FAVORITISM AND REFEREE BIAS IN EUROPEAN SOCCER: EVIDENCE FROM THE SPANISH LEAGUE AND THE UEFA CHAMPIONS LEAGUE

BABATUNDE BURAIMO, ROB SIMMONS and MAREK MACIASZCZYK

In this paper, we test for, and find evidence of, referee bias in favor of home teams in European football using minute-by-minute analysis to control for within-game events. The context for the analysis is Spain's Primera Liga and the Union of European Football Association (UEFA) Champions League. We find that the award of sanctions by Spanish referees in the Champions League are not significantly different to those of the referees from other countries and as such are subject to the same sources of bias. In Primera Liga matches where the crowd is separated from the pitch by running tracks, we find that the probability of the award of a yellow card to the home team is higher and that of the away team is lower compared to matches played at stadia without running tracks. Similar results are found in the Champions League, where efforts are made to hire "neutral" referees. Referee behavior is also influenced by the size of the crowd in attendance. (JEL D8, J2)

I. INTRODUCTION

A growing literature has highlighted the possibility of favorable treatment of some agents offered by principals. Such behavior can lead to inefficiency in principal-agent relationships and hence inefficiency in particular labor markets. If agents' performances in workplaces are subjectively assessed then key decisions on worker careers, such as promotions, may be biased in such a way that the most productive workers may not be optimally allocated to their best jobs and assignments within organizations. This problem has been analyzed theoretically by Prendergast (1999).

Although the theory of favoritism in principal-agent relationships is well established, testing for both the presence of favoritism and for unintended adverse effects on organizational performance is very difficult in most labor markets as the required data are simply not available. Questionnaires are unlikely to elicit honest and reliable responses to questions about favoritism. Assessing worker performance in many organizations, particularly in the service sector, is fraught with difficulties because of complexities of team production.

Hence, many researchers have turned to professional team sports as a vehicle for analyzing favoritism and its effects on organizations. Team sports offer precise, detailed, and publicly available performance data for both players and teams. In this paper, we investigate favoritism as applied by soccer referees in response to social pressure applied by crowds inside stadia. Soccer authorities are aware of the possibilities of bias when soccer referees make judgements during matches and these referees are trained to avoid such bias. In one of the competitions that we shall analyze, the pan-European Union of European Football Association (UEFA) Champions League, the organizers appoint neutral match officials from outside the countries represented by the teams in any match. We shall show that referee bias persists even in this case. In our second case, the Spanish Primera Liga, we also find evidence of referee bias. This case comprises teams that vary considerably in size

Buraimo: School of Sport, Tourism and the Outdoors, University of Central Lancashire, Preston PR1 2HE, UK. Phone +44-1772-894916, E-mail baburaimo@uclan.ac.uk

Simmons: Department of Economics, Lancaster University Management School, Lancaster University, Lancaster LA1 4YX, UK.

Maciaszczyk: School of Sport, Tourism and the Outdoors, University of Central Lancashire, Preston PR1 2HE, UK.

and also have to travel long distances to games in a large country.

We aim to assess bias by referees in the form of awards of sanctions against players who are penalized for infractions against the rules of professional soccer. These rules are set by the world governing body of soccer, Fédération Internationale de Football Association (FIFA), and not by competition organizers. The role of soccer referees is to ensure that players perform within the rules of soccer and, when rules are broken, that players are sanctioned appropriately. Sanctions vary from free kicks against the offending team to yellow cards ("cautions") to the ultimate sanction of a red card, denoting expulsion from the game without replacement. Two yellow cards awarded to a player in a match results in automatic expulsion and so can be seen as equivalent to a red card.

In team sports worldwide, referees and umpires are employed to make critical judgements about player behavior that could have consequences for game outcomes. In baseball, umpires decide whether a pitcher's ball is in a "strike zone" defined as an area within which a batter has the potential to hit the ball. In basketball, referees assess whether fouls are committed. In each of these sports, investigators have detected racial bias. In Major League Baseball, pitchers with same race or ethnicity as the umpires give up fewer hits, strike out more batters, and improve their teams' probability of winning (Parsons et al. 2011). In the National Basketball Association, it appears that more personal fouls are called against players when the refereeing crew is opposite-race rather than own-race (Price and Wolfers 2010).

In Spanish soccer, Garicano, Palacios-Huerta, and Prendergast (2005) found that referees added more discretionary "injury time" at the end of normal time when home teams were behind in score compared to when home teams were ahead in score. This suggests a form of home team bias by referees. But although Rickman and Witt (2008) found a similar type of bias for English Premier League soccer, this bias disappeared with the onset of annual salary contracts for referees, which the authors term "professionalism." The wider implication is that bias can be removed by appropriate financial incentives.

Although the studies of "injury time bias" are persuasive, with credible econometric modeling, the ability of teams to score in time

added on at the end of a match would appear to be rather limited. An alternative form of bias can be sought in the form of sanctions (cards) given to players during matches. The award of a caution (yellow card) implies that a player's aggression, which could be physical or verbal, needs to be curbed, otherwise a second yellow card and a dismissal will follow. Ridder, Cramer, and Hopstaken (1994) have shown that teams reduced from 11 to 10 players because of a dismissal will tend to lose games played against teams with a full complement of players. Players and fans know this, of course. This is why players surround soccer referees, and fans shout, to encourage referees to award cards to opponents after fouls or other misdemeanors.

A growing literature has detected favorable treatment to home teams in the award of cards. Dawson et al. (2007) analyzed a large sample of English Premier League games over the period 1996–2003 and controlled for referee fixed effects and a number of covariates reflecting relative strength. The conclusion emerged that away teams received more cautions and sendings off than home teams, even after conditioning for home advantage effects on match results. In a sequel, Dawson and Dobson (2010) consider referee bias in the UEFA Champions League and the UEFA Cup. Again controlling for relative strengths and game characteristics they find that home teams receive less sanctions than away teams.

