ORIGINAL ARTICLE

The Effect of Crowd Support on Home-Field Advantage: Evidence from European Football

Tuğbay Inan*
Faculty of Sport Sciences, Dokuz Eylul University, Izmir, Turkey.

Submitted 07 October 2019; Accepted in final form 21 January 2020.

ABSTRACT

Background. Home field advantage exists in all professional games including professional football. Studies have revealed persistent home advantage across countries and divisions. Although crowd effects have regularly been cited as one of the potential causes of home advantage, the effect of crowd size and crowd density on home advantage has not been demonstrated in any football league. Objectives. The aim of this study was to demonstrate this effect using a new theoretical framework and binary logistic regression. Methods. Data collected for four seasons of five major European Leagues was processed. For this reason, 4 seasons and n = 8530 matches were analyzed starting from 2015 to 2019. Results. The model results suggest the presence of home-field advantage for all major European Leagues with varying degrees and varying shares attributable to the crowd support and crowd density. Conclusion. In the present study, it was found that crowd support and crowd density are of great importance for home-field advantage.

KEYWORDS: Home Advantage, European Football, Crowd Effects, Crowd Size, Crowd Density, Soccer

INTRODUCTION

Conventional wisdom of players, coaches, commentators and fans declares that in professional team sports, home side has an advantage. Regardless of country and division of league, including international games, home field advantage can be seen. Major professional team sports in England and North America have all been studied with this perspective and the results have confirmed this finding (1). Moreover, a paper by Pollard and Gomez studied 157 national soccer leagues and found that the advantage persists in all continents (2). Home advantage is valid for the long term as well. Inan applied Pollard’s methodology to 30 seasons of the Turkish Super League and found that home advantage was present in all seasons (3).

The exact causes of home advantage are still an issue of investigation and an important topic of ongoing research in sports economics. The factors leading to the home field advantage and how these factors affect the outcome of a game have yet not been well understood.

Dowie was the first to apply the concept to soccer and came forward with three possible causes. He mentioned fatigue, familiarity and fans, however he did not reach clear-cut conclusions (3).

The 1986 study of Pollard set the most commonly used system to analyze home advantage. Home field advantage has since been expressed by the percentage of points gained or goals scored at home with respect to total points gained or total goals scored. Pollard summarized the likely causes as local crowd support, travel fatigue, familiarity with local conditions, referee bias, special tactics and psychological factor (1).
Pollard later added territoriality and rule factors to the likely causes (4). Other causes may also come into play. Marek & Vavra used a different approach by using goals, instead of points gained, and compared home field advantage in different leagues (5). Different studies have tried to dissect various causes and found all to be effective to different degrees. One notable exception is the rule factor. Starting with 1980s, the system where 2 points was awarded to the winner and 1 point to both teams in the case of draw was changed throughout the world and 3 points started to be awarded to winners. Moreover all leagues use the same rules. Therefore rule factors were meaningful before but have now become obsolete (2).

Travel effects performance of the teams negatively. Pollard, Silva, & Medeiros found that the length of the travel decreases the performance in soccer games in Brazil (6). In international games, jet lag may take a toll. Another possible factor is the impact of travel on sleep and food. The mode of travel, ie. Plane, train or bus and the quality of accommodation may all have negative effects. Many people would prefer to sleep at home in his own bed rather than in a hotel. These are all questions yet to be analyzed in depth. Just the opposite, in local derbies, home advantage decreases. Some of this is probably because of lack of travel fatigue, rest may come from more balanced support. Seckin and Pollard found that home advantage was significantly lower in Istanbul derbies, where teams based in Istanbul played against each other (7). London Clarke & Norman, Lisbon, Paris, Madrid and Milan Pollard & Gomez, have all been shown as examples Seckin and Pollard also tried to assess home advantage for several teams in remote locations by considering travel fatigue and territoriality as potential factors here (1-5, 8-20). The study found that especially in Van, Rize and Diyarbakir, home advantage was higher than average, with Van having the most advantage (7). As athletes are thought to lose performance at altitudes higher than 1,500 meters, it is probably not coincidental that Van has an altitude of over 1,700 meters, although other factors may also have been influential (21).

