van Damme, Nils; Baert, Stijn

Article

Home advantage in European international soccer: Which dimension of distance matters?

Economics: The Open-Access, Open-Assessment E-Journal

Provided in Cooperation with:
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: van Damme, Nils; Baert, Stijn (2019) : Home advantage in European international soccer: Which dimension of distance matters?, Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 13, Iss. 2019-50, pp. 1-17,
https://doi.org/10.5018/economics-ejournal.ja.2019-50

This Version is available at:
https://hdl.handle.net/10419/209407

Economics
The Open-Access, Open-Assessment E-Journal

Vol. 13, 2019-50 | December 09, 2019 | http://dx.doi.org/10.5018/economics-ejournal.ja.2019-50

Home advantage in European international soccer: which dimension of distance matters?

Nils Van Damme and Stijn Baert

Abstract

The authors investigate whether the home advantage in soccer differs by various dimensions of distance between the (regions of the) home and away teams: geographical distance, climatic differences, cultural distance, and disparities in economic prosperity. To this end, the authors analyse 2,012 recent matches played in the UEFA Champions League and UEFA Europa League by means of several regression models. They find that when the home team plays at a higher altitude, they benefit substantially more from their home advantage. Every 100 meters of altitude difference is associated with an increase in expected probability to win the match, as the home team, by 1.1 percentage points. The other dimensions of distance are not significantly associated with a higher or lower home advantage. By contrast, the authors find that the home advantage in soccer is more outspoken when the number of spectators is higher and when the home team is substantially stronger than the away team.

JEL L83 J44 Z00

Keywords Soccer; home advantage; cultural distance; UEFA Champions League; UEFA Europa League

Authors

Nils Van Damme, Ghent University, Ghent, Belgium

Stijn Baert, Ghent University, Ghent, Belgium; Research Foundation – Flanders, Brussels, Belgium; University of Antwerp, Antwerp, Belgium; Université catholique de Louvain, Louvain-la-neuve, Belgium; Institute of Labor Economics (IZA), Bonn, Germany; Global Labor Office (GLO), Stanford, USA; IMISCOE, Rotterdam, Netherlands, Stijn.Baert@UGent.be

Citation Nils Van Damme and Stijn Baert (2019). Home advantage in European international soccer: which dimension of distance matters? *Economics: The Open-Access, Open-Assessment E-Journal*, 13 (2019-50): 1–17. http://dx.doi.org/10.5018/economics-ejournal.ja.2019-50

Received February 1, 2019 **Published as Economics Discussion Paper** March 4, 2019
Revised October 28, 2019 **Accepted** November 25, 2019 **Published** December 9, 2019

1 Introduction

The home advantage in team sports is a phenomenon that has been widely studied in peer-reviewed literature. Courneya and Carron (1992, p. 13) defined this home advantage in their review article as: "the consistent finding that home teams in sports competitions win over 50.0% of the matches played under a balanced home and away schedule." More concretely, the home advantage has been documented as a key determinant of sports game outcomes in a broad range of different team sports, including American football (Pollard and Pollard, 2005b), basketball (Ribeiro et al., 2016), field hockey (Smith et al., 2000), and ice hockey (Bray, 1999). However, this phenomenon has been studied most widely in soccer. Numerous research has centred around analysing the home advantage in soccer matches at the national level, moving from country-specific studies in Australia (Goumas, 2014a), Brazil (Pollard et al., 2008), England (Clarke and Norman, 1995; Nevill et al., 1996; Carmichael and Thomas, 2005), Germany (Oberhofer et al., 2010), Greece (Armatas and Pollard, 2012), Spain (Sánchez et al., 2009; Saavedra et al., 2015), and Turkey (Seckin and Pollard, 2008), among others, to cross-country investigations (Pollard, 2006a, 2006b; Pollard and Gómez, 2014; Leite and Pollard, 2018). Additionally, research on the home advantage in soccer has been conducted based on World Cup data (Torgler, 2004; Pollard and Armatas, 2017), international club competitions data (Page and Page, 2007; Poulter, 2009; Goumas, 2013, 2014b), and data on international football games played in South America (McSharry, 2007).

Several of the aforementioned studies have investigated the moderators of the home advantage in soccer. Among the most discussed factors influencing this home advantage are: (i) crowd effects (Nevill et al., 1996; Pollard and Pollard, 2005a; Sánchez et al., 2009; Oberhofer et al., 2010; Goumas, 2013; Ponzo and Scoppa, 2018; Pollard and Armatas, 2017); (ii) referee bias (Nevill et al., 1996; Sutter & Kocher, 2004; Nevill et al., 2013); (iii) territoriality effects (Neave and Wolfson, 2003; Pollard, 2006a, 2006b; Pollard et al., 2008; Seckin and Pollard, 2008; Armatas and Pollard, 2012; Pollard and Gómez, 2013, 2014; Pollard et al., 2017); (iv) travel effects (Clarke and Norman, 1995; McSharry, 2007; Pollard et al., 2008; Oberhofer et al., 2010; Armatas and Pollard, 2012; Bäker et al., 2012; Goumas, 2014a, 2014b; Pollard and Armatas, 2017); and (v) familiarity effects (Pollard, 2002; Watson and Krantz, 2003; Pollard and Gómez; 2014; Pollard and Armatas, 2017).

Moderators (i), (ii), and (iii) each relate to the fact that the home team typically receives stronger support from the audience, which motivates the players of the home team, and which tends to influence the referee's decisions in favour of this team. Therefore, not surprisingly, many studies have found that the larger the audience, the greater the home advantage. In addition, countries with a higher sense of territoriality, like those in the Balkan region, are generally found to have a greater home advantage (Pollard, 2006a, 2006b; Pollard and Gómez, 2013, 2014).

