I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12142

Twelve Eyes See More Than Eight. Referee Bias and the Introduction of Additional Assistant Referees in Soccer

Olivier Verstraeten
Stijn Baert

FEBRUARY 2019

DISCUSSION PAPER SERIES

IZA DP No. 12142

Twelve Eyes See More Than Eight. Referee Bias and the Introduction of Additional Assistant Referees in Soccer

Olivier Verstraeten
Ghent University

Stijn Baert
Ghent University, Research Foundation – Flanders, University of Antwerp, Université-catholique de Louvain, IZA and IMISCOE

FEBRUARY 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Twelve Eyes See More Than Eight. Referee Bias and the Introduction of Additional Assistant Referees in Soccer

We are the first to investigate whether the introduction of additional assistant referees in the UEFA Europa League (in season 2009-2010) and the UEFA Champions League (in season 2010-2011) was associated with lower referee bias. To this end, we analyse a unique database with pre- and within-game characteristics of all games in seven recent seasons in these leagues by means of bivariate probit regression models. We find evidence for substantial referee bias before the introduction of additional referees, while no such evidence is found after the introduction. Furthermore, additional assistants go hand in hand with more yellow cards for both home and away teams. We show that these findings are robust to multiple operationalisations of referee bias and that they are not just picking up a general time evolution towards less referee bias or the effect of parallel reforms.

JEL Classification: L83, J44, Z00

Keywords: Soccer, referee bias, assistant referees, bivariate probit model, UEFA Champions League, UEFA Europa League

Corresponding author:
Stijn Baert
Ghent University
Sint-Pietersplein 6
9000 Gent
Belgium

E-mail: Stijn.Baert@UGent.be

I. Introduction

Soccer referees are often accused of bias in their decision-making. For example, after the return of the quarter final of the Union of European Football Association (UEFA) Champions League season 2016-2017 between Real Madrid and Bayern Munich, the referee was criticised for giving an unjustified second yellow card to Bayern Munich player Vidal and not giving a second yellow card to Real Madrid player Casemiro (The Guardian, 2017). Moreover, Real Madrid player Ronaldo scored in extra time from an offside position. These criticisms suggest that the referee decisions were biased. This bias of referees—incorrect decisions may result from conscious bias as well as from human error or incompetence—may have had a big economic impact. Real Madrid won the tournament and received 23 million euros more prize money than Bayern Munich. However, it does not end there. Other revenues, such as television rights, are divided based on performance as well. Thus, Real Madrid earned more money, partly as a result of referee decisions. Other examples are the return of the quarter final of the UEFA Champions League 2016-2017 between FC Barcelona and Paris Saint-Germain (Brown, 2017) and the return of the quarter final of the UEFA Champions League 2012-2013 between Dortmund and Malaga (BBC, 2013). In these games as well, it is obvious that referee decisions had serious economic consequences.

Therefore not surprisingly, many academics have investigated this popular idea of referee bias in European soccer (Buraimo, Forrest, & Simmons, 2010; Buraimo, Simmons, & Maciaszszyk, 2012; Dawson & Dobson, 2010; Dohmen, 2008; Garicano, Palacios-Huerta, & Prendergast, 2005; Pettersson-Lidbom & Priks, 2007; Rickman & Witt, 2008; Scoppa, 2008; Sutter & Kocher, 2004). As reviewed by Dohmen and Sauermann (2016), these academics' research confirms that this bias is not a myth: it does exist. For instance, Garicano et al. (2005), Dohmen (2008) and Scoppa (2008) showed, based on data for Spanish, German and Italian national competitions, that the injury time awarded by the referee in the second half of a soccer game is significantly longer if the away team is one goal ahead compared with when the home team is leading by one goal. As a result, home teams are afforded more time to come back from a losing position than away teams when they are losing by one goal. In addition, Dawson and Dobson (2010) showed that teams from the big five European competitions (English Premier League, Spanish Primera Division, French Ligue 1, German Bundesliga and Italian Serie A) receive fewer yellow cards when playing against a team outside of the big five in European international club soccer, keeping team strength constant. Finally, Pettersson-Lidbom and Priks (2007), Buraimo et al. (2010) and Buraimo et al. (2012) showed that home teams receive fewer yellow cards than away teams, especially when crowd size is substantial and/or there is no running track in the stadium (so the crowd is close to the pitch), confirming the popular thought that the crowd exerts pressure on the referee.