Familiarity with local conditions works for the home team. An interesting study has shown that home advantage decreases when a team moves to a new stadium (16). Territoriality is a rather subjective cause as well. We believe it can be analyzed in the same manner with special tactics and physiological factors. Home advantage may to some extent be a self-perpetuating phenomenon. Because teams expect to do better in home turf, or because they believe they should play well in their home city, in front of their own fans, they do better. They adopt different tactics or act more aggressively hoping that the referee, under crowd pressure, will be more lenient. This feeling of having to do better may be heightened because of political issues, usually analyzed under territoriality.

Referee bias in favor of the home team has been well established (21). Boyko et al., examined 5244 English Premier League (EPL) match results including 50 referees and found that referees cause for some part of the observed home advantage in the EPL, though home bias differs between referees (22). Lovell, Newell, & Parker also showed referee bias does exist, even that may be the main reason for home advantage (17). Even though referee bias is deemed as a determinant for home-field advantage, yet that should be taken as a result at the intermediary stage in a sequential process which is triggered by the crowd. In this regard, the presence of crowd noise influenced referees’ decisions to favor the home team, observed by Nevill, Balmer, and Williams (18). Anders and Rotthoff showed that the home advantage was greater in teams with greater prospect of fan violence (23). However, the dynamics still need exploration. This may be a very human reaction to crowd support but corruption, gambling interests and fear of violence from home team supporters may all play a role. It is not easy if even possible to study these effects. However, crowd effect has been paid attention recently to investigate further.

Home teams show more performance, visiting teams less, referees may show bias towards home team, of course, yet that is stemming mainly from crowd support. Boudreaux, Sanders, & Walia examined the crowd effect upon home field advantage, and found that is quite meaningful for motivating home team to win and almost entire home advantage is attributable to the crowd effect (24). If we take a closer look at the crowd affect in home advantage, its size, intensity, density and proximity can be expected to affect both teams and the referees (21).

Intensity and Density: Greer analyzed the effect of “booing” on team performances in American College Basketball and reached the
conclusion that booing adversely affected the performance of the visiting team (13) the size of the crowd, which is the focus of this study, is an indicator of the booing potential. However, the effect on booing may not be similar in professional games, mainly because those teams contain relatively older and more experienced players compared to college teams. Hence they would be expected to be affected less by the booing. Moreover, the dynamics and venues of basketball are not the same as football. Therefore, caution is advised in making comparisons.

Size: It should be noted that magnitude of home advantage varies a lot among countries (21). Interestingly, past research has shown that home advantage does not vary much among divisions of the same league, although crowd sizes differ a lot (25). The crucial point here is all teams generally have smaller crowds in lower divisions and the size of stadiums can also be smaller. Although significant positive association between crowd size and home advantage was shown in European competition by Goumas (12). And no relationship between home advantage and crowd size has been demonstrated within a single league (21).

Proximity: Existence of running tracks in stadiums decrease proximity. Home advantage decreases in stadiums without running tracks, according to the study done by Armatas & Pollard for Greek football (26).

The aim of this study is to demonstrate this effect in major leagues in Europe by using a new framework and backward multiple linear regression, an econometric technique not previously used in this field, to the best of our knowledge. We will differentiate team capabilities that are always present regardless of the venue, and so we can find the effect of home advantage as shown in the Figure 1 below. This study focuses especially on crowd effects in home advantage by examining football matches from major leagues: Bundesliga (Germany), La Liga (Spain), Ligue 1 (France), Premier League (United Kingdom) and Serie A (Italy).

**MATERIALS AND METHOD**

A new theoretical approach will be presented and tested with empirical data in this study. Rather than evaluating home field advantage by the widely used percentage ratio of home points gained, we will prefer investigating the determinants of the win home points per game by each team in each match. In this examination we divide the determinants of home field average points in two categories: The factors crowd size, and crowd density that are directly related to playing at home game by each team in each match. An elaboration of this theoretical framework is presented in Figure 1.

Indeed, performance indicators at home games and other capability indicators are important indicators showing the capability of a team to affect the outcome of a match even without any crowd support. The crowd size and crowd density will also be present in the home games and contribute to the points gained at home. The more capable a team is, the more points it is expected to gain at home even in the lack of crowd support and any other related home field advantage. And we may decompose the magnitude of the actual effect of crowd support on the points gained at home.