Moderators (iv) and (v) address the fact that the away team may experience fatigue due to travel-related factors and that the home team has the advantage of being familiar with the circumstances in the city of the stadium, both resulting in a higher relative productivity of the home team. Crucial with respect to (iv) and (v) are various aspects of *distance* between home and away teams. In this respect, small—but significant—positive associations between home advantage and distance travelled are found in England (Clarke and Norman, 1995), Brazil

(Pollard et al., 2008), Germany (Oberhofer et al., 2010), and in international European soccer matches (Goumas, 2014b), but not in Greece (Armatas and Pollard, 2012) nor Australia (Goumas, 2014a). Relatedly, Seckin and Pollard (2008), Bäker et al. (2012), and Leite and Pollard (2018) indicate that the home advantage is substantially smaller or even completely vanishes whenever a match is a derby. In addition, McSharry (2007) and Pollard and Armatas (2017) report there is a significant association between home advantage and altitude, with each 1,000 m in altitude difference worth, on average, an increase in the goal difference by half of a goal according to the first study and 0.115 of a point's advantage for the home team according to the second study. Last, Pollard et al. (2017) report that playing in high humidity increases home advantage.

However, this literature on the relationship between home advantage in soccer and distance between the home and away teams is characterised by an important gap. That is, all mentioned studies investigate one or two variables related to geographical distance while making abstraction of other dimensions of distance. In other words, they neglect that distance between two teams can go beyond mere measurable miles. From an empirical point of view, their approach may result in an omitted variable bias. Indeed, the included (geographical) distance measures may pick up the moderating effect of other dimensions of distance that are not included. For instance, the travel length variables included in previous studies may pick up the effect of temperature differences between the cities of the home and away teams (to which away players have to adapt).

The present study aims to fill this gap. We investigate the association between home advantage in European international soccer and multiple perspectives of the factor of distance between home and away teams. More concretely, we investigate whether home advantage in soccer is heterogeneous by (a) geographical distance (travel length and difference in altitude); (b) climatic differences (with respect to temperature and precipitation); (c) cultural distance; (d) and disparities in economic prosperity between the regions of the home and away teams, keeping heterogeneity in the home advantage by the number of spectators, the derby status of the match, the home advantage at the national competition level, and the teams' relative strength constant. We are not aware of any previous work investigating the importance of distance factors (c) or (d) in the home advantage in soccer, let alone previous work investigating them within one statistical framework.

To this end, we analyse 2,012 matches in the Union of European Football Associations (UEFA) Champions League and UEFA Europa League between 2008 and 2016. The match data are merged with country and city-level data. These data also allow us to test, as a first study, whether or not the home advantage in international soccer matches is different in derbies and whether or not an elevated home advantage in the national leagues in the Balkan translates into a higher home advantage for Balkan teams in international matches.

2 Methods

2.1 Data

The basis of our dataset was formed by match reports from all matches in the UEFA Champions League between 2008 and 2016, and all matches in the UEFA Europa League between 2011 and 2016—before 2011, another competition format was used for the latter competition. These data were collected from the official website of the UEFA (UEFA; http://www.uefa.com). The UEFA Champions League, which is the most prestigious club competition in the world, and the UEFA Europa League both begin with a group stage of 32 and 48 teams, respectively, divided into groups of four teams, where each team plays against the other once at home and once away. The group stage of each season is played from September to December. The teams finishing first and second in each group proceed to the knock-out stage of their competition. Additionally, the teams finishing third in each group of the UEFA Champions League enter the UEFA Europa League knock-out stage. The knock-out stage of both competitions is played from February to May. During this phase, teams meet each other in one home and one away match after which the team with the positive goal difference over these two matches (potentially after additional time and penalties) advances to the following round. In total, 125 and 205 matches are played in each season of the UEFA Champions League and UEFA Europa League, respectively, which totals of 2,025 matches within the mentioned time frame. However, the 13 final matches were excluded from our analyses, given that they were played on a neutral pitch (without home advantage). Consequently, our analyses are based on 2,012 match reports. For more information on the regulations of the two competitions and the rules of a soccer match, we refer to the UEFA website (http://www.uefa.com) and to FIFA (2017).

Following the approach used by Ponzo and Scoppa (2018), we considered each match twice in our data, one time from the perspective of the home team and one time from the perspective of the away team. This generated a total of 4,024 observations at the team-match level. As the outcome variables are closely related for the observations of the home and away teams at the match level, we clustered the standard errors in our regression analyses at this level. In addition, as a robustness check, we redid our analyses after randomly assigning each match either to the home or to the away team, thereby considering each match only once. However, this alternative approach did not yield different empirical conclusions. In what follows, we will always refer to a match between a ‘team’ and its ‘opponent’, where ‘team’ is the home team and ‘opponent’ is the away team if the match is viewed from the perspective of the home team, and vice versa.

Table 1 presents descriptive statistics for all of the variables used in the regression analysis below, together with their definitions and their respective sources. Panel A describes the variables used as dependent variables in our analysis. We constructed three distinctive variables capturing the outcome of the match at full time from the perspective of the team under concern: (i) goal difference, (ii) victory, and (iii) number of points. The mean value of 0.000 for (i) is a direct consequence of the construction of our dataset, where, as aforementioned, we considered each match twice. Using the mean value of victory, equal to 0.379, we can deduce that 24.2%, i.e. $1 - 2 \times 0.379$, of the matches ended in a draw. As a victory yields three points and a draw yields one point, not surprisingly, each team obtained about 1.379 points per match on average.