Since referee bias is real in European soccer, and given the financial consequences that accompany it, the necessity to decrease referee bias is evident. A logical approach is to add additional assistant referees to obtain better judgement. Before the 2009-2010 season, there were four referees at every game in the UEFA Champions League and the UEFA Europa League, i.e. the two most prestigious international club soccer competitions worldwide as measured by prize and television money at stake (UEFA, 2017). Two assistant referees, also known as linesmen, assisted the (head) referee particularly in decisions regarding offside and whether the ball leaves the field of play. A fourth official assisted the referee in ensuring that substitutions were in accordance with the Laws of the Game. In season 2009-2010, the UEFA introduced two additional assistant referees as a trial during all games played in the UEFA Europa League. In the following season, this approach was extended to all games played in the UEFA Champions League. The two additional referees work behind the goal line and assist the referee in deciding situations in or near the penalty area. As the UEFA believes that the additional assistant referees contribute to better control of the game, partly because they allow the other four (assistant) referees to focus on their core tasks, the additional assistant referees are now ubiquitous in European international club soccer (UEFA, 2014). In this study, to the best of our knowledge, we are the first to investigate whether this belief corresponds with reality.

More concretely, we answer the following research question: Did referee bias in European international club soccer decrease since the introduction of a fifth and sixth referee? To do so, we analyse a unique dataset consisting of all games played in the UEFA Champions League and UEFA Europa League between the 2007-2008 and 2013-2014 seasons. We replicate the analyses in Buraimo et al. (2010), Buraimo et al. (2012) and Dawson and Dobson (2010) based on these more recent data and investigate whether their measures of referee bias are heterogeneous by the introduction of additional assistant referees.

II. Methods

II.1 Data

To answer our research question, we constructed a dataset containing all games played in the UEFA Champions League and UEFA Europa League between the 2007-2008 and 2013-2014 seasons except for final games (for which no home and away games are designated). Both competitions are organised by the UEFA, with the aim to let the best European teams from different competitions compete against each other. Each national football association receives a number of entry tickets based on their teams' performances over the past years in both competitions. These tickets can either be direct entry tickets or entries for one of the preliminaries and are allocated to the teams that finished highest in the domestic competitions in the previous season. The highest placed teams in these national soccer competitions receive the UEFA Champions League tickets. The UEFA Europe League tickets go to the teams next in line.

After the preliminaries, which start with 78 teams, the UEFA Champions League starts with a group stage of 32 teams divided into groups of four teams. Every team plays home and away against the other teams in its group. The first and second place teams qualify for the knock-out stage of the UEFA Champions League and the third place teams proceed to the knock-out stage of the UEFA Europa League. Since the 2009-2010 season, the UEFA Europa League starts the season with a group stage of 48 teams divided into groups of four teams. The winners and runners-up of each group go through to the knock-out stage. During seasons 2007-2008 and 2008-2009, the format of the UEFA Europa League, then called 'UEFA Cup', was different. The group stage started with 40 teams divided into groups of five teams. In contrast to the current format, there was only one encounter between the teams, with each team playing twice at home and twice away. The three first placed teams for each group qualified for the knock-out stage.

In the knock-out stage of both competitions, two teams play against each other home and away, and the team that scores the most goals over those two games proceeds to the next round. In case of a draw, the team that scores the most goals when playing away proceeds to the next round. If the goals scored away are equal, then 30 minutes of extra time are added. If the score is still equal after those 30 minutes, a penalty shoot-out decides which team proceeds to the next round. In contrast, the final is played in one match on neutral ground. In case of a draw, 30 minutes of extra time and, if necessary, a penalty shoot-out decide which team wins the league.

In total, our dataset contains 2,168 games, of which 868 are from the UEFA Champions League and 1,300 are from the Europa League. These games are played between 189 different soccer teams. Unlike most studies on referee bias (e.g. Dohmen, 2008; Scoppa, 2008), we do not use the game as the unit of observation. Instead, we analyse our data on the minute level, in line with Buraimo et al. (2010) and Buraimo et al. (2012). The main advantage of this approach is that we can take the flow of events that occur during the game into account. In the end, two games can have the same number of goals and yellow and red cards although the order of events is completely different. For example, the advantage of a particular team during the game may have caused the opponent to play more aggressively and receive a red card, whereas in another game, a team may have received a red card that resulted in the other team scoring more easily. This approach results in a dataset of 195,120 observations at the game-minute level. We explain below how we take into account that (i) these observations are clustered at the game level and (ii) minute 45 and minute 90 may take more than one minute due to injury time.