A potential problem with the studies by Dawson and co-authors is that they do not consider within-game dynamics, particularly the progress of score during the game and likely responses by teams. As Buraimo, Forrest, and Simmons (2010) show, teams that are behind in score, which are most likely to be away teams, tend to adopt more aggressive behavior that incurs more cautions. A complete analysis of referee bias should then control for events during the game that might cause referees to award more cards, without necessarily being associated with bias to any particular type of team. Buraimo, Forrest, and Simmons find evidence of bias toward home teams by referees in both English Premier League and German Bundesliga 1, after controlling for a number of pregame and within-game characteristics as well as referee fixed effects. In particular, referee bias can be identified by the removal of running tracks around stadia in Germany. When tracks were removed, or teams relocated to stadia without tracks, it

was found that home teams received fewer yellow and red cards than when tracks were in place. The authors interpret the removal of running tracks as leading to an increase in social pressure which might in turn affect referee behavior.

The present paper uses the same methodology of within-game analysis to consider two important soccer competitions: Spain's Primera Liga and UEFA's Champions League. The Spanish case is notable for its League structure where three large teams (Real Madrid, Barcelona, and Valencia) are dominant. This is a League in which some very small teams compete with some very large teams. Also, Spain has large distances between some locations meaning that the composition of attendance will be more skewed toward home fans than in England, where distances between teams are relatively small.

UEFA's Champions League is of interest because UEFA makes a conscious decision to appoint neutral match officials who are not from countries of competing teams in any given match. This is a deliberate attempt to combat potential bias. But this attempt to remove bias may still be unsuccessful if neutral referees are still susceptible to social pressure and if referee behavior is inconsistent across national identities (Dawson and Dobson 2010).

II. METHODOLOGY

To test for home advantage and refereeing bias, matches from Spain's Primera Liga and the UEFA Champions League from 2003–2004 to 2006–2007 are used. Referees in the Primera Liga are drawn from a list of match officials nominated by the Spanish Football Association. In contrast, UEFA deliberately ensures that referees and their assistants are taken from countries outside of those containing the contestants in a given Champions League match. That is, UEFA attempts to install "neutral" referees, precisely because of the risk that referees from a country that one or both teams belong to may be prone to implicit bias. If this hiring policy is fully successful, then we expect any referee bias observed in domestic competitions to disappear in the Champions League.

Most studies on referees' behavior have used match as the unit of observation (e.g., see Boyko, Boyko, and Boyko 2007; Dawson et al. 2007; Dawson and Dobson 2010; Scoppa, 2007). Assessing refereeing sanctions using match as the unit of observation is imprecise. Although some factors do not vary across individual matches, a significant proportion of sanctions are likely to depend on events that occur during the match itself. For this reason, analyses should take account of events that occur during the match. Buraimo, Forrest, and Simmons's (2010) study of refereeing bias in English and German football is a precursor in this respect. They use minute as the unit of observation and incorporate within-game information. Similarly, we assess the probability of a card being awarded to either team in any given minute using a bivariate probit model. The use of minute as the unit of observation is relatively new in the analysis of games (Buraimo, Forrest, and Simmons 2010; Corral, Rodriguez, and Simmons 2010) but a very appropriate unit of analysis. By using minute as the unit of observation, the precise order and effects of events occurring during a match can be taken into account. For example, two matches may report the same number of cards and goals; however, in one match, the goals may have preceded the cards and even contributed to the cards. In the other, the cards may have preceded the goals; however, using match as the unit of observation would not sufficiently deal with the order of events as they occur during a match. This means that each of the 90 minutes across 1,520 games in the Primera Liga form the basis of the analysis. This also applies in the case of the UEFA Champions League, however, for a small number of matches in this league, the matches have cause to extend beyond the normal 90 minutes of play and an additional 30 minutes are played. These are normally during the second leg of the knockout stage in which the outcome is limited to either a home or away win; draws are not permitted as one team must progress to the next stage.

A bivariate probit model fits maximum-likelihood for two probit models. The conjecture in using a bivariate probit model is that the error terms from the two probit estimates are correlated. In the case of independence, the two equations reduce to two univariate probit models. If the error terms are correlated, the use of a bivariate probit model produces efficient parameter estimates of the independent variables. For a more formal presentation of the model, see Greene (2008). The application of the bivariate probit model is appropriate on theoretical as well as empirical grounds. First, the award of sanctions to either team is not independent

of those awarded to or events involving the opposing team. Furthermore, the null hypothesis that the correlation between the two error terms is 0 is rejected in all cases. Hence the probabilities of sanctions should be jointly modeled (see Buraimo, Forrest, and Simmons 2010; Dawson et al. 2007; Dawson and Dobson 2010).

For both competitions, the dependent variables are pairs of binary indicators. The first pair represents the award of yellow cards to home and away teams in any given minute of the match. These variables take the value of 1 if a card is awarded during that minute or 0 otherwise. A second pair of variables is constructed but this time for dismissals. For dismissals, no distinction is made between the award of a second yellow card and a straight red. Preliminary analysis showed that there was very little distinction between models of second yellow and straight red cards. Furthermore, combining second yellow and red cards is justifiable given that they result in the same outcome, a player being sent off.

A. Pregame Variables

To model the dependent variables, two groups of variables are noted: *pregame* and *withingame*. *Pregame* information is based on data available prior to the start of the game and remains constant throughout the match. The first of these is *track*. In Spain, it is common for games to be played on pitches surrounded by running tracks. Dohmen's (2008) analysis of German football found that the presence of a running track influenced the amount of extra time that referees allocated at the end of the second half. *Track* is a binary variable that takes the value of 1 if a track is present and 0 otherwise.

As noted in the previous section, crowd pressure is a significant part of the literature. The conjecture is that the greater the crowd size, the more likely decisions will favor home teams. The extent of favoritism shown to the home team is likely to depend on crowd composition. The crowd may not be entirely made up of home team fans and a significant portion may be neutral or fans of the opposing team. In the absence of information on crowd composition, crowd size is a reasonable proxy for home support. To capture the influence of support and crowd pressure, *log attendance*, the log of match attendance, is included as an independent variable.