Table 1: Data: Summary Statistics

Variable name	Definition	Source	Mean	SD	N
A. Dependent variables					
Goal difference	Number of goals at full time for team minus number of goals for opponent	http://www.uefa.com	0.000	1.913	4,024
Victory	Equal to 1 if team wins the match, 0 otherwise	http://www.uefa.com	0.379	-	4,024
Number of points	Number of points obtained by team at full time (3 for a victory, 1 for a draw, and 0 for a loss)	http://www.uefa.com	1.379	1.324	4,024
B. Independent variable					
Home	Equal to 1 if team is the home team, 0 otherwise	http://www.uefa.com	0.500	-	4,024
C. Dimensions of potential heterogeneity in home effect					
Distance: travel length	Distance (in 1,000 km) in bird's eye view between city of team and opponent	http://www.2travel2.nl	1.623	0.922	4,024
Distance: altitude	Difference in meters above sea level (in 100 m increments) between city of team and opponent	http://www.weernetwerk.nl	0.000	2.359	4,024
Distance: temperature	Difference in mean temperature (in °C) between city of team and opponent in month of match	http://www.timeanddate.com	0.000	6.365	4,024
Distance: precipitation	Difference in mean precipitation (in 100 mm) between city of team and opponent in month of match	http://www.timeanddate.com	0.000	0.350	4,024
Distance: culture	Cultural Distance Index between country of team and opponent	Spolaore and Wacziarg (2016)	−27.078	29.078	3,832
Distance: wealth	Difference in GDP (in 1,000 $) per capita between country of team and opponent in year of match	http://www.worldbank.org	0.000	2.558	4,024
Absolute distance: altitude	Absolute value of 'Distance: altitude'	http://www.weernetwerk.nl	1.613	1.721	4,024
Absolute distance: temperature	Absolute value of 'Distance: temperature'	http://www.timeanddate.com	5.037	3.889	4,024
Absolute distance: precipitation	Absolute value of 'Distance: precipitation'	http://www.timeanddate.com	0.241	0.253	4,024
Absolute distance: wealth	Absolute value of 'Distance: wealth'	http://www.worldbank.org	1.988	1.610	4,024
Spectators	Number of spectators (in 1,000 persons) present	http://www.worldfootball.net	31.101	20.431	4,024
Derby	Equal to 1 if team and opponent are from same country, 0 otherwise	http://www.uefa.com	0.021	-	4,024
Balkans	Equal to 1 if team comes from Balkan country, 0 otherwise	Pollard (2006b)	0.045	-	4,024
Northern Europe	Equal to 1 if team comes from Northern European country, 0 otherwise	Pollard (2006b)	0.158	-	4,024
Relative strength	Difference between UEFA coefficient of team and opponent in year of match	http://www.uefa.com	0.000	52.914	4,024
D. Selection variables					
Own stadium	Equal to 1 if home team plays match in own stadium, 0 otherwise	http://www.uefa.com	0.033	-	4,024
Without competitive value	Equal to 1 if match has no competitive value for team or opponent, 0 otherwise	http://www.uefa.com	0.147	-	4,024
Knock-out stage	Equal to 1 if match is knock-out stage match, 0 otherwise	http://www.uefa.com	0.285	-	4,024

Notes. Some abbreviations are used: GDP (gross domestic product) and UEFA (Union of European Football Associations). No standard deviations are reported for binary variables.

Panel B of Table 1 presents the main independent variable, i.e. the home team status of the considered team. Given the construction of our dataset, half of the observations capture match events from the perspective of the home team. Panel C shows the variables by which the advantage of this team (over the away team) may differ. As mentioned in the introduction, we included six such variables that relate to the multi-dimensional 'distance' between the home and away teams.

First, geographical distance is captured by the variables 'Distance: travel length' (average distance determined using a bird's eye view between the city of the home team and its opponent) and 'Distance: altitude' (difference in meters above sea level between the two cities). The highest travel length (of 6,173 km) is observed between the stadiums of Benfica (Lisbon, Portugal) and FC Astana (Nur-Sultan, Kazakhstan). The stadium with the lowest altitude is that of Qarabağ (Baku, Azerbaijan; 7 meters below sea level), while the stadium with the highest altitude is that of FC St. Gallen (St. Gallen, Switzerland; 779 meters above sea level).

Second, distance regarding climatic differences between the cities of the home team and the away team are determined by their temperature and precipitation differences, both measured against the month of the match. The lowest and highest (average) temperatures are measured in Kazan in February (FC Rubin Kazan, Russia; −10°C) and Tel-Aviv in September (Maccabi Tel Aviv FC and Hapoel Tel Aviv FC, Israel; 27°C). In Tel-Aviv, also the precipitation level is the lowest (0 mm in September); it is the highest in San Sebastián (Real Sociedad, Spain; 181 mm in November).

Third, the teams' cultural distance is based on the Cultural Distance Index constructed by Spolaore and Wacziarg (2016). Using the answers to a questionnaire containing enquiries related to six different value-related categories collected from people in 71 countries, Spolaore and Wacziarg (2016) calculated the cultural variance index for 2,701 pairs of countries. This index is not available for 192 observations—matches with teams from Israel are overrepresented in these observations. The smaller the value of the cultural distance index, the smaller the cultural distance between the two countries under review. For example, the smallest cultural distance in our dataset is that observed between Russian and Ukrainian teams, equal to –89.820—the same value is used when two teams from the same country play against each other. We find the biggest cultural distance between Denmark and Turkey, with a value of 81.670.

Fourth and last, the disparity in economic prosperity between the country of the home and away teams is operationalised by their difference in gross domestic product (GDP) per capita. The lowest GDP level is measured in Ukraine in 2015 (FC Dnipro, FC Dynamo Kyiv, and FC Shakhtar Donetsk; 2,125 euro) while the highest level is measured in Norway in 2013 (Tromsø IL; 102,910 euro).

We believe these four dimensions ensure a focus on the most relevant aspects of distances in Europe. Substantial correlations are found between these dimensions. In particular, teams that are at a large distance in bird's eye view are often characterised by a high cultural distance (*Pearson's* $r = 0.541$). In addition, other significant correlations (at the 5% significance level) are those between (i) altitude difference and temperature difference ($r = -0.097$), (ii) altitude difference and precipitation difference ($r = -0.058$), (iii) altitude difference and wealth difference ($r = 0.126$), and (iv) precipitation difference and wealth difference ($r = 0.165$).