In statistics, the logistic model (or logit model) is used to model the probability of a certain class or event existing such as pass/fail, win/lose. Each object being detected in the image would be assigned a probability between 0 and 1 one. Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. A very simple Machine Learning algorithm which will come to your rescue is logistic regression. In regression analysis, logistic regression is estimating the parameters of a logistic model (a form of binary regression). Mathematically, a binary logistic model has a dependent variable with two possible values, such as win/lose which is represented by an indicator variable, where the two values are labeled "0" and "1". The binary logistic regression model has two levels of the dependent variable.

To this end, we constructed the statistical model formulated below:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

The variables in this model are explained below:

- $Y_i$ = Dependent variable,
- $\beta_0$ = Y intercept
- $\beta_1$ = Slope coefficient independent variable
- $\varepsilon_i$ = Error term for observation i.

Due to data limitations, the home advantage was tested with the size of the stadium, the number of the spectators in the home field, and the intensity of spectators in the stadium. In this context, the effect of the number and the density of the spectators in the stadium on the points
scored by the home team only in its own field was examined. Such factors as the total value of the team's squad, previous achievements, brand value, in-game performance indicators as well as physical and psychological factors were not taken into consideration. However, we will test this theoretical framework by only performance indicators at away games as controlling for capability and crowd support as basic home field factor. We compiled seasonal average values from match level data of five major European Leagues (Bundesliga (Germany), La Liga (Spain), Ligue 1 (France) Premier League (United Kingdom), Serie A (Italy) and for four seasons (2015-2016, 2016-2017, 2017-2018 and 2018-2019).

In this model, coefficients of spectator numbers variables measure the effect of crowd support and crowd density on points gained at home team by each match (Table 1).

Theoretically, we expect both the constant term and the coefficient of normalized number of spectators and spectators density to be positive implying that an increase in crowd support is expected to increase the home points gained.

We applied the same procedure for all these seven divisions by running this model with binary logistic regression model of the SPSS (24.0) software and the results are presented in the next section. The data used in the experiments are publicly available. Match results, stadium capacity and number attendance were downloaded from www.transfermarkt.com.

RESULTS
Model results are shown in Table 2 below. The Home team gained point observed in between 2014-2019 seasons that means, after 1225 games played, was 70.45%, i.e. in 863 from those 1225 games in Bundesliga, the team playing at home obtained a win or a draw. After 1520 games played, was 71.65% in La Liga, after 1520 games played, was 71.51% in Ligue 1 after 1520 games played, was 69.67% in Premier League, after 1520 games played, was 69.07% in Serie a.
In each turn of the binary logistic regression method of SPSS software, the variable with the highest p-value is excluded from the base model. This process continues until all the variables remaining in the model are statistically significant on 10 percent level of significance. Models shown in Table 2 are finally reached models for each country where the coefficients of all variables except the constant term are statistically significant on 5 percent level of significance. Although the independent variable (crowd density) is not statistically significant for some countries, we keep it in the model throughout the process both to avoid biased predictions.

Interestingly, it was found out that stadium capacity and the number of spectators in the stadium are important for the model in all leagues, while the density of the spectators is only important for La Liga and Ligue 1. It seems that given the sample size, these two variables are critical to reflect the effect of capability of a team in lack of crowd support on the points gained at home. On the other hand, the predicted overall percentage of accuracy in the model is over 65% for all leagues. This result shows us that the independent variables are correctly classified for the model. Interestingly, the highest values obtained as 71.9% for La Liga and 71.7% for Ligue 1 show that the model is correctly categorized as seen in Table 2.

For all of the divisions, the presence of home field advantage is empirically validated. The signs of all coefficients are in line with our theoretical expectations. The share of average home points earned attributable to home field factors, however, varies among divisions with a maximum of 71.65 percent for La Liga and with a minimum of 69.07 percent for Serie A. Interestingly, in addition to the low level of home field effect for Serie A, the crowd support also does not have any statistically significant effect on average home points earned. Also for Bundesliga, crowd support does not have any statistically significant effect on home points earned, however, 70, 45 percent of home points earned are attributable to home field conditions. This implies that physical or psychological home field factors are more important for Bundesliga compared to the effect of crowd support.