Matches that involve local rivals (e.g., Real Madrid and Athletico Madrid) or are of historical significance (e.g., Real Madrid and FC Barcelona) may induce greater levels of foul play. These matches carry greater importance beyond league standings as rival fans seek to gain bragging rights over their counterparts. Consequently, additional pressure exerted by fans may induce extra effort on the part of players which may spill over into foul play punishable by a caution or a dismissal. A dummy variable *derby* takes the value of 1 if the match involves such rivalry and 0 otherwise. For the Champions League, *derby* takes the value of 1 if both teams are from the same country, otherwise its value is set to 0.

The award of sanctions will vary depending on the relative qualities of the teams. If one team is significantly superior to its rival, the incentive for its players to engage in foul play is reduced as the team is expected to outplay its opponent. The less favored team, however, stands to benefit from foul play as it looks to stifle the opponent. To capture team quality, we use odds taken from the betting market. The abolition of tax on betting and increased competition in the betting market mean that bookmakers have to set odds that are both competitive and an accurate reflection of the likely outcome of events (see Forrest, Goddard, and Simmons 2005). Using odds data from William Hill, the probabilities of a home win, away win, and draw are computed by taking the reciprocal of the published returns to a unit stake. The sum of the three probabilities normally exceeds 1 because of the inbuilt over-round reflecting the bookmaker's profits. The three probabilities are adjusted by dividing by the over-round so that they sum to 1. To capture the expected closeness of the game, the difference in home and away win probabilities (*difference in bookmaker probability*), along with its square are included. As the difference increases emphasizing the home team's superior quality, the likelihood of the home team being awarded a card is expected to reduce. For the away team, the likelihood of a card is expected to increase at a decreasing rate. One would expect that as the home team's quality increases relative to the away team's quality, the benefits that accrue to the away team from foul play diminish.

As the format of the Champions League is different to domestic leagues, additional control variables are included to reflect this. The reference stage is the last 16 and dummy variables

for the main group stage, as well as the quarter- and semifinals are included and denoted by *group stage*, *quarterfinal*, and *semifinal*, respectively. The stages of the tournament are likely to reflect the propensity for teams to engage in foul play punishable by a sanction. The finals of the Champions League are omitted as these are played at neutral stadia. Additionally, dummy variables are also included to capture group stage matches in which either team has already qualified for the knockout stages (*home team qualified from group stage* and *away team qualified from group stage*). Players of such teams are less likely to engage in behavior that will result in a sanction. Finally, dummy variables for home teams, away teams, referees, and seasons are included as control variables.

B. Within-Game Variables

The main characteristic of the within-game variables is that they vary throughout the game. *Minute* and its square are included as control variables. We hypothesize that the probability of receiving a card will increase with *minute* but at a decreasing rate. In addition, dummy variables *45th minute* and *90th minute* are included and take the value of 1 if the match is in the 45th or 90th minute. Our sources of information on times does not provide details of extra time allocated and all events occurring in injury time are denoted as having taken place in the last minute of each half. The 45th and 90th are therefore *longer* minutes.

Goal difference, the difference in home and away team goals prior to current minute, along with its squared term, is included to capture the current state of the game. The behaviors of players and referees are likely to be influenced by this. As the difference in home and away goals increases, the incentives and benefits from committing fouls are reduced.

The probability of a card being issued to either team is likely to depend on the number of cards awarded so far. A pair of variables, *home yellow last 3 min* and *away yellow last 3 min*, reflects the number of yellow cards awarded recently, in the last 3 minutes prior to the current minute. These variables, depending on the sign and significance of their coefficients may capture an even-out effect in which referees, having recently sanctioned one team, feel inclined to sanction the other. Alternatively, it may also capture retaliation in which players respond to being fouled by the opposition by committing fouls themselves soon after. Complementing these are *home yellow prior* and *away yellow prior* which are the total number of yellow cards awarded prior to the last 3 minutes to the home and away teams, respectively.

The last pair of variables is *home second yellow and red* and *away second yellow and red* and are the total number of second yellow and red cards received by the home and away teams, respectively, prior to the minute under consideration. Tables 1–4 show the summary statistics for the award of cards across the two competitions along with those of the independent variables.

Our bivariate probit model for yellow cards in the Primera Liga is:

Pr*(home yellow card, away yellow card) = f (track, log attendance, derby, difference in bookmaker probability, difference in bookmaker probability squared, minute, minute squared, 45th minute, 90th minute, goal difference, goal difference squared, home yellow last 3 min, away yellow last 3 min, home yellow prior, away yellow prior, home second yellow and red, away second yellow and red).*

For the Champions League, the variables: *group stage, quarterfinal, semifinal, home team qualified from group stage, away team qualified from group stage* are included in the independent

TABLE 1
Distribution of Cards Per Game by Season in the Primera Liga

Season	Home Yellow	Away Yellow	Home Second Yellow	Away Second Yellow	Home Red	Away Red	Number of Games
2003–2004	2.53	2.84	0.08	0.11	0.09	0.09	380
2004–2005	2.29	2.77	0.06	0.10	0.06	0.07	380
2005–2006	2.48	2.85	0.07	0.12	0.10	0.09	380
2006–2007	2.46	3.05	0.03	0.04	0.11	0.12	380
Total	2.45	2.89	0.06	0.09	0.09	0.09	1,520

TABLE 2

Distribution of Cards Per Game by Season in the Champions League

Season	Home Yellow	Away Yellow	Home Second Yellow	Away Second Yellow	Home Red	Away Red	Number of Games
2003–2004	1.17	1.90	0.06	0.09	0.04	0.02	125
2004–2005	1.45	1.72	0.02	0.07	0.02	0.05	125
2005–2006	1.29	2.02	0.03	0.12	0.03	0.03	125
2006–2007	1.53	1.98	0.02	0.07	0.02	0.06	125
Total	1.35	1.90	0.03	0.09	0.03	0.04	500

TABLE 3

Summary Statistics for Primera Liga ($n = 135,720$)

Variable	Mean	Standard Deviation	Minimum	Maximum
Track	0.19	0.39	0	1
Attendance	28,973.42	18,127.63	2,500	98,200
Derby	0.06	0.24	0	1
Difference in bookmaker probability	0.18	0.25	−0.58	0.78
Home yellow last 3 min	0.08	0.28	0	4
Away yellow last 3 min	0.09	0.31	0	5
Home yellow prior	0.76	1.07	0	8
Away yellow prior	0.97	1.22	0	8
Home second yellow and red	0.04	0.20	0	3
Away second yellow and red	0.05	0.23	0	3
Goal difference	0.15	1.06	−6	5

variables. For second yellow and red cards, the dependent variables are replaced with *home second yellow and red card* and *away second yellow and red card* in the Primera Liga and Champions League models.