Two of the six distance variables are equal for the home and away teams at the match level: travel distance and cultural distance. The four other distance variables have a direction, so that their value for the home team is the opposite of that of the away team (and their average value is, by construction of our data, 0): altitude difference, temperature difference, precipitation difference, and wealth difference. For the latter variables, we also constructed the corresponding distance in absolute values. These variables are added to the regression model in our extended analysis. Including these absolute values makes it possible to determine whether it is a difference (or shock) in these variables that determines the home advantage, irrespective of its direction, or whether it is a difference *in a certain direction* that yields an additional home premium.

The other variables in Panel C are match characteristics with a potential influence on the home advantage that are not related to the distance between the home and away teams. First, to capture crowd effects in a direct way, we included a variable capturing the number of spectators. As mentioned in our introduction, this variable often recurs in the literature as a factor that increases the home advantage. The average number of spectators in the analysed matches was 31,101. Second, we adopted a derby variable to check, as a first study, whether the home advantage varies by this variable in European international soccer as it does in national matches in Germany (Bäker et al., 2012). Third, we included indicators for teams from the Balkans and Northern Europe. Following Pollard (2006b), the home advantage in national leagues in the Balkans is generally higher than elsewhere in Europe, while the home advantage in Northern Europe (including the Baltic states, Scandinavian countries, Iceland, and the five countries of the British Isles) is lower than average. By means of our regression framework, we can test whether the higher (lower) home advantage in national leagues in the Balkan (Northern Europe) is also reflected in a higher (lower) home advantage for Balkan (Northern European) teams in international matches. A final potential moderator of home advantage that we investigate is the relative strength of the team and its opponent. This relative strength is captured by the teams' difference in UEFA coefficient—the UEFA coefficient of a team is based on its participation and results in the previous five seasons of the UEFA Champions League and UEFA Europa League.

The variables in Panel D of Table 1 are used to confirm whether the performed analyses are robust for (i) the exclusion of matches in which the home team does not play in their own stadium; (ii) the exclusion of matches without a competitive value for the team or its opponent; and (iii) the exclusion of matches in the knock-out stage. A team does not play in its own stadium if its stadium does not meet the requirements of the UEFA, for example, at Zulte Waregem (Belgium), when participating to the Europa League in 2013, or if there are security concerns, as at Shakhtar Donetsk (Ukraine), when participating to the Champions League in 2014. In those instances, the home team has only a pseudo home status. This occurred, however, only in 3.3% of the analysed matches. Next, we define a match without competitive value to be a match in the group stage where it was mathematically impossible for the team and/or its opponent to change their qualification status for the next stage. A third robustness check is performed to see whether the home advantage patterns in our data remain when matches in the knock-out phase are excluded. This is considered given that, as aforementioned, additional time and a penalty shoot-out may be added to the return match of a round, potentially resulting in an additional home advantage for the home team of such matches.

2.2 Statistical approach

We analysed the data presented in the former subsection using linear regression models. All models that are estimated can be represented by means of the following general equation:

$$Y_{i,n} = \alpha + \beta D_{i,n} + \boldsymbol{\gamma} D_{i,n} \boldsymbol{X_{i,n}} + \varepsilon_{i,n}.$$

In this equation, $Y_{i,n}$ is the dependent variable: the outcome of the *n*th match, from the point of view of a team *i*. $D_{i,n}$ is the dummy variable capturing the home team status of team *i* in match *n*. $\boldsymbol{X_{i,n}}$ is a vector of distance-related and other variables according to which the association of $D_{i,n}$ with $Y_{i,n}$ may be heterogeneous. α is the intercept of the model, β is the coefficient related to $D_{i,n}$, $\boldsymbol{\gamma}$ is a vector of coefficients associated with $D_{i,n}\boldsymbol{X_{i,n}}$ and $\varepsilon_{i,n}$ is the error term. As aforementioned, we clustered the standard errors at the match level to correct for the correlation between the error terms due to the two observations per match. In addition, this clustering of the standard errors corrects for their heteroscedasticity due to the fact that our dependent variables, 'Victory' in particular, are not normally distributed (Angrist and Pischke, 2008; Baert and Amez, 2018; Van Den Broucke and Baert, 2019). However, we also estimated (ordered) logit models, yielding the same research conclusions.

It is important to notice that we did not include the $\boldsymbol{X_{i,n}}$ without interaction with $D_{i,n}$ (*ergo*, as a control variable). Correlation between $D_{i,n}$ and the $\boldsymbol{X_{i,n}}$ is impossible given the construction of our dataset, where for every combination of teams, there is always a match where one team is the home team and the other one is the away team, and a match where the opposite is true. For the same reason, controlling for team fixed effects is not rational. Furthermore, it is not desirable to include the $\boldsymbol{X_{i,n}}$ as such because we would then consistently divide the total home effect into an effect of home advantage and away disadvantage, which would not be consistent with the literature mentioned in the introduction.

The $\boldsymbol{X_{i,n}}$ in $D_{i,n}\boldsymbol{X_{i,n}}$ were mean-centred so that throughout the regression models, we can interpret β as the average effect of playing at home. For each of the models presented in the Results section, we computed multicollinearity diagnostics leading to variance inflation factors substantially lower than 5.

3 Results

3.1 Main analysis

Table 2 presents the results of our benchmark analysis. In regressions (1) – (3), we regress goal difference (model (1)), victory (model (2)) and number of points (model (3)) on the home status of the team only. In regressions (4) – (6), we redo the same analyses after adding the interactions between the home status of the team and the variables from Panel C of Table 1.

