Influence of match location in the spanish Copa del Rey
Efecto localización en la Copa del Rey del fútbol español

*Miguel Pic Aguilar, **Julen Castellano Paulis
*Universidad de La Laguna (España), **Universidad del País Vasco (España)

Abstract. This research studies second leg home advantage in football. Spanish Copa del Rey football eighth-finals (round of 16), quarter-finals (round of 8) and semi-finals (round of 4) from 1940-41 to 2013-14 have been analyzed, adding up to 2056 matches and 1028 play-offs. A non-parametric binomial contrast was applied to determine the influence of playing at home, while a binary logistic regression was used to show the explanatory proportion of goals after winning the qualifying round following the first confrontation. The results confirmed significant differences between playing as local or visitor (location) in the first or second match of the tie. Learning to manage the goals margin according to location with possible trips prior to the match (Clarke & Norman, 1995) and tactical-strategic familiarity with the place (Barnett & Hilditch, 1993), referee bias (Boyko, Nicolás & García, 2001), territoriality (Neave & Wolfson, 2003), psychological (Sánchez, Gónzalez, Ruiz, San-Juan, Abando, de Pablo, & Tena, 2015; Pollard & Pollard, 2005) the way teams face competitions (Page & Page, 2007), home and away teams approach games differently. Traditionally, regular models of football competitions have been investigated (Pollard & Gómez, 2014). That is, during a season, teams play once, both home and away, against the rest of the teams that make up the league, obtaining two results, a system which is more independent than using a two-legged tie format. Therefore, using a competition model composed by knockouts (Page & Page, 2007) gives information about the asymmetric character of the competition. The fact that an extended time or penalties are only necessary in the second match, that the teams have time to prepare strategies facing a hypothetical adverse result in the first leg, that the first match conditions the tie against the crucial importance of the second match, or local teams mobilizing fans (Agnew & Carron, 1994) in the face of transcendental matches, could make us think that playing the second game locally would be advantageous for the local team.

However, on the other hand, the away goals rule would favour the visiting team and would determine the teams’ strategic approach. Except for a few contributions, two-legged ties remain a research challenge. In the current Champions League championship, period 1994-95 / 2009-10, minor trends in the round of 32 have been pointed out, confirming HA for local teams in the second match (Eugster, Gertheiss, & Kaiser, 2010). The HA was also subsequently endorsed (Pic & Castellano, 2016) in the quarter-finals of the same tournament. The authors of the previous study, analyzed a sample adding up to 642 matches and 336 play-offs. A non-parametric binomial contrast was applied and, even though no statistical evidences were found, it was pointed out that there are differences between knockout quarterfinals and semifinals. Specifically, HA could become reversed, at least in the semifinals UEFA Champions League.

The way in which the number of goals can affect (Flores, Forrest, de Pablo, & Tena, 2015; Pollard & Pollard, 2005) the way teams face the second leg, or the insignificance of the regularity of the teams in a competition format to overcome play-offs, both contribute to characterize this competition format. Two-legged ties may constitute specific confrontation models. Therefore, perhaps the values of HA found in football (Pollard, 1986) can be modified by researching the ‘knockout’ or ‘match’ units. Researchers Page and Page (2007) analyzed a total of 6182 qualifying rounds, from 1955 to 2006, concluding that playing the second leg at home favoured local teams. In a similar approach (Lidor, Bar-El, Arnon, & Bar-El, 2010) the existence of HA in European...
qualifying rounds was confirmed. Recently (Flores et al., 2015), it has been tried to model the goal in relation to the quality of the teams in European competitions over two-legged ties. The importance of the number of goals needed to progress to the next round did not offer a static behaviour, but a changeable one. It had already been investigated in the football World Cups (Castellano, Perea, & Hernández-Mendo, 2008; Castellano, 2009).

Based on the foregoing, the present study aims to deal with a dual purpose. On the one hand, 1) to determine the effect of playing at home the first or second match of the Copa del Rey two-legged ties. On the other hand, 2) to reveal if the goals margin scored in the first match constitutes a significant advantage regarding the probability of overcoming the round.