III. RESULTS

Starting with a descriptive analysis, we test whether the mean number of sanctions per game by card (yellow, second yellow, and red cards) awarded to home and away teams are equal (see Tables 1 and 2). The null hypothesis of equality is rejected ($p < .01$) in both the Primera Liga ($n = 1,520$) and the Champions League ($n = 500$). The fact that home teams receive fewer sanctions compared with away teams provides some insight into the degree of home advantage enjoyed by hosts. However, the presence of refereeing bias is inconclusive based on this evidence. It could be argued that home teams generally adopt an attacking style of play and this induces fouls play from visiting teams. Therefore, visiting teams may justifiably receive a higher proportion of sanctions.

Another aspect within our comparative analysis is the behavior of referees from Spain officiating in the Champions League compared with those from other countries. Might it be that Spanish referees have a propensity to offer a higher or lower amount of cards to home and away teams relative to their non-Spanish counterparts? To perform this comparative analysis, a vector of mean values for home and away yellow cards, and home and away dismissals for Spanish referees is compared with that of non-Spanish referees using the paired Hotelling's *t*-squared test. The test determines whether a set of means is equal between the two groups. The null hypothesis of equality between Spanish and non-Spanish referees in the Champions League cannot be rejected (*t*-squared = 5.01 with $p = .29$) suggesting that the behavior of Spanish referees in the allocation of cards is not significantly different to their counterparts. This is an interesting result in that if there was a difference, the award of cards might be deemed to be dependent on the referees' countries of origin; however, in the absence of any statistical difference, the source of any bias is consistent amongst referees irrespective of their countries.

Given that the behavior of Spanish referees in the Champions League is similar to those referees from other countries, how might the

TABLE 4
Summary Statistics for Champions League

Variable	Mean	Standard Deviation	Minimum	Maximum
Champions League group stage (n = 31,500)				
Track	0.22	0.41	0	1
Attendance	39,094.55	18,086.47	4,000	98,000
Derby	0.01	0.08	0	1
Difference in bookmaker probability	0.21	0.31	−0.60	0.81
Home team qualified from group stage	0.01	0.09	0	1
Away team qualified from group stage	0.01	0.12	0	1
Goal difference	0.26	1.12	−5	7
Home yellow last 3 min	0.04	0.21	0	2
Away yellow last 3 min	0.06	0.24	0	2
Home yellow prior	0.47	0.76	0	5
Away yellow prior	0.67	0.95	0	5
Home second yellow and red	0.01	0.12	0	2
Away second yellow and red	0.04	0.20	0	2
Champions League knockout stage (n = 10,500)				
Track	0.17	0.37	0	1
Attendance	50,643.08	19,718.78	15,000	98,436
Derby	0.07	0.26	0	1
Difference in bookmaker probability	0.19	0.22	−0.29	0.65
First leg goal difference	−0.20	0.87	−3	3
Quarter final	0.28	0.45	0	1
Semi final	0.14	0.35	0	1
Goal difference	0.27	1.01	−3	6
Home yellow last 3 min	0.05	0.21	0	2
Away yellow last 3 min	0.07	0.25	0	2
Home yellow prior	0.52	0.83	0	6
Away yellow prior	0.75	0.96	0	6
Home second yellow and red	0.02	0.13	0	1
Away second yellow and red	0.03	0.18	0	2

behavior of Spanish referees compare across the two competitions? Again using the paired Hotelling's *t*-squared test, this time there is a significant difference in the allocation of cards in the two competitions and the null hypothesis being rejected (*t*-squared = 57.22 and *p* = .00). The evidence in Tables 1 and 2 suggests that a lower amount of cards are awarded in the Champions League compared with Primera Liga; however, isolating Spanish referees in this context suggests that they allocate more cards in their domestic league. The difference might be explained by player behavior. In the Champions League, the costs of sanctions are higher compared with domestic leagues. This is particularly the case in the group stages in which there are only six games per team and more so in the knockout phase in which the fate of teams are decided in two matches, one at home and one away; in the domestic team, comparative cost can be distributed across a much larger number of games.

A. Pregame Variables and the Propensity for Yellow Cards

To establish whether bias exists, close examination of the results in Tables 5–8 is necessary. Table 5 shows the results of the bivariate probit model for home and away team yellow cards in the Primera Liga. Starting with the pregame variables, the coefficient of *track* is significantly different from zero for both the home ($p < .1$) and away ($p < .01$) teams. If a track is present, controlling for other factors, the probability of the home team receiving a yellow card increases. In contrast, the probability of the away team receiving a yellow card decreases. This can be viewed as evidence of crowd pressure. The farther the crowd is from the action, the less inhibited the referee feels and so the propensity to award the home team yellow cards increases. Added to this, the reduced pressure from the crowd manifests itself in a reduced probability of a yellow card for the away team. The impact of track on the

TABLE 5
Bivariate Probit Model for Yellow Cards in Spain's Primera Liga[a]