Table 2: Results: Benchmark Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Home	0.817*** (0.083)	0.184*** (0.019)	0.553*** (0.057)	0.811*** (0.080)	0.183*** (0.019)	0.550*** (0.056)
Home × Distance: travel length				−0.010 (0.052)	−0.001 (0.015)	0.007 (0.039)
Home × Distance: altitude				0.050*** (0.018)	0.011** (0.005)	0.032*** (0.012)
Home × Distance: temperature				0.002 (0.006)	0.002 (0.002)	0.004 (0.005)
Home × Distance: precipitation				−0.146 (0.118)	−0.013 (0.031)	−0.051 (0.080)
Home × Distance: culture				0.002 (0.002)	0.001* (0.000)	0.002 (0.001)
Home × Distance: wealth				0.002 (0.016)	0.005 (0.005)	0.011 (0.012)
Home × Spectators				0.009*** (0.002)	0.002*** (0.001)	0.007*** (0.001)
Home × Derby				−0.291 (0.244)	−0.030 (0.082)	−0.092 (0.215)
Home × Balkans				−0.225 (0.221)	−0.063 (0.054)	−0.212 (0.149)
Home × Northern Europe				−0.134 (0.113)	−0.024 (0.032)	−0.077 (0.083)
Home × Relative strength				0.012*** (0.001)	0.002*** (0.000)	0.007*** (0.001)
Intercept	−0.409*** (0.042)	0.287*** (0.010)	1.103*** (0.028)	−0.408*** (0.043)	0.287*** (0.010)	1.102*** (0.029)
Dependent variable: Goal difference	Yes	No	No	Yes	No	No
Dependent variable: Victory	No	Yes	No	No	Yes	No
Dependent variable: Number of points	No	No	Yes	No	No	Yes
R^2	0.046	0.036	0.044	0.124	0.093	0.102
N	4,024	4,024	4,024	3,832	3,832	3,832

Notes. A definition of the included variables can be found in Table 1. The variables interacted with 'Home' are mean-centred. The presented statistics are linear regression model estimates and standard errors, clustered at the match level, in parentheses. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance level.

The estimation results concerning the average effect of playing at home are robust across the six regression models. A highly significant (p = 0.000) positive association is found between playing at home and the outcome of the match in terms of our three dependent variables. After including the interaction variables, we find that playing at home increases (i) the expected goal difference at full time by 0.811 goals, (ii) the expected probability of a victory by 18.3 percentage points, and (iii) the expected number of points by 0.550, all other variables held constant.

With respect to the importance of the multi-dimensional distance between the home and away teams, only the interaction with the altitude difference between the teams has a significant coefficient. Every additional 100 m above sea level is associated with (i) an increase in the goal difference by 0.050 goals (p = 0.006), (ii) an increase in the chance of a victory by 1.1 percentage points (p = 0.014), and (iii) an increase in points by 0.032 (p = 0.008) for the home team. The increase in the goal difference is exactly of the same magnitude as that reported in McSharry (2007), namely an increase of about half a goal for each 1,000 meter additional distance in altitude. In contrast, the increase in points we find (10×0.032 per 1,000 meter) is more than three times as high as the 0.115 additional points for each 1,000 meter additional distance in altitude reported by Pollard and Armatas (2017). This higher association may be related to the fact that Europeans might be less used to substantial differences in altitude. Given this striking difference in magnitude, we estimated several alternative regression models with non-linear specifications (adopting, for instance, the natural logarithm of our distance measure) but these did not seem to capture our data better.

With respect to the other distance variables, we identify a small and weakly significant (p = 0.053) coefficient for the interaction between the home status of a team and its cultural distance to the away team in regression (5), but not in regressions (4) and (6), which indicates that this may be a statistical artefact.

Regarding the other interaction variables, we find that the home advantage is consistently higher when the number of spectators is higher and when the relative strength of the home team is more substantial. Per 1,000 additional spectators (one unit increase in the relative strength index), the goal difference in favour of the home team increases with 0.009 (0.012), the chance of a home win increases by 0.2 (0.2) percentage points, and (iii) the number of points obtained by the home team increases by 0.007 (0.007). We do not find evidence for the home advantage to be heterogeneous by the derby status of the match or the region of the country of the team (Balkan, Northern European, or other).

Table 3 presents the results of an extended analysis where we include the absolute values of the distance variables with a direction. As aforementioned, this allows us to check whether it is a shock in these distances that determines the home advantage, or a shock only heading in a certain direction. Regarding the altitude difference, we see that the direction of this difference is important. The regular altitude distance variable is very comparable to that in Table 2 after including its absolute value, while this absolute value is not significant. So, again, when the home team plays at a higher (lower) altitude, they benefit more (less) from their home advantage. Furthermore, we notice a significantly positive association between home advantage and the absolute wealth difference between the competing teams. An additional difference in wealth between the country of the home team and the country of the away team of 1,000 dollars