**Method**

**Participants**

During the season 1940-41, the two-legged ties classification format for rounds of eighth-finals, quarter-finals and semifinals was established, and it has remained unchanged until today. Once this first requirement of organizational stability was achieved, the second criterion was the existence of a two-legged tie. Direct classification of some teams to the quarter-finals during the seasons 1950-51, 1951-52, 1953-54 and 1954-55 reduced the number of rounds of 16. The sample composition is already a limitation that does not serve a balanced number (n) in the qualifying rounds. However, the study considered 1028 qualifying knockouts for the Copa del Rey (2056 matches), grouped in 584 round of 16 knockout (1168 matches), 296 quarter-final knockouts (392 matches) and 148 semifinal knockouts (286 matches). The penalty goals scored by the teams after finishing the second match were excluded.

**Variables**

The location (home and away) and tie (eighth-finals, quarter-finals and semifinals) variables were taken. The dependent variable success was associated to overcoming the tie. Both as a predictor and to reveal the teams’ odds of passing qualifying rounds, the goal margin was included; that is the goal difference between home and away teams, referring only to the first match of the tie.

**Procedure**

The data were downloaded from the official website of the Spanish Football Federation; [http://www.rfef.es/](http://www.rfef.es/) and [www.linguasport.com/](http://www.linguasport.com/). To calculate the success rate, teams that played the first match of the tie away were selected, together with the proportion of teams which classified to the next round by playing the first match at home. Each match was labelled as an independent unit, due to the great variability of organizational stability was achieved, the second criterion was the existence of a two-legged tie. Direct classification of some teams to the quarter-finals during the seasons 1950-51, 1951-52, 1953-54 and 1954-55 reduced the number of rounds of 16. The sample composition is already a limitation that does not serve a balanced number (n) in the qualifying rounds. However, the study considered 1028 qualifying knockouts for the Copa del Rey (2056 matches), grouped in 584 round of 16 knockout (1168 matches), 296 quarter-final knockouts (392 matches) and 148 semifinal knockouts (286 matches). The penalty goals scored by the teams after finishing the second match were excluded.

A binary logistic regression was applied (Gómez, Lorenzo, Ibáñez, & Sampaio, 2013; Vinson, Padley, Croad, Jeffreys, Brady, & James, 2013; Pic & Castellano, 2016). When the first match was over, the goal margin was included, both for locals and visitors, with the winning teams taking positive values and with negative values for losing teams.

**Data analysis**

Statistical package SPSS v.18 for Window was used, and a binomial non-parametric contrast was applied. Through hypothesis testing, and taking as null hypothesis the sample proportion being equal to 50% (\( \Pi = 0.5 \)), while it was considered as an alternative hypothesis the sample proportion being smaller than 50% (\( \Pi \) different to .5). For the significance level (\( \alpha = 0.05 \)), if (\( p < .05 \)) the alternative hypothesis is accepted (the test is significant), and if (\( p > .05 \)) the null hypothesis would not be rejected (the test is not significant).

In order to find the influence of the goal margin on the opponent team in the first leg, taken as an independent variable or covariable, a binary logistic regression was applied to the probability of passing the tie taken as a dependent variable. While the value 0 indicated that the teams had tied, positive values appointed the winning team, with negative values indicating the losing team.

From the obtained models, the influence of goals margin to overcome the knockout was estimated. Two logistic regressions were applied, one for local teams and one for visitors. This distinction will report on the margin goal profitability for home and away teams during the first match.

The following coding is used for the dependent variable (y): 1) It does not pass the knockout: internal value 0, and 2) It passes the knockout: internal value 1. The resulting regression model took the chance of passing the knockout depending on covariate (x). The model equation was: \( P(y = 1) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 * x)} \).

The specificity and sensitivity of the models are defined as: 1) Sensitivity: ability of the model to predict that a team passes the round; and 2) Specificity: adequate predictive capacity of the model for not getting through the tie. To establish the constants and coefficient model significance, Wald statistic was used. For all cases, the cutoff was set at 0.5.

Before addressing the regression model, the Mann-Whitney U test was performed to prove the relationship between the covariate and the dependent variable. Through Hosmer and Lemeshow test the adequacy of the data to the model was checked. The goal margin between teams after the first game was also analyzed. From this analysis, the figures of the teams depending on location can be compared.

**Results**

**Home advantage**

Table 1 shows the sample size (n = total number of teams that played the first match of the round away), the sample proportion unit (number of these teams which passed the round divided by n), and the p-bilateral value or significance of the test.

After analyzing the 1028 playoffs with a unit sample proportion (\( pm = 0.55 \)), levels of significance (\( p < .002 \)) were obtained. The data significantly confirmed that playing the second leg at home was profitable for the teams. These results are shown in detail in Table 1, taking into account each knockout.