Independent Variables	Home Team		Away Team	
	Coefficient	Absolute t Statistic	Coefficient	Absolute t Statistic
Pregame				
Track	0.115*	(1.91)	−0.444***	(8.34)
Log attendance	0.019	(0.48)	0.074*	(1.87)
Derby	0.036	(1.21)	−0.007	(0.28)
Difference in bookmaker probability	−0.247***	(3.12)	0.164**	(2.19)
Difference in bookmaker probability squared	−0.103	(0.94)	−0.255**	(2.36)
Within-game				
Minute	0.011***	(8.72)	0.015***	(12.49)
Minute squared	−0.000013	(1.02)	−0.000069***	(5.66)
45th minute	−0.314***	(3.45)	−0.344***	(4.07)
90th minute	−0.768***	(7.13)	−0.666***	(6.49)
Goal difference	−0.005	(0.74)	−0.004	(0.53)
Goal difference squared	−0.022***	(6.05)	−0.021***	(5.91)
Home yellow last 3 min	−0.140***	(5.09)	0.088***	(3.69)
Away yellow last 3 min	0.047*	(1.94)	−0.194***	(7.58)
Home yellow prior	−0.076***	(8.99)	0.017**	(2.18)
Away yellow prior	0.008	(1.07)	−0.080***	(10.01)
Home second yellow and red	0.005	(0.13)	0.053	(1.64)
Away second yellow and red	0.038	(1.33)	0.001	(0.05)
Constant	−2.559***	(6.75)	−3.009***	(7.81)
Home team dummies			Yes	
Away team dummies			Yes	
Referee dummies			Yes	
Season dummies			Yes	
Rho			0.088*** (5.69)	
Log-likelihood			−34644	
n			135,720	

Note: Robust z statistics in parentheses.

[a]Given the large number of independent variables, the issue of multicollinearity should be addressed. Variance inflation factors (VIF) for each of the variables in all the regression models are examined. With the exception of minutes and minute squared which are approximately 18 in all four models, the highest VIF measure is 2.39. All the VIF measures are considerably below the threshold of 30 that is normally used. Consequently, multicollinearity is not a problem.

*Significant at .1%; **significant at .05%; ***significant at .01%.

probability of a yellow card in the Champions League is not significantly different from zero (see Table 6).[1]

On this dimension, a bias that is present in the domestic league (Spain) is not found in the Champions League, in which neutral

referees are appointed. Hence, the appointment of neutral referees may well have had a beneficial effect in terms of reduction of bias. However, Dawson and Dobson (2010) report convincing evidence, in the form of significant referee nationality effects, that country of origin does induce variation in the disciplinary sanctions awarded by supposedly neutral referees in the Champions League. These country of origin data were unavailable to us in the present study.

The size of the crowd (*log attendance*) matters to some extent in both the Primera Liga and the Champions League. As the size of the crowd increases, the probability of the away team receiving a yellow card increases (*p* value of estimated coefficient <.10) in the Primera

---

1. Dawson and Dobson (2010) report results from the Champions League showing that presence of a running track is associated with greater disciplinary points, taken as a measure that combines yellow and red cards. This significant effect becomes zero, however, when they introduce control variables for what they term "within-game dynamics," which include total fouls and total shots during a game. Recall that a key difference between our analysis and Dawson and Dobson is that we use minute of game as our unit of observation and hence we are able to control for events in games as they occur, thus avoiding the charge of endogeneity of control variables.

TABLE 6
Bivariate Probit Model for Yellow Cards in the Champions League

	Home Team		Away Team	
Independent Variables	Coefficient	Absolute t Statistic	Coefficient	Absolute t Statistic
Pregame				
Track	−0.020	0.17	−0.05	0.47
Log attendance	0.133*	1.83	0.126*	1.81
Derby	0.063	0.43	0.10	0.70
Difference in bookmaker probability	−0.403	1.60	0.24	1.11
Difference in bookmaker probability squared	0.244	0.92	−0.658***	2.66
Group stage	−0.005	0.08	−0.07	1.27
Home team qualified from group stage	0.404	1.19	−0.456**	2.11
Away team qualified from group stage	0.263	1.25	−0.14	0.98
Quarter final	0.054	0.61	−0.04	0.56
Semi final	−0.109	0.84	−0.11	0.94
First leg goal difference (second leg home team—second leg away team)	−0.104**	2.28	0.01	0.17
Within-game				
Minute	0.017***	5.15	0.018***	7.36
Minute squared	−0.00010***	3.00	−0.00010***	4.06
45th minute	0.411***	3.82	0.472***	4.99
90th minute	0.212	1.64	−0.027	0.20
Goal difference	−0.020	0.77	−0.023	1.32
Goal difference squared	0.010	0.72	−0.002	0.37
Home yellow last 3 min	−0.501***	5.89	0.187***	2.84
Away yellow last 3 min	0.122*	1.80	−0.451***	6.67
Home yellow prior	−0.236***	7.75	0.049**	2.10
Away yellow prior	0.073***	3.04	−0.212***	9.30
Home second yellow and red	−0.020	0.12	0.053	0.69
Away second yellow and red	−0.040	0.49	−0.217**	2.29
Constant	0.017***	5.15	0.018***	7.36
Home team dummies			Yes	
Away team dummies			Yes	
Referee dummies			Yes	
Season dummies			Yes	
Rho			0.114*** (2.79)	
Log-likelihood			−7050	
n			41,610	

Note: Robust z statistics in parentheses.
*Significant at .1%; **significant at .05%; ***significant at .01%.

Liga. However, the probability of home team yellow card is unaffected by crowd size. In the Champions League, a larger crowd increases the probability of the home team receiving a yellow card ($p < .05$). There is a similar, though now marginally significant, positive effect with $p < .10$, of crowd size on probability of the away team receiving a yellow card. One explanation of this finding is that the crowd are exerting pressure on the home team players whose extra efforts result in illegal play punishable by a yellow card. Larger crowds also result in higher probability of a yellow card to away team players. This could be due to crowd pressure exerted on the referee or it could be away team player's response to increased support for the home team in the form of increased effort, some of which is illegal in the form of fouls and other yellow card offences. Our findings on crowd effects are similar to those for the Champions League reported by Dawson and Dobson (2010), although we again note that we have controlled for within-game events as the match progresses in our minute-by-minute analysis.