Table 3: Results: Extended Analysis

	(1)	(2)	(3)
Home	0.809*** (0.080)	0.183*** (0.019)	0.550*** (0.056)
Home × Distance: travel length	0.005 (0.061)	0.001 (0.017)	0.014 (0.045)
Home × Distance: altitude	0.051*** (0.018)	0.011** (0.005)	0.032*** (0.012)
Home × Absolute distance: altitude	0.024 (0.024)	0.004 (0.006)	−0.002 (0.017)
Home × Distance: temperature	0.003 (0.006)	0.002 (0.002)	0.004 (0.005)
Home × Absolute distance: temperature	−0.002 (0.013)	0.000 (0.003)	−0.002 (0.009)
Home × Distance: precipitation	−0.146 (0.115)	−0.013 (0.031)	−0.050 (0.080)
Home × Absolute distance: precipitation	0.158 (0.158)	0.014 (0.041)	0.040 (0.106)
Home × Distance: culture	−0.000 (0.002)	0.000 (0.001)	0.001 (0.002)
Home × Distance: wealth	0.001 (0.016)	0.005 (0.005)	0.010 (0.012)
Home × Absolute distance: wealth	0.076** (0.031)	0.013 (0.009)	0.031 (0.023)
Home × Spectators	0.009*** (0.002)	0.003*** (0.001)	0.007*** (0.001)
Home × Derby	−0.294 (0.245)	−0.032 (0.083)	−0.101 (0.215)
Home × Balkans	−0.298 (0.223)	−0.074 (0.054)	−0.234 (0.150)
Home × Northern Europe	−0.128 (0.115)	−0.022 (0.032)	−0.082 (0.084)
Home × Relative strength	0.012*** (0.001)	0.002*** (0.000)	0.007*** (0.001)
Intercept	−0.408*** (0.043)	0.287*** (0.010)	1.102*** (0.029)
Dependent variable: Goal difference	Yes	No	No
Dependent variable: Victory	No	Yes	No
Dependent variable: Number of points	No	No	Yes
R^2	0.126	0.094	0.103
N	3,832	3,832	3,832

Notes. A definition of the included variables can be found in Table 1. The variables interacted with 'Home' are mean-centred. The presented statistics are linear regression model estimates and standard errors, clustered at the match level, in parentheses. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance level.

per capita increases the home advantage in terms of goal difference by 0.076 goals. However, this interaction is not significant in regressions (2) and (3), so, again, this result should be interpreted with caution.

3.2 Robustness checks

As our main finding of a higher home advantage for home teams playing at a higher altitude could be driven by a few outliers, we discuss an outlier analysis. More concretely, in Table 4 we replicate regressions (4), (5), and (6) of Table 2 after excluding matches with a distinctive distance in altitude. That is, for columns (1), (2), and (3) of Table 4, we exclude matches where the distance in terms of altitude is more than three standard deviations higher or lower than the average of 0 (so, with a distance of 707.7 meters or more). In addition, for columns (4), (5), and (6), we exclude matches where this distance is more than two standard deviations higher or lower than the average. As a consequence, the number of observations is reduced from 3,832 observations to 3,804 observations and 3,660 observations in the first three and last three columns, respectively. However, the regressions results are very comparable to those presented in Table 2.

In addition, as mentioned in the Data subsection, we redid our benchmark analysis after (i) the exclusion of matches when the home team does not play in their own stadium, (ii) the exclusion of matches without any competitive value for the team or their opponent, and (iii) the exclusion of matches in the knock-out stage. However, none of these analyses, the results of which can be obtained upon request, has led to other insights than those of the benchmark analysis.

4 Conclusion

This study contributed to the literature about home advantage in soccer in several ways. Former contributions to this literature investigated how this home advantage varies by the geographical distance between the home and away teams neglecting other dimensions of distance (and the related, potential omitted variable bias in their estimates). In contrast, we investigated heterogeneity in the home effect by geographical distance (travel length and difference in altitude), climatic differences (temperature and precipitation), cultural distance as well as disparities in economic prosperity between the region of the home and away teams. In addition, we allowed the measured home advantage to vary by the number of spectators, the derby status of the match, the home advantage at the national competition level, and the teams' relative strength. To this end, 2,012 matches in the UEFA Champions League and UEFA Europa League between 2008 and 2016 were analysed.

We found, first, in line with the literature, there is a highly significantly positive association between playing at home and ending the match in a favourable position. Second, the altitude difference stood out as the one major distance-related moderator of this home advantage. Each

Table 4: Results: Outlier Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Home	0.798*** (0.080)	0.181*** (0.019)	0.544*** (0.056)	0.808*** (0.081)	0.185*** (0.019)	0.557*** (0.057)
Home × Distance: travel length	−0.011 (0.052)	−0.002 (0.015)	0.006 (0.039)	−0.020 (0.052)	−0.003 (0.015)	0.004 (0.040)
Home × Distance: altitude	0.064*** (0.019)	0.013*** (0.005)	0.039*** (0.013)	0.047** (0.022)	0.010* (0.006)	0.031** (0.015)
Home × Distance: temperature	0.004 (0.006)	0.002 (0.002)	0.005 (0.005)	0.004 (0.006)	0.002 (0.002)	0.005 (0.005)
Home × Distance: precipitation	−0.119 (0.118)	−0.009 (0.031)	−0.042 (0.081)	−0.127 (0.120)	−0.011 (0.031)	−0.043 (0.082)
Home × Distance: culture	0.002 (0.002)	0.001** (0.000)	0.002 (0.001)	0.002 (0.002)	0.001* (0.000)	0.002 (0.001)
Home × Distance: wealth	0.001 (0.016)	0.005 (0.005)	0.010 (0.012)	0.003 (0.017)	0.005 (0.005)	0.010 (0.013)
Home × Spectators	0.009*** (0.002)	0.003*** (0.001)	0.007*** (0.001)	0.009*** (0.002)	0.003*** (0.001)	0.007*** (0.002)
Home × Derby	−0.290 (0.243)	−0.027 (0.082)	−0.087 (0.214)	−0.365 (0.249)	−0.065 (0.085)	−0.168 (0.221)
Home × Balkans	−0.239 (0.221)	−0.065 (0.054)	−0.219 (0.150)	−0.259 (0.227)	−0.066 (0.055)	−0.232 (0.154)
Home × Northern Europe	−0.132 (0.114)	−0.025 (0.032)	−0.079 (0.083)	−0.125 (0.114)	−0.024 (0.032)	−0.077 (0.083)
Home × Relative strength	0.012*** (0.001)	0.002*** (0.000)	0.007*** (0.001)	0.012*** (0.001)	0.002*** (0.000)	0.007*** (0.001)
Intercept	−0.402*** (0.043)	0.288*** (0.010)	1.105*** (0.029)	−0.428*** (0.043)	0.280*** (0.011)	1.084*** (0.030)
Dependent variable: Goal difference	Yes	No	No	Yes	No	No
Dependent variable: Victory	No	Yes	No	No	Yes	No
Dependent variable: Number of points	No	No	Yes	No	No	Yes
Exclusion of matches with distance: altitude > 3×SD(distance: altitude) or distance: altitude < −3×SD(distance: altitude)	Yes	Yes	Yes	No	No	No
Exclusion of matches with distance: altitude > 2×SD(distance: altitude) or distance: altitude < −2×SD(distance: altitude)	No	No	No	Yes	Yes	Yes
R^2	0.123	0.092	0.101	0.114	0.088	0.098
N	3,804	3,804	3,804	3,660	3,660	3,660

