HOSTED BY

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Sport and Health Science 6 (2017) 321–326

www.jshs.org.cn

Original article

Modelling home advantage for individual teams in UEFA Champions League football

Chris Goumas

Sydney School of Public Health, The University of Sydney, Sydney, NSW 2006, Australia

Received 21 April 2015; revised 24 July 2015; accepted 8 September 2015
Available online 24 December 2015

Abstract

Background: Home advantage (HA) is well documented in a wide range of team sports including association football (soccer). Although much attention has been paid to differences in the overall magnitude of HA between football competitions and across time, few studies have investigated HA at the team level.

Methods: A novel method of estimating HA for individual teams, based solely on home performance, was used to compare HA between the highest performing teams and countries in the Union of European Football Associations (UEFA) Champions League over a 10-year period (2003/2004 to 2012/2013). Away disadvantage (AD) was also estimated based on each team's performance away from home. Poisson regression analysis was used to estimate covariate adjusted HA and AD in terms of the percentage of goals scored at home (HA) and conceded away from home (AD).

Results: When controlling for differences in team ability, HA did not vary significantly between the 13 selected teams. There was evidence ($p < 0.1$), however, of between-team variation in AD, ranging from 45% (away advantage) to 68% (away disadvantage). When teams were grouped into the 11 selected countries, both HA and AD varied significantly ($p < 0.02$) between countries: HA ranged from 52% for Turkish teams to 70% for English teams, while AD ranged from 52% (France) to 67% (Turkey).

Conclusion: Differences in style of play and tactical approaches to home and away matches may explain some of the variation in HA and AD between teams from different countries.

Keywords: Football; Home advantage; Modelling; Soccer

1. Introduction

Home advantage (HA) is the tendency for sporting teams to perform better at their home ground than away from home, and its existence has been well established in a wide range of team sports including association football.[1,2] Although much attention has been paid to differences in the overall magnitude of HA between football competitions[3–5] and across time,[2] relatively few studies have investigated HA at the individual team level, and this is the focus of the present study.

In a sporting competition where each team plays each other the same number of times both home and away, differences in team ability balance out over the season and therefore do not bias calculations of total HA across the whole competition.[6] HA for an individual team, however, will be largely determined by its ability relative to other teams in the competition; that is, a stronger team will be expected to win more matches at home than a weaker team. Team ability therefore needs to be controlled for when estimating HA for individual teams.

In the first comprehensive investigation into HA at the team level in football, Clarke and Norman[6] compiled 10 years of match data for 94 teams across 4 divisions of English football. HA for individual teams was calculated as a function of home and away goal difference and total HA across the whole division. Although this results in each team's HA being influenced by the HAs of all the other teams in the competition, the authors showed that this method controls for differences in team ability. A regression analysis found some evidence of variation in HA between teams, as well as for London teams having lower HA. There was strong evidence, however, for HA being higher for teams playing on artificial pitches, suggesting that home ground familiarity was playing a role.

Peer review under responsibility of Shanghai University of Sport.
E-mail address: goumas@tpg.com.au

http://dx.doi.org/10.1016/j.jshs.2015.12.008

Another method of estimating HA for individual sporting teams was developed by Pollard and Gomez,[7] originally for basketball, and subsequently applied to football teams in 4 countries in South-West Europe.[8] To control for team ability, the raw calculated HA—based on both home and away performance—for each team was regressed on a measure of that team's ability. The residual value for each team (i.e., the difference between its observed HA and the expected HA for a team of that ability) was then added or subtracted from the total HA for all teams in the competition. Like the Clarke and Norman[6] method described above, this approach produces HA estimates which are influenced by the HAs of other teams. Highly significant differences in HA between teams were observed in France, Italy, and Portugal, although there was little evidence of variation in HA between Spanish teams. Teams from ethnically and/or culturally distinct locations in France and Italy had greater HA than the rest of the teams in those countries, suggesting that territoriality and/or travel factors may be playing a role. However, such regional effects were not observed for teams in Portugal or Spain. Teams from capital cities in each country except Italy had significantly lower HA than teams from other areas.