Turning to our control variables, *derby* has no noticeable effect on the likelihood of cautions for either team in either competition. The control

TABLE 7

Bivariate Probit Model for Second Yellow and Red Cards in Spain's Primera Liga

Independent variables	Home Team		Away Team	
	Coefficient	Absolute t Statistic	Coefficient	Absolute t Statistic
Pregame				
Track	0.001	0.02	0.003	0.06
Log attendance	0.071	1.47	0.108**	2.43
Derby	−0.035	0.41	−0.078	0.90
Difference in bookmaker probability	−0.042	0.35	−0.055	0.46
Difference in bookmaker probability squared	−0.704**	2.41	−0.429*	1.66
Within-game				
Minute	−0.003	0.73	0.004	0.94
Minute squared	0.000107***	2.72	0.000035	0.88
45th minute	−0.114	0.39	−0.251	0.86
90th minute	−0.382**	2.40	−0.276*	1.85
Goal difference	−0.049***	2.60	0.048**	2.57
Goal difference squared	−0.013	1.36	−0.008	1.13
Home yellow last 3 min	0.328***	5.75	0.091	1.36
Away yellow last 3 min	0.139**	2.17	0.294***	5.73
Home yellow prior	0.110***	6.18	0.016	0.77
Away yellow prior	0.030	1.55	0.101***	6.13
Home second yellow and red	−0.126	1.60	0.074	1.03
Away second yellow and red	0.071	1.18	−0.034	0.62
Constant	−4.152***	8.15	−4.516***	9.89
Home team dummies		No		
Away team dummies		No		
Referee dummies		Yes		
Season dummies		Yes		
Rho		0.339*** (5.98)		
Log-likelihood		−3380		
n		135,720		

Note: Robust z statistics in parentheses.
*Significant at .1%; **significant at .05%; ***significant at .01%.

variables for team quality, *difference in bookmaker probability* and its square, largely behave as expected. The more favored the home team is in the Primera Liga, the probability of a caution decreases at a constant rate. For the away team, the likelihood of a caution increases but at a decreasing rate. The most likely explanation is again player response. The more favored the home team is by the betting public, the more likely they are to be attacking and are therefore less likely to be in receipt of yellow cards. The away team, conversely, will be defending and this increases the likelihood of yellow cards as effort spills over into illegal activity. In the Champions League, the effect of team quality influences only the away team's propensity to receive a yellow card. The negative relation in the quadratic of difference in bookmaker probability suggest that as the home team becomes more favored to win the match, the likelihood

of a yellow card being shown to the away team reduces. Given the greater superiority of the home team, there would seem little point in engaging in behavior punishable by a yellow card.

For those pregame variables specific to the Champions League, of note is the impact of group stage games in which either the home or away teams have already qualified for the knockout stage. When the home teams have already qualified, the away team is less likely to be cautioned. The likely explanation is that there is less incentive on the home team's part to be attacking. Home team players are likely to be cautious as they try to avoid injury or miss subsequent Champions League matches. Consequently, the away team's need to defend is curtailed somewhat and so the probability of yellow cards to the away team reduces. Another Champions League specific variable is the score

TABLE 8
Bivariate Probit Model for Second Yellow and Red Cards in the Champions League

Independent Variables	Home Team		Away Team	
	Coefficient	Absolute *t* Statistic	Coefficient	Absolute *t* Statistic
Pregame				
Track	0.148	1.11	−0.197*	1.65
Log attendance	−0.106	1.07	−0.026	0.28
Derby	0.021	0.07	−4.480***	20.71
Difference in bookmaker probability	0.907	1.49	−0.181	0.74
Difference in bookmaker probability squared	−2.180*	1.81	0.130	0.28
Group stage	0.095	0.51	0.103	0.8
Home team qualified from group stage	−4.580***	20.41	−4.610***	25.12
Away team qualified from group stage	−4.620***	29.51	0.274	0.99
Quarter final	0.035	0.16	0.140	0.7
Semi final	0.203	0.67	0.030	0.09
First leg goal difference	−0.169	1.59	0.159*	1.84
Within-game				
Minute	0.016	1.55	0.002	0.25
Minute squared	−0.0000504	0.55	0.0000271	0.31
45th minute	0.799***	3.04	0.505*	1.94
90th minute	0.252	0.88	0.144	0.64
Goal difference	−0.131*	1.95	0.033	0.61
Goal difference squared	−0.090***	2.68	−0.014	0.65
Home yellow last 3 min	0.436**	2.46	0.008	0.05
Away yellow last 3 min	−0.146	0.71	0.543***	3.95
Home yellow prior	0.223***	4.30	−0.145**	2.49
Away yellow prior	−0.036	0.73	0.302***	6.65
Home second yellow and red	−0.048	0.17	0.139	0.61
Away second yellow and red	0.093	0.34	−0.079	0.65
Constant	−2.960***	2.72	−3.310***	3.24
Home team dummies			No	
Away team dummies			No	
Referee dummies			No	
Season dummies			Yes	
Rho			0.644*** (3.95)	
Log-likelihood			−592	
n			41,610	

Note: Robust *z* statistics in parentheses.
*Significant at .1%; **significant at .05%; ***significant at .01%.

line from the first leg in the knockout phase of the competition. In instances where the home team has a positive score line from the first leg, the greater this score line, the less likely it is to receive a yellow card. For the home teams, the benefits of home advantage and a lead on aggregate diminishes the need for playing strategies that are likely to result in yellow cards. In the case of the away team, the score line from the first leg has no significant effect.

Among the pregame variables, we find some evidence of home advantage and refereeing bias in the form of *track* in the Primera Liga. The effects captured by *log attendance* in the

Champions League cannot be conclusively put down to referee bias and are likely to be partly due to behavioral responses on the part of players. The focus of the following section is on within-game variables and the effect on the likelihood of yellow cards in both the Primera Liga and the Champions League.

B. Within-Game Variables and Yellow Cards

Minute and its square are control variables. Their effects are somewhat consistent across the two competitions and teams. As the match progresses, the likelihood of a yellow card increases but at a diminishing rate, with the exception

of the home team in the Primera Liga where the likelihood of a yellow card increases at a constant rate. In the Primera Liga, teams are less likely to receive a caution in the 45th minute, 90th minute, and any extra time at the end of each half. Conversely, in the Champions League, both teams are likely to be cautioned during the 45th minute and first half injury time. However, this only marginally applies to the home team at the end of the second half. This would suggest that, controlling for other factors, players' behaviors with respect to cautions are dependent on the stage reached in the match.