Notes. A definition of the included variables can be found in Table 1. The variables interacted with 'Home' are mean-centred. The presented statistics are linear regression model estimates and standard errors, clustered at the match level, in parentheses. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance level.

100 m of rising above sea level is associated with an increase of the home advantage by 0.032points. A possible explanation for this may be that the available oxygen decreases with increasing altitude. Home team players are likely to be more adapted to performing well in the condition of low oxygen levels. Other explanations, as mentioned by an anonymous referee commenting the original version of the present study (Van Damme and Baert, 2019), might be a decreased air friction and a higher ball velocity. Third, we found that the home advantage in soccer is more outspoken when the number of spectators is higher and when the home team is substantially stronger (in terms of UEFA coefficient) than the away team. These findings are consistent with Nevill et al. (1996), Goumas (2013), Ponzo and Scoppa (2018), and Pollard and Armatas (2017). Finally, no significant association was found with variables capturing derby matches and variables portraying the home advantage at the national level. The latter finding is remarkable, especially for the countries in the Balkans, because the higher home advantage identified in these countries' national leagues often recurs in the literature (Pollard, 2006a, 2006b; Pollard and Gómez, 2013, 2014).

We end this study by acknowledging its main research limitation. By means of investigating how the home advantage in soccer is associated with a broad spectrum of distance-related variables, we took a step forward in measuring the unbiased, independent importance of these determinants of the home advantage. Yet, the related coefficient estimates mentioned in this article cannot be given a causal interpretation. This is the case as there might be still other factors that we did not include in our study but may correlate with our distance dimensions and with performance in soccer. Therefore, we are in favour of (i) future empirical work that exploits (quasi-)experimental variation in one or more of these dimensions to investigate their genuine causal impact and (ii) qualitative research on the mechanisms underlying the reported association between home advantage and the altitude difference between the cities of the home team and the away team.

References

Angrist, J.D., and Pischke, J.S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Armatas, V., and Pollard, R. (2012). Home advantage in Greek football. *European Journal of Sport Science*, 14(2), 116–122. https://doi.org/10.1080/17461391.2012.736537

Baert, S., and Amez, S. (2018). No better moment to score a goal than just before half time? A soccer myth statistically tested. *PLOS ONE*, 13, e0194255. https://doi.org/10.1371/journal.pone.0194255

Bäker, A., Mechtel, M., and Vetter, K. (2012). Beating thy neighbor: Derby effects in German professional soccer. *Journal of Economics and Statistics*, 232, 224–246. http://dx.doi.org/10.2139/ssrn.1718355

Bray, S.R. (1999). The home advantage from an individual team perspective. *Journal of Applied Sport Psychology*, 11(1), 116–125. https://doi.org/10.1080/10413209908402954

Carmichael, F., and Thomas, D. (2005). Home-field effect and team performance. *Journal of Sports Economics*, 6(3), 264–281. https://doi.org/10.1177/1527002504266154

Clarke, S.R., and Norman, J. M. (1995). Home advantage of individual clubs in English soccer. *Journal of the Royal Statistical Society—Series D*, 44(4), 509–521. https://www.jstor.org/stable/2348899

Courneya, K.S., and Carron, A.V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology*, 14, 13–27. https://doi.org/10.1123/jsep.14.1.13

FIFA (2017). *Laws of the game 2016/2017*. Zurich: FIFA.

Goumas, C. (2013). Home advantage and crowd size in soccer: A worldwide study. *Journal of Sport Behavior*, 36(4), 387–399. https://www.questia.com/library/journal/1G1-351081108/home-advantage-and-crowd-size-in-soccer-a-worldwide

Goumas, C. (2014a). Home advantage in Australian soccer. *Journal of Science and Medicine in Sport*, 17(1), 119–123. https://www.sciencedirect.com/science/article/abs/pii/S1440244013000467

Goumas, C. (2014b). Tyranny of distance: Home advantage and travel in international club football. *International Journal of Performance Analysis in Sport*, 14(1), 1–13. https://doi.org/10.1080/24748668.2014.11868698

Leite, W., and Pollard, R. (2018). International comparison of differences in home advantage between level 1 and level 2 of domestic football leagues. *German Journal of Exercise and Sport Research*, 48(2), 271–277. https://doi.org/10.1007/s12662-018-0507-2

McSharry, P.E. (2007). Effect of altitude on physiological performance: a statistical analysis using results of international football games. *BMJ*, 335, 1278–1281. https://doi.org/10.1136/bmj.39393.451516.AD

Neave, N., and Wolfson, S. (2003). Testosterone, territoriality, and the 'home advantage'. *Physiology and Behavior*, 78(2), 269–275. http://doi.org/10.1016/s0031-9384(02)00969-1

Nevill, A.M., Newell, S.M., and Gale, S. (1996). Factors associated with home advantage in English and Scottish soccer matches. *Journal of Sports Sciences*, 14(2), 181–186. http://doi.org/10.1080/02640419608727700

