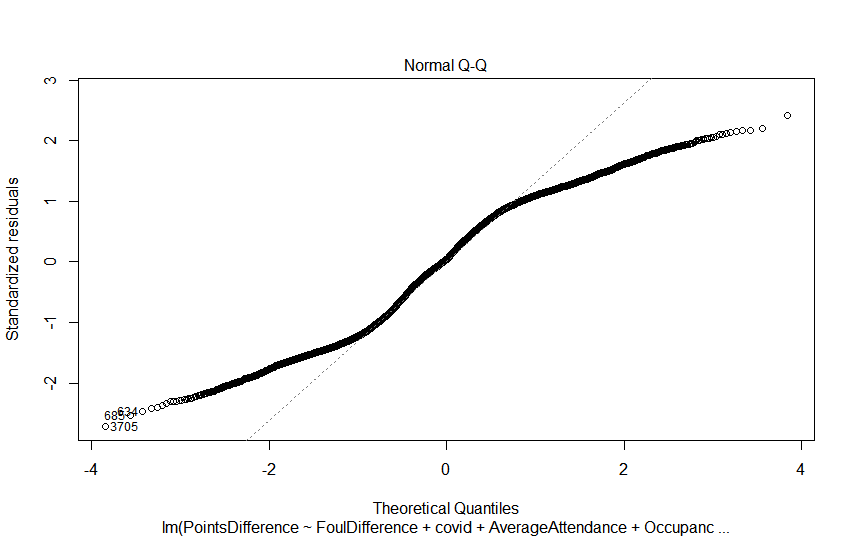
**Appendix 5c: normality direct path Goal Difference:**



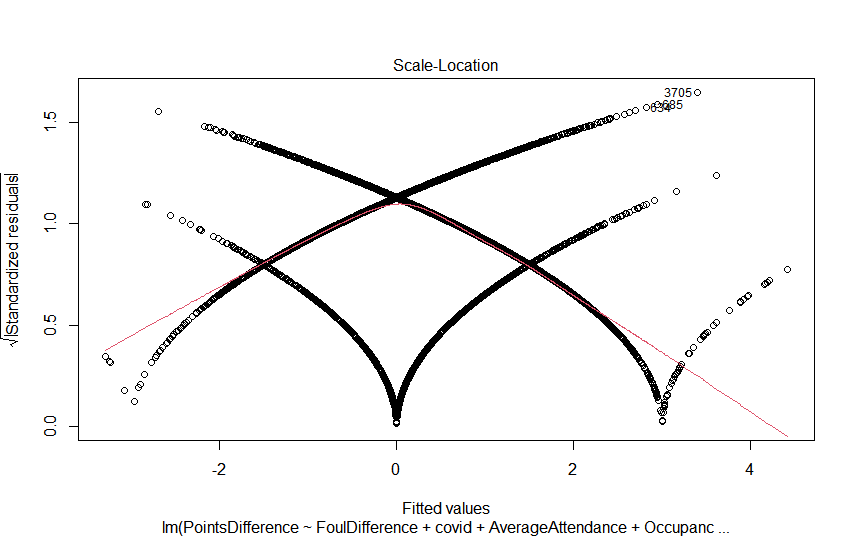
**Appendix 5d: Homoscedasticity direct path Goal difference**



**Appendix 6a: Normality direct path Points difference**

****

**Appendix 6b: Homoscedasticity direct path Points difference**

****

**Appendix 7: Correlation table**



Appendix 8a:**VIF table goal difference** 

**Appendix 8b:** **VIF table points difference**



**Appendix 9a**



A**ppendix 9b**



**Appendix 9c**



Yellow Card Goal



Yellow Card Point



Point with red card



1. <https://projects.fivethirtyeight.com/soccer-predictions/> [↑](#footnote-ref-0)
2. https://www.football-coefficient.eu/money/ [↑](#footnote-ref-1)
3. https://www.skysports.com/football/news/11661/11358620/the-cost-of-relegation-what-is-the-financial-impact-of-dropping-out-of-the-premier-league [↑](#footnote-ref-2)