The state of the match with respect to score line matters in the Primera Liga. For every additional goal by which the home team leads, the likelihood of a card reduces at an increasing rate for both teams. This is a matter of incentives. As the game becomes further out of reach for the away team, there is little point in players of either team engaging in foul play. The contrast is also true in that for every additional goal the home team finds itself behind, the likelihood of a caution decreases for either team. The effect of score line in the Champions League on the probability of a card being awarded is not significantly different from 0. The likely explanation is the complexity of the competition's structure, which includes the group phase, as well as knockout phase, first and second legs, and the importance of away goals. Consequently, the score line is not absolute but relative to a set of factors.

The history of sanctions within the game prior to the subject minute is a determinant of whether teams are likely to be awarded a yellow card. In both the Primera Liga and Champions League, if the home team had received a yellow card in the last 3 minutes, the likelihood of receiving a yellow card is reduced ($p < .01$ in both cases). Interesting is that for both competitions, the likelihood of the away team receiving a card increases ($p < .01$ in both models). This may be reflecting an evening-out of sanctions by the referee. It could also be capturing retaliatory behavior on the part of the away team. In the case of the former, this is an indication of home advantage and bias. However, it is not possible to separate out home advantage and bias from retaliation. If the away team has been awarded a card in the previous 3 minutes, this increases the likelihood of the home team being awarded a card and reduces the likelihood of the away team being awarded a card in both competitions. These results are replicated for yellow

cards awarded to the home team prior to the last 3 minutes; in both competitions, they reduce the home team's likelihood of receiving a yellow card while increasing that of the away team's. The impact of dismissals is limited to away team in the Champions League. If the away team suffers a dismissal earlier during the match, the likelihood of receiving a yellow card reduces. The history of sanctions during the match clearly has the dual role of influencing player and referee behavior to the point that refereeing bias cannot be rejected.

C. Dismissals in Primera Liga and Champions League

For dismissal, we turn attention to Tables 7 and 8. The impact of *track* is different from the effects noted on the likelihood of yellow cards. There is no significant effect on home or away teams' likelihood of dismissals in Primera Liga. The same finding applies to home team likelihood of dismissal in the Champions League. However, for the Champions League, the presence of a running track reduces the likelihood of a dismissal for the away team, albeit with marginal significance ($p < .10$). Again, this shows the effect of crowd pressure and its manifestation in home advantage and refereeing bias.

The size of the crowd (*log attendance*) increases the likelihood of a dismissal for the away team in Primera Liga ($p < .05$) but has no discernable effect on either home or away team in the Champions League. Causality could be running from crowd to players in which case, there is no bias. It could be that the increased crowd is inducing more attacking play from the home team players which then results in foul play on the part of away team players which is then followed by dismissal. Alternatively, it could be that the crowd is able to influence the behavior of the referee in the award of a red card. In this case, it is not possible to distinguish between players' behavioral response and referee bias. Another possibility is that UEFA's policy of neutral referees ensures that referee policy toward dismissals is independent of crowd size (so referees are not at all responsive to social pressure on this dimension). In contrast, faced with larger crowds in Spain, referees tend to exhibit a greater propensity to award red cards to away teams, over and above the effects captured by pregame and within-game control variables.

For derby matches, a significant impact is noted only in the Champions League. The probability of the away team player being dismissed is reduced for matches of local rivalry. The effect of *difference in bookmaker probability* is also significant. As difference in bookmaker probability increases, so that bookmakers' assessment is of a less close contest, ex ante, the likelihood of home team dismissals declines for Primera Liga matches, ($p < .05$). A similar effect is noted in the Champions League but with a lower degree of significance ($p < .1$ for *difference in bookmaker probability squared*). Further, for Champions League group phase matches, whether the home or away teams have qualified for the knockout phase is significant and influences the behavior of the teams as to be expected, except for the effect of *away team qualified from group stage* on away team dismissal. The status of the contests also matters in the knockout phase of the Champions League. If the home team has an advantage from the first leg, the away team has a greater likelihood of receiving a red card ($p < .1$).

With respect to within-game control variables, *minute squared* is significant and positive showing that home team players are more likely to be dismissed, increasingly, as the match draws to a close. For both teams' players, the likelihood of dismissal increases in the *45th minute* which includes first half injury time in the Champions League. The effects of injury time at the end of the second half is only noticeable in the Primera Liga in which players of both teams are less likely to be dismissed.

As goal difference increases, the home team is less likely to suffer a dismissal in both competitions. For the away team in the Primera Liga, players are likely to be shown a second yellow or red card. This applies only to the likelihood of a dismissal for the away team in the Primera Liga. However, no effect is noted for the likelihood of away team dismissal in the Champions League.

With respect to cards, dismissal of the home team players are influenced by yellow cards to the home teams in the preceding 3 minutes in both competitions. If the away team has been awarded a card in the last 3 minutes, the home and away teams are likely to have a player dismissed in the Primera Liga. This only applies to the away team in the Champions League. For yellow cards awarded to the home team prior to the last 3 minutes, this again increases the

likelihood of dismissal to the home team in both competitions. Contrastingly, the probability of a dismissal for the away team in the Champions League falls. As expected, yellow cards awarded to the away team prior to the last 3 minutes are likely to result in dismissals for the away team. Taken overall, the role of cards awarded during the match have the expected effects on dismissals.

For dismissals in both competitions, the extent of home advantage and referee bias is not as high compared with those noted in the case of yellow cards. One reason for this is that the consequences of dismissal are serious and are likely to have a dramatic effect on the match compared with the issue of a caution. However, this is not to suggest that there is no evidence of referee bias in dismissing players; the effects reported through the variables *track*, in the Champions League, are indicative of bias.

IV.   CONCLUSIONS

In this paper, we test for, and find evidence of, referee bias in favor of home teams in European football. The context for the analysis is Spain's Primera Liga and the UEFA Champions League. As with Buraimo, Forrest, and Simmons (2010), we adopt an analytical approach in which match minute is the unit of observation. By using minute as the unit of observation, we are able to control for a set of covariates capturing pregame information and for a set of control variables representing events that unfold during the course of the match. The results of the analysis show that pregame and within-game information both contribute to referee behavior in terms of likelihood of awarding a card in a given minute.