<table>
<thead>
<tr>
<th>Play Off</th>
<th>Period</th>
<th>n</th>
<th>pm</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1941-49</td>
<td>72</td>
<td>0.53</td>
<td>0.724</td>
</tr>
<tr>
<td></td>
<td>1950-59</td>
<td>72</td>
<td>0.56</td>
<td>0.410</td>
</tr>
<tr>
<td></td>
<td>1960-69</td>
<td>80</td>
<td>0.31</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td>1970-79</td>
<td>80</td>
<td>0.56</td>
<td>0.314</td>
</tr>
<tr>
<td></td>
<td>1980-89</td>
<td>80</td>
<td>0.56</td>
<td>0.314</td>
</tr>
<tr>
<td></td>
<td>1990-99</td>
<td>80</td>
<td>0.65</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>2000-09</td>
<td>80</td>
<td>0.64</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>2010-14</td>
<td>40</td>
<td>0.60</td>
<td>0.268</td>
</tr>
</tbody>
</table>

Global 584 0.58 0.000

<table>
<thead>
<tr>
<th>Play Off</th>
<th>Period</th>
<th>n</th>
<th>pm</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1941-49</td>
<td>36</td>
<td>0.90</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1950-59</td>
<td>40</td>
<td>0.48</td>
<td>0.875</td>
</tr>
<tr>
<td></td>
<td>1960-69</td>
<td>40</td>
<td>0.60</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>1970-79</td>
<td>40</td>
<td>0.58</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>1980-89</td>
<td>40</td>
<td>0.48</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>1990-99</td>
<td>40</td>
<td>0.58</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>2000-09</td>
<td>40</td>
<td>0.43</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>2010-14</td>
<td>20</td>
<td>0.40</td>
<td>0.583</td>
</tr>
</tbody>
</table>

Global 256 0.54 0.222

<table>
<thead>
<tr>
<th>Play Off</th>
<th>Period</th>
<th>n</th>
<th>pm</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1941-49</td>
<td>18</td>
<td>0.44</td>
<td>0.815</td>
</tr>
<tr>
<td></td>
<td>1950-59</td>
<td>20</td>
<td>0.50</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1960-69</td>
<td>20</td>
<td>0.65</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>1970-79</td>
<td>20</td>
<td>0.50</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td>1980-89</td>
<td>20</td>
<td>0.45</td>
<td>0.824</td>
</tr>
<tr>
<td></td>
<td>1990-99</td>
<td>20</td>
<td>0.50</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>2000-09</td>
<td>20</td>
<td>0.50</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>2010-14</td>
<td>10</td>
<td>0.30</td>
<td>0.344</td>
</tr>
</tbody>
</table>

Global 158 0.57 0.002

Global 1,028 0.55 0.002

Statistical significance decreased when the variable playoff was included. In the last 16 rounds (n = 584; pm = 0.58, p = .000),
quarterfinals (n = 296; pm = 0.54, p = .222) and semifinals (n = 148; pm = 0.47, p = .460). In the rounds of last 16, significance came from the decades 1990–99 (p = .010) and 2000–09 (p = .018), noting that it was more likely to get through the tie when playing the second game at home. Although there was no significant evidence during the period 2010–14 (p = .268), a trend favoring the classification of teams in the same sense was observed. This trend which appeared in the last period was supported by what had been noted in the two previous decades (1990–99/2000–09). In the round of last 16, the classification of local teams during the second match was more likely, with a global significance (n = 584; pm = 0.58, p = .000).

In the quarterfinals, the values indicated significance in the decade from 1980 to 1989 (n = 40; pm = 0.68, p = 38) through the relation between playing the second game at home and winning the tie. Greater proportions were found, although not significant, of local teams which in the second match progressed to semifinals 1960–1969 (p = .268), 1970–1979 (p = .430), 1980–1989 (p = .38), 1990–99 (p = .430). In contrast, during the decade 1950–1959 (p = .875), 2000–09 (p = .430) and the period 2010–14 (p = .503), the proportion of teams that passed the quarterfinal knockout when visitors in the second game was higher. Although the referred proportions did not rest on statistical significance, they reported on the proportion of teams which continued in the competition (p = .222).