Nevill, A., Webb, T., and Watts, A. (2013). Improved training of football referees and the decline in home advantage post-WW2. *Psychology of Sport and Exercise*, 14(2), 220–227. http://doi.org/10.1016/j.psychsport.2012.11.001

Oberhofer, H., Philippovich, T., and Winner, H. (2010). Distance matters in away games: Evidence from the German football league. *Journal of Economic Psychology*, 31(2), 200–211. https://doi.org/10.1016/j.joep.2009.11.003

Page, L., and Page, K. (2007). The second leg home advantage: Evidence from European football cup competitions. *Journal of Sports Sciences*, 25(14), 1547–1556. https://doi.org/10.1080/02640410701275219

Pollard, R. (2002). Evidence of a reduced home advantage when a team moves to a new stadium. *Journal of Sports Sciences*, 20(12), 969–973. https://doi.org/10.1080/026404102321011724

Pollard, R. (2006a). 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*, 29(2), 169–189. https://www.questia.com/library/journal/1G1-146122554/home-advantage-in-soccer-variations-in-its-magnitude

Pollard, R. (2006b). Worldwide regional variations in home advantage in association football. *Journal of Sports Sciences*, 24(3), 231–240. https://doi.org/10.1080/02640410500141836

Pollard, R., and Armatas, V. (2017). Factors affecting home advantage in football World Cup qualification. *International Journal of Performance Analysis in Sport*, 17(1–2), 121–135. https://doi.org/10.1080/24748668.2017.1304031

Pollard, R., Armatas, V., and Zamani Sani, S.H. (2017). Home advantage in professional football in Iran: Differences between teams, levels of play and the effects of climate. *International Journal of Science Culture and Sport, 5,* 328–339. https://doi.org/10.14486/IntJSCS696

Pollard, R., and Gómez, M.Á. (2013). Variations in home advantage in the national basketball leagues in Europe. *Revista de Psicologia del Deporte*, 22, 263–266. https://www.rpd-online.com/article/view/1338

Pollard, R., and Gómez, M.Á. (2014). Comparison of home advantage in 157 national leagues worldwide. *International Journal of Sport and Exercise Psychology*, 12(3), 218–233. https://doi.org/10.1080/1612197X.2014.888245

Pollard, R., and Pollard, G. (2005a). Home advantage in soccer: A review of its existence and causes. *International Journal of Soccer and Science,* 3, 28–38. https://doi.org/10.2174/1875399X00801010012

Pollard, R., and Pollard, G. (2005b). 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. https://doi.org/10.1080/02640410400021559

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

Pollard, R., Silva, C.D., and Medeiros, N.C. (2008). Home advantage in football in Brazil: Differences between teams and the effects of distance travelled. *Brazilian Journal of Soccer Science*, 1, 3–10. https://www2.dti.ufv.br/seer/rbf/index.php/RBFutebol/article/view/6

Ponzo, M., and Scoppa, V. (2018). Does the home advantage depend on crowd support? Evidence from same-stadium derbies. *Journal of Sports Economics*, 19(4), 562–582. https://doi.org/10.1177/1527002516665794

Poulter, D. (2009). Home advantage and player nationality in international club football. *Journal of Sports Medicine*, 27(8), 797–805. URL https://doi.org/10.1080/02640410902893364

Ribeiro, H.V., Mukherjee, S., and Zeng, X.H.T. (2016). The advantage of playing home in NBA: Microscopic, team-specific and evolving features. *PLOS ONE*, 11, e0152440. https://doi.org/10.1371/journal.pone.0152440

Saavedra, M., Gutiérrez, Ó., Fernández, J.J., and Sa, P. (2015). Measuring home advantage in Spanish football (1928-2011). *Revista Internacional de Medicina y Ciencias de la Actividad Fisica y del Deporte*, 15(57), 181–194. http://cdeporte.rediris.es/revista/revista57/artventaja550.htm

Sánchez, P., Garcia-Calvo, T., Leo, F., Pollard, R., and Gómez, M.Á. (2009). An analysis of home advantage in the top two Spanish professional football leagues. *Perceptual and Motor Skills*, 108(3), 789–797. https://doi.org/10.2466/PMS.108.3.789-797

Seckin, A., and Pollard, R. (2008). Home advantage in Turkish professional soccer. *Perceptual and Motor Skills*, 107(1), 51–54. https://doi.org/10.2466/PMS.107.5.51-54

Smith, D.R., Ciacciarelli, A., Serzan, J., and Lambert, D. (2000). Travel and the home advantage in professional sports. *Sociology of Sport Journal*, 17(4), 364–385. https://doi.org/10.1123/ssj.17.4.364

Spolaore, E., and Wacziarg, R. (2016). Ancestry, language and culture. In: Ginsburgh, V., and Weber, S. (eds.), *The Palgrave Handbook of Economics and Language*. London: Palgrave Macmillan.

Sutter, M., and Kocher, M. (2004). Favoritism of agents – The case of referees' home bias. *Journal of Economic Psychology*, 25(4), 461–469. https://doi.org/10.1016/S0167-4870(03)00013-8

Torgler, B. (2004). The economics of the FIFA football World Cup. *Kyklos*, 57(2), 287–300. https://doi.org/10.1111/j.0023-5962.2004.00255.x

Van Damme, N., and Baert, S. (2019). Home advantage in European international soccer: Which dimension of distance matters? Economics Discussion Papers, No 2019-19, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2019-19

Van Den Broucke, L., and Baert, S. (2019). And at the end, the Germans always win, don't they? An evaluation of country-specific scoring behaviour in the dying seconds of international club soccer games. *PLOS ONE*, 14, e0202852. https://doi.org/10.1371/journal.pone.0202852

Watson, J.C., and Krantz, A.J. (2003). Home field advantage: New stadium construction and team performance in professional sports. *Perceptual and Motor Skills*, 97(3), 794–796. https://doi.org/10.2466/pms.2003.97.3.794