In Primera Liga matches where the crowd is separated from the pitch by running tracks, we find that the probability of the award of a yellow card to the home is higher and that of the away team is lower compared to matches played at stadia without running tracks. This finding is similar to that of Buraimo, Forrest, and Simmons (2010) for England and Germany. Our additional evidence on running track effects in Spain gives greater force to the charge of refereeing bias created by social pressure. In the Champions League the probability of an away team player being dismissed is reduced in the presence of a running track, which is again indicative of referee bias.

Referee behavior is also influenced by the size of the crowd in attendance. Although we have no evidence on the composition of the crowd, stadium attendance can be used as a proxy for home team support. We find that the greater the crowd size, the more likely the away team is to be awarded a yellow card in the Primera Liga. This provides evidence that the crowd, and more specifically its size, is able to exert enough social pressure on officials to the extent that decisions favor the home team.

Cautions are more likely to be given to away team players, in both competitions, if the home team has received one in the previous 3 minutes; the likelihood of the home team receiving a card is reduced under the same circumstances. Similarly, cautions to the away team influence how cautions are awarded to the home and away team and the presence of bias cannot be ruled out. In some instances, it is difficult to separate referee bias from player behavior. For instance, the increased probability of a yellow card to the away team following a card to the home team may be a result of retaliation from away team players rather than bias. However, bias cannot be ruled out given the collective evidence emerging from the impact of *track* and crowd size.

The Primera Liga relies on referees drawn from the domestic market. In contrast, UEFA is able to draw on the *best* referees across major European leagues for its prestigious Champions League competition. In light of this, one would have expected the degree of refereeing bias to be less evident in the Champions League than in the Primera Liga. On this, we have conflicting evidence. On the one hand, the positive impact of track presence on propensity to award home yellow cards that is present in Spain does not carry over into the European competition with neutral referees. However, for red cards, we find that the likelihood of player dismissal is lower in the presence of the track, suggesting that some bias against away teams does remain.

Europe's premier competition is subject to similar significant impacts of social pressure through crowd size as in the domestic competition, specifically in relation to a positive effect on the award of yellow cards to visiting teams. Interestingly, in the Champions League, we find a significant positive effect of crowd size on probability of award of a home team yellow card that was not present in the domestic competition.

This suggests a contrarian approach by referees in response to crowd pressure, where they overcompensate by means of a higher probability of award of a yellow card to the home team. More reassuringly, from UEFA's perspective, significant crowd effects on probability of red card to either team are not revealed in the Champions League.

Overall, we have some partial grounds for optimism to UEFA on the effectiveness of its policy of appointing neutral referees. However, our findings from the Champions League on the response of probability of away team red card to presence of running track and on the response of probability of home team yellow card to crowd size suggest that elements of bias do persist, despite the appointment of neutral referees. Further research on these questions is needed. One important extension will be to examine the incentives for referee performance in terms of remuneration in different competitions and also the career implications of good performance in the Champions League and domestic competitions (Bryson, Buraimo, and Simmons 2011). Another useful extension would be to model the subjective ratings of refereeing performance as provided by newspapers and specialist football magazines (Frick, Gürtler, and Prinz 2008).

REFERENCES

Boyko, R., A. Boyko, and M. Boyko. "Referee Bias Contributes to Home Advantage in English Premiership Football." *Journal of Sports Sciences*, 25, 2007, 1185–94.

Bryson, A., B. Buraimo, and R. Simmons. "Do Salaries Improve Worker Performance?" *Labour Economics*, 18, 2011, 424–33.

Buraimo, B., D. Forrest, and R. Simmons. "The 12th Man? Refereeing Bias in English and German Soccer." *Journal of the Royal Statistical Society, Series A*, 173, 2010, 431–49.

Corral, J., J. Rodriguez, and R. Simmons. "The Effects of Incentives on Sabotage: The Case of Spanish Football." *Journal of Sports Economics*, 11, 2010, 243–60.

Dawson, P., and S. Dobson. "The Influence of Social Pressure and Nationality on Individual Decisions: Evidence from the Behaviour of Referees." *Journal of Economic Psychology*, 31, 2010, 181–91.

Dawson, P., S. Dobson, J. Goddard, and J. Wilson. "Are Football Referees Really Biased and Inconsistent? Evidence from the English Premier League." *Journal of the Royal Statistical Society, Series A*, 170, 2007, 231–50.

Dohmen, T. "The Influence of Social Forces: Evidence from the Behavior of Football Referees." *Economic Inquiry*, 46, 2008, 411–24.

Forrest, D., J. Goddard, and R. Simmons. "Odds Setters as Forecasters." *International Journal of Forecasting*, 16, 2005, 551–64.

Frick, B., O. Gürtler, and J. Prinz. "Men in Black: Monitoring and Performance of German Soccer Referees," in *Soccer: Economics of a Passion*, edited by H. Dietl, E. Franck, and H. Kempf. Schorndorf: Hofmann, 2008, 309–22.

Garicano, L., I. Palacios-Huerta, and C. Prendergast. "Favoritism under Social Pressure." *Review of Economics and Statistics*, 87, 2005, 208–16.

Greene, W. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall, 2008.

Parsons, C., J. Sulaeman, M. Yates, and D. Hamermesh. "Strike Three: Umpires' Demand for Discrimination." *American Economic Review*, 101, 2011, 1410–35.

Prendergast, C. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37, 1999, 7–63.

Price, J., and J. Wolfers. "Racial Discrimination among NBA Referees." *Quarterly Journal of Economics*, 125, 2010, 1859–87.

Rickman, N., and R. Witt. "Favouritism and Financial Incentives: A Natural Experiment." *Economica*, 75, 2008, 296–309.

Ridder, G., J. S. Cramer, and J. Hopstaken. "Down to Ten: Estimating the Effect of a Red Card in Soccer." *Journal of the American Statistical Association*, 89, 2004, 1124–27.

Scoppa, V. "Are Subjective Evaluations Biased by Social Factors or Connections? An Econometric Analysis of Soccer Referee Decisions." *Empirical Economics*, 35, 2007, 123–40.