Knockout goals

To analyse the importance of the goal margin, the sample analysis was filtered according to the round. Significance in the relationship between the covariate and the dependent variable was fulfilled. The results brought by Mann–Whitney U test were the following ones: in the round of 16, at home (z = -14.592, p < .001) and away (z = -14.518, p < .001); quarterfinals, at home (z = -10.916, p < .001) and away (z = -10.675, p < .001); in the semifinals, at home (z = -8.225, p < .001) and away (z = -8.225, p < .001). The non significance of the Hosmer and Lemeshow test in the round of 16 at home (p = .084) and as visitor (p = .111), quarterfinals as local (p = .856) and as visitor (p = .896), in semifinals as local was (p = .157) and as visitor (p = .157); so the null hypothesis which shows that the model was well adjusted is not rejected. The model has proven to be reliable.

Table 2 shows that the B coefficient for the covariate was significant and the confidence interval for the Exp(B) did not contain the unit. Nagelkerke’s R² reports on the explained variance. The covariate had a significant influence on the occurrence of the event. Although not shown in table 2, the constants were also significant.

The corresponding models for each knockout are shown according to the condition of playing at home or as visitor (Table 3). In the same table it can be seen in which way the goal margin after the first match was decisive to overcome the tie.

The interpretation of Exp(B), odds ratio, for the rounds of 16 or eighth-finals:

- For each advantage goal scored by the team at home during the first match, the probability of passing the round increases 3.021 times (302 %).
- For each advantage goal scored by the team away during the first match, the probability of passing the round increases 2.991 times (299 %).
- The interpretation of Exp(B), odds ratio, for the semifinals:
  - For each advantage goal scored by the team at home during the first match, the probability of passing the round increases 3.262 times (326 %).
  - For each advantage goal scored by the team away during the first match, the probability of passing the round increases 3.038 times (303 %).
- The interpretation of Exp(B), odds ratio, for the semifinals:
  - For each advantage goal scored by the team at home during the first match, the probability of passing the round increases 3.410 times (341 %).
  - For each advantage goal scored by the team away during the first match, the probability of passing the round increases 3.410 times (341 %).

Teams at home obtained identical values in relation with the goal margin to reach the final. A comparative description of the three knockouts showed the irregular margin goals depending on the knockout and location of the match, with the exception of semifinals, where there was no appreciable difference. Thus, an intra-knockout description showed that the goal margin was almost non-existent for the rounds of 16 and semifinals, with 23 percentage points in the quarterfinals.

Without taking location into account, the goal value increased when passing knockouts from rounds of 16 (300 %), quarterfinals (315 %) and semifinals (341 %). The knockout with bigger percentages offered less probabilities of modifying the result after the first match.

Discussion

The aim of this paper was analysing the overcoming of two-legged ties by Spanish football teams in the Copa del Rey, and its relationship with the HA. The effect of playing the second match at home, together with the goal margin during the first match, were the main goals of this research. The main result of the study showed statistically significant differences on home advantage when facing the second game (or the first as a visitor) depending on the tie. In addition, the predictive properties

![Figure 1. Temporal evolution (by decades) of the proportion of teams which classified as visitors during the first match, and as local during the second match, for the round of last-eight-finals, quarter-finals and semifinals. Through binary logistic regression (p = .05).](image-url)
of the goal margin after the first match were pointed out. The temporarily irregular behaviour of the variables in the study showed the factorial complexity of the studied phenomenon (Pollard, 2008; Pollard & Pollard, 2005).

The data analysed in the present study confirmed that playing the second game at home, location effect, increased the odds of overcoming the tie. These effects (Page & Page, 2007), were unequally distributed in the different knockouts. It would be advisable to exercise extreme caution and not to generalize the effect of HA to all qualifying rounds in the study because it did not respond to the researchers’ conclusion, except when referring to the round of 16. In this knockout, the significance found during 1990-99 (p = 0.010) and 2000-09 (p = 0.010), without forgetting the trend in the same way between 2000 and 2014, would help to highlight the importance of location effect, intensifying its effects since the nineties until nowadays. Despite this, the goal margin between home and away teams in the round of 16 (first match) was similar, which might give greater importance to the location effect against the relative importance of the goal in a competition format based on two-legged ties.

Regarding the quarterfinals and semifinals, there are slight statistical trends estimated in European competitions, and it is not possible to confirm the existence of HA due to the fact that the second game is played at home (Eugster et al., 2010; Lidor et al., 2010; Pic & Castellano, 2016). Despite this, and taking into account that we are discussing statistical trends, both studies agreed to point out the reverse HA for semifinals of the UEFA Champions League, as previously noted by other researchers (Baumeyer & Steinhilber, 1984; Wright & Vogel, 1995).