The Pollard and Gomez[7] method of estimating HA for individual teams has also been used in studies of Brazilian and Greek football. In the First Division of the Brazilian football league, significant variation in HA between teams was observed.[9] In particular, teams in the north and south of Brazil had significantly higher HA than those from the central region; effects of travel and change in climate were suggested as possible explanations. Significant between-team variation in HA was also observed in the Greek Superleague,[10] with teams based in Athens showing less HA than those in the rest of Greece; the authors suggested that the sense of territorial protection may be less for teams in capital cities.

The present study introduces a novel method of estimating HA for individual teams, based solely on home performance and hence independent of the HAs of other teams in the competition. Away disadvantage (AD) is also estimated, based on a team's performance away from home. Multivariate regression techniques are used to control for team ability. Ten years of match data from the Union of European Football Associations (UEFA) Champions League—an international club competition featuring the best teams from over 25 European countries—were used to estimate HA and AD for a selection of the highest performing teams and countries over this period. This is the first study to investigate HA for individual teams in the UEFA Champions League.

2. Materials and methods

2.1. Data

The data used in this study were matches from the 2003/2004 to 2012/2013 seasons of the UEFA Champions League. Entry into this competition is based on a team's performance in their domestic league the previous season. The Champions League consists of a round-robin group stage, followed by a knock-out finals stage. All matches except the final are played in pairs, 1 at each team's home ground. Each of the group stage matches is decided on its own, whereas matches in the finals stage (except the final) are decided over 2 "legs". The final match in each season was excluded from the analysis as it was played at a neutral venue where there is no HA to be gained. Matches between AC Milan and Inter Milan ($n = 2$) were also excluded as these teams share the same home ground. Teams playing at least 50 matches over the 10-year study period (13 teams; 1058 matches) and countries whose teams played at least 75 matches between them (11 countries; 2028 matches) were chosen for the analysis. This selection method maximises the statistical power of the analysis, and results in the highest performing teams and countries being chosen. All match data were downloaded from the official UEFA website (www.uefa.com).

2.2. Analysis

HA for each individual team was estimated as the percentage of goals scored in home matches by that team. For example, if a team scored 50 goals in their home matches and conceded 30, then their unadjusted HA would be $50/(50 + 30) \times 100\% = 62.5\%$. Correspondingly, AD for each team was estimated as the percentage of goals conceded in away matches. HA greater than 50% represents superior performance in home matches, whereas AD greater than 50% represents inferior performance in away matches. Since crude calculations of HA and AD are influenced by differences in team ability, multivariate regression analysis was used to control for its confounding effect.

To model HA a paired design was used whereby each match contributed 2 observations, 1 for the home team and 1 for the away team. A repeated measures regression analysis using log-link Generalised Estimating Equations[11] in STATA Version 11 (STATA Corp., College Station, TX, USA) was used to estimate the mean number of goals scored by home and away teams. Repeated measures analysis is used when observations occur in pairs or groups and the outcome of interest is likely to be correlated within each group. In the present study the "groups" were the individual matches and the "observations" were the number of goals scored by each of the 2 opposing teams. As this outcome is a discrete count Poisson errors were specified for the regression model. Robust estimation of variance was used, as this produces valid standard errors even if the within-group correlations deviate from the correlation structure specified in the model.[12] An additional advantage of robust variance is that it prevents under-estimation of standard errors when count data are over-dispersed. This modelling strategy has previously been used to investigate HA in terms of goals scored[13] and disciplinary sanctions issued by referees[14] in football, and is described in greater detail by Goumas.[15]

To determine places and seedings in its club competitions, UEFA allocates points to each European football team based on previous performance in these competitions (www.uefa.com/memberassociations/uefarankings). To control for difference in home and away team ability a linear term for the number of points allocated to each team in each season of the Champions League was added to the regression model described above. Departure from linearity was tested for using quadratic and