For the other 3 divisions for which crowd support has statistically significant effects, the share of home field advantage attributable to crowd support varies between 69, 67 percent (Premier League) and 71, 51 percent (Ligue 1). La Liga appears to be the division for which crowd support has the highest effect on points earned at home (71, 65%). It can also be seen in Table 2 that the increase in the number of spectators has a positive effect on the home team to score point or points. As the number of spectators attending the matches of the home
team increased, it was determined that the home teams would score more points in the Premier League and Ligue 1 than the visiting teams. On the other hand, when the density of the spectators is examined, it is interesting to see that as the intensity increases in Ligue 1 and La Liga, the home teams leave the stadium with score or scores in hand.

**DISCUSSION**

The results of the model used in this study suggest that there is a statistically significant level of home field advantage for all major European leagues. However, the share of home field advantage attributable to crowd support differs dramatically among divisions and for Serie A and Bundesliga crowd support does not even have any statistically effect on points earned at home. This might be due to the fact that Serie A has the most capable teams of the Europe and the differences in capabilities are much more important than the effects of home field conditions including crowd support. Or it can also be explained by the fact that Serie A has the lowest density of spectators among all leagues, which may explain the low home advantage.

In his study investigating the games played in terms of home field advantage in 72 leagues around the world, Pollard (2006) found that among the countries in Europe, especially Bosnia (78.95%) and Albania (77.20%) had a much higher home field advantage than the average. This rate is 63.83% in Italy, 62.81% in Germany, 61.19% in England and 63.90% in Spain, the leading countries of European football (20). The home field advantage was found in Brazil as (64.45%), in Argentina as (57.95%), and the highest rate was found in Bolivia with 74.16%. In the study conducted by Garcia et al. (2013), the average home ownership advantage was 63.18% in Albania, 60.99% in Bulgaria, 61.29% in Croatia, 59.40% in Romania and 59.80% in Greece (10). In their study, Armatas and Pollard (2014) analyzed a total of 2160 matches and examined the home field advantage in the Greek Super League, determining that number of shots, corner kicks, steals and goals scored are higher in favour of the home teams than the visiting teams. In addition, according to data related to the teams’ home field advantages covering at least 4-seasons, they found that the point advantage scored by the home teams ranged from 60.14% to 74.54% (26). According to Goumas’s (2014) study, the advantage of being a home team in terms of scored points reached the peak with 65.5% in the 2011/2012 season, while the average of the 7 seasons examined was found to be 57.7%.

In 2015-2016 season (12), Leitte, in a study conducted in (2017), examined football leagues of ten different countries in order to determine the home field advantage rates. According to the survey carried out in this study examining a total of 3223 matches from 10 football leagues in Europe, the mean home field advantage was found as 61.2% in Spain, 61% in Turkey, 60.7% in Belgium, 59.3% in Italy, the Netherlands 57.5%, 57.4% in France, 56.4% in Portugal, 56.2% in Germany, and 55.3% in Russia (15). Gould et al. (1999) stated that players are motivated when they get the chance to show their talents and enthusiasm to their supporters (11). Backing this idea, Carmichael & Thomas (2005) says that football teams playing in their own stadiums are more effective in terms of shoots and attacks (27).

Many factors that could contribute to home advantage (HA) in football have been identified, and the impact of crowd support on both player performance and referee decision-making process has been examined (20). Although many studies have investigated the relationship between crowd support and HA in football, most have dwelled only on one aspect of crowd support which is the crowd size, and many have not compared the independent effects of different crowd factors. The aim of this study was to measure the extent to which the crowd effect (size and density) contributed to HA in football in the major European football leagues and to investigate to what extent crowd support contributed to home teams in term of scoring points. In football matches, a team must fulfill many conditions in order to win a match. A summary of the variables from which the home team might be positively, while the visiting team negatively affected is presented in Figure 1.