The results have supported the existence of HA according to the number of goals scored by home and away teams (p < 0.01) in any qualifying round (Table 2). However, in a competition system with two-legged ties, goal margin would have a relative importance, when compared with the existence of HA in the rounds of last 16, as it has been pointed out.

The changeable value of the goals depending on the knockout (Flores et al., 2015) was proved. The goal margin for local teams (first match) in the rounds of last 16 and semifinals was scarce. However, while a higher proportion of teams in the round of 16 reached the next round as local during the second match, in contrast semifinals pointed in the opposite direction, or reverse HA (home advantage away in the second match of the tie). Therefore, the complexity of a phenomenon with many factors, interactions (Pollard & Pollard, 2005) and interpretations was clearly stated.

Considering the goals margin in Table 3, if a local team wins the first match of the quarterfinals with a one goal margin, the chances to move to the next round would triple (3.0 times). However, getting the same result when playing away would slightly reduce their chances of progressing in the rankings. In quarterfinals, if a team got a score of 2-0 as the home team, they would increase their chances to reach the final (6.5 times), while in an identical situation, a visiting team would reduce their chances to (6.0 times). In this sense, the home-scored goals margin was more profitable to pass the eliminator than the visiting team’s one. The location-based goals margin profit was identical in the semifinals, while in recent studies (Pic & Castellano, 2016) the value of the visiting team goals exceeded the value of local goals in the same semifinals round. The standard double value of goals in case of a tie could explain part of these results because teams could strategically plan the second leg, taking strategic risks (Pollard, 2008) in order to rectify adverse outcomes.

The role of double value of away goals in case of a tie, was invariably introduced during the 93-94 season, but it made an appearance for the first time in the 70-71 season, although it was removed later. However, such legislation was applied to almost half of the seasons studied. Even though a decrease in HA during the second game was noted (Figure 1), statistical significance was nonexistent in the quarterfinals and semifinals from the 90s, which matched previous research (Pic & Castellano, 2016). The fact that the statistical significance rested on the second round, supporting HA during the second game, points to the need to further define an interaction of factors model, including ties.

Among the limitations of this study, the low equity in the distribution of (n) by knockout stands out. Increasing the sample or the inclusion of the result of the second match would offer a more complete understanding of the competition through double matches. Other limitation would be the necessity of going beyond the goals margin concept, since a 1-0 score in the first match is not equal to a 4-3 score. However, it should not be forgotten that a substantial (r) is required in order not to jeopardize the results. Taking into account the quality of the teams, and considering including “regulatory changes” covariables such as the number of players changed per match or offside regulation, present themselves as challenges for future research. Stimulating the discussion about the controversial rule of the value of goals scored in the rival field could constitute a challenge for future research.

The specificity of this competition model (Veroz, Yagüe, & Taberner, 2015) directly affects the coach and the preparation of routines with possible consequences on the model or style of play. While the most well-known model is the regular season format, all against all, we must put the focus on what specific modifications could be changed in order to get the best preparation for high competition football teams. Being able to move to the actual training situation concepts such as the limited relevance of victory or the relative importance of the goal, are tasks that should not go unnoticed in the preparation of players, if a team competes in two-legged ties.

Perhaps this contribution justifies the strategic thinking of football (Pollard, 2008) coaches, especially in the Copa del Rey round of last 16, taking into account that the approaches which have been put into practice seem to have statistically benefited the home team during the second match.

Conclusions

The location effect was irregular in the studied knockouts. In the round of last 16, the team that was significantly favoured to reach the next tie was the home team (second game) which had a narrow goal margin. In quarterfinals the trends pointed in the same direction as the ones found in the round of last 16, but with a bigger goal margin. However, the goal margin disappeared in semifinals, and a reverse HA trend was identified. So, it was confirmed that the visiting team was favoured during the second match.

Acknowledgments

We gratefully acknowledge the support of two Spanish government projects (Ministerio de Economía y Competitividad): 1) La actividad física y el deporte como potenciadores del estilo de vida saludable: Evaluación del comportamiento deportivo desde metodologías no intrusivas [Grant number DEPI2015-66069-P, MINECO/FEDER, UE]; 2) Avances metodológicos y tecnológicos en el estudio observacional del comportamiento deportivo [PSI2015-71947-REDP, MINECO/FEDER, UE].

References


Retos, número 31, 2017 (1º semestre) - 205 -