Crowd effect can be defined as noise or cheer, created in different ways by the spectators and the number and density of spectators in order to affect the performances of the players on the field. In particular, the presence of a large number of spectators in the stadiums is said to have a positive impact on the performance of the home team. Schwartz and Barsky (1977) state that the more the the number of spectators in a match, the higher the probability of success in favor of the home team (28). According to Pollard and Pollard...
(2005), it is stated that this advantage can reach up to 12% from time to time (25).

In many studies in the literature, it was stated that the number of spectators of the home team would exhibit a positive advantage in favor of the home team. Similar results were also found in Ligue 1 (France) and Premier League (England) in our study. Overall, most of the current studies suggest that collective support towards the home team is a factor that positively affects the performance of the home teams and also, the high number of spectators in the home matches may affect the visiting team psychologically and this may have a negative effect on the performances of the visiting team players. On the other hand, the pressure of the spectators in the stadiums may affect the decisions of the referees. In a study about ice hockey competitions, Agnew, & Carron (1994) examined the effect of the factors related to the presence of crowd support (the number of spectators, arena capacity compared to the spectator size). They mentioned that there was a 61.6% home advantage when draws were removed from the matches. On the other hand, they found that crowd density was significantly related to the home field advantage, and that the higher the density, the higher the home field advantage (29). In a study of Scottish football leagues, although the mean crowd size in the top two leagues was four times larger compared to the third league, in the top two leagues the same level of HA was observed (about 60%), while the figure was found to be (51%) in the third league attracting fewer spectators (19). Analyzing English football leagues, Pollard found that HA was higher in the first four leagues (about 60%) than the lower five leagues (55%) as a percentage of home team points scored (4). When past studies are examined, it is observed that less attention was paid to the relationship between crowd density and HA. In a study of the first four leagues of the English football league, it was observed that there was no corresponding difference in HA, although the mean crowd density varied from 20% in Division Four to 70% in Division One. Boyko et al. (2007) analyzed the aforementioned English Premier League and found no evidence related to the relationship between crowd density and HA (22). In our study, it was found that the number of spectators in the top league of England was an effective factor in scoring points in favor of home team but that the crowd density was not effective. The reason why crowd density is not effective for this league is believed to be due to the fact that approximately 96% of the stadium is full in matches. The fact that a similar situation is observed in the Bundesliga (90%) shows us that there are several limitations in clearly measuring the effect of crowd density on the home advantage in the near-full leagues. On the other hand, while in La Liga, the Crowd density was found to be effective on the home team's scoring points and in Ligue 1 (France), the crowd effect (Crowd size and Crowd Density) was found to be effective in the home team's scoring points. These results show us that binary regression analysis can be used to determine the crowd effect in leagues with an average occupancy rate of 70% in football leagues.

**CONCLUSION**

The idea presented and empirically tested here may be applied to bigger data sets with more observations. Other aspects of home field advantage such as travel fatigue, aggressiveness of supporters, tactical preferences made by teams at home and away and confidence levels of players may be integrated to the base model by gathering the required data for future study.

In the present study, It has been determined that crowd support and crowd density are important variables that contribute to the home field advantage in football. Developing different models in leagues with high crowd density, the effect can be attempted to be clearly separated.

However, extending this methodology to investigate further the determinants of home field advantage has some inherent difficulties. Many of the factors affecting home field advantage are strongly interrelated and some are quite difficult to be expressed in quantitative terms without making subjective declarations and assumptions. Multicollinearity makes it difficult to measure the isolated effects of the determinants of home field advantage on the home points gained. Still expressing the relationship between home field advantage and its causes as econometric models has a potential to shed more light on the determinants of the very obviously existing home field advantage in many team sports.

**APPLICABLE REMARKS**

- The results of the model used in this study suggest that there is a statistically significant level of home-field advantage for all major European leagues.
According to the results, the share of home-field advantage attributable to crowd support differs dramatically among divisions and for Serie A and Bundesliga crowd support does not even have any statistical effect on points earned at home.

- These results show us that binary regression analysis can be used to determine the crowd effect in leagues with an average occupancy rate of 70% in football leagues.

In the present study, it has been determined that crowd support and crowd density are important variables that contribute to the home-field advantage in football.

REFERENCES