In line with Buraimo et al. (2010), Buraimo et al. (2012) and Dawson, Dobson, Goddard and Wilson (2007), in our benchmark analysis we assume that home teams are favoured by referees regarding the award of yellow cards. This form of favouritism can result in a lower probability for the home team to receive a yellow card as well as a higher probability for the away team to receive a yellow card in a certain minute. These expectations have been related to the aforementioned idea that the crowd can exert pressure on the referee. Consequently, in line with Buraimo et al. (2010) and Buraimo et al. (2012), we expect that home teams will be less (or not at all) favoured if there is a track separating the crowd from the field. Without such a track, the referee will be more intimidated by the home crowd. Therefore, in cases of referee bias, home teams are expected to have a higher probability to receive a yellow card in a certain minute in games where a track separates the crowd from the field. In contrast, for away teams we assume that they have a lower possibility to receive a yellow card in a certain minute if there is a track. More importantly, we assume that these effects will be lower (if not undone) in games where a fifth and a sixth referee are present because we believe that these two extra pairs of eyes will lead to better judgement and thus less referee bias.

Subsequently, consistent with Dawson and Dobson (2010), in our alternative analysis we assume that teams from the big five European soccer competitions will be favoured in the award of yellow cards as a result of the prestige of these competitions. More specifically, our assumption is that the big five teams will have less chance to receive a yellow card in a certain minute when playing against non-big five teams, keeping the teams' relative strength constant. On the other hand, teams from outside the big five are expected to have a higher chance to receive a yellow card in a certain minute when playing against teams from the big five. Again, we expect that this favouritism will be lower (if not undone) in games where additional assistant referees are present.

To test these expectations—and thereby answer our research question—numerous variables were collected with respect to the included games. The first and most important source were the games reports available on the official UEFA website (http://www.uefa.com). Since the crowd size during these games and the name of the referee were not always available in these reports, we augmented the data from these reports with additional information from Worldfootball (http://www.worldfootball.net). Table 1 presents the descriptive statistics for the variables used in our statistical analyses. Panel A presents the dependent variables, Panel B presents the independent variables related to referee bias and Panel C presents the control variables.

In line with Buraimo et al. (2010) and Buraimo et al. (2012), we use a pair of dependent variables, i.e. the award of a yellow card in a certain minute to the home and away teams, respectively. Both are dummy variables and thus take the value 1 if a yellow card has been awarded to the team under review and 0 otherwise. As can be seen in Panel A of Table 1, a yellow card is awarded to a home team in approximately one minute out of fifty, while an away team receives a yellow card in one minute out of forty. We investigate whether this home advantage in receiving yellow cards can be related to referee bias (versus whether it is fully explained by other reasons for home advantage; Van Damme & Baert, 2018; Boyko, Boyko,

& Boyko, 2007; Johnston, 2008).

Our first independent variable is a binary variable that captures the presence of a running track between the field and the spectators. In total, 377 of the 2,168 analysed games are played in a stadium with a running track. The second and third independent variables indicate whether the home and away teams come from one of the big five European soccer competitions: the English Premier League, Spanish Primera Division, Italian Serie A, German Bundesliga and the French Ligue 1 (Dawson and Dobson, 2010). Given that all teams play as many games at home and away, the mean value for both variables is the same (and equal to 0.454). The fourth and last independent variable captures the presence of a fifth and sixth referee. These additional assistant referees are present in 69.9 percent of the analysed games.

Finally, Panel C of Table 1 shows various control variables. In line with Buraimo et al. (2010) and Buraimo et al. (2012), we included these variables because they capture other elements that may influence the chance of a yellow card award. Moreover, they are probably correlated with the independent variables. These control variables can be divided in pre-game and within-game influences.

The first pre-game control variable captures crowd size. On average, approximately 30,083 individuals are in the stadium for UEFA Champions League and UEFA Europa League games. In our analyses, we include the logarithm of the attendance, like Dawson et al. (2007), Buraimo et al. (2010) and Buraimo et al. (2012). A second pre-game determinant of receiving yellow cards is the teams' relative strength. In this respect, we followed Baert and Amez's (2018) definition to create our proxy of relative strength based on the competing teams' UEFA team coefficient. This coefficient is based on the teams' performances over the past five editions of the UEFA Champions League and UEFA Europa League. More concretely, it is the quotient of the UEFA team coefficient of the home team divided by the coefficient of the away team for that season plus 1. A third pre-game control variable is an indicator for derby games.

In line with Buraimo et al. (2010) and Buraimo et al. (2012), we assume that a game between two clubs located close to each other will be a more intense encounter due to stronger rivalry between players and supporters. Consequently, we assume that there is a higher possibility that a player receives a yellow card in a derby. Because all games in our dataset are international, like Van Damme and Baert (2018), we defined a derby as a game between two teams from the same country.

Turning now to the within-game influences, an obvious one is the goal difference at the start of the analysed minute, calculated by subtracting the number of goals scored by the away team from the number of goals scored by the home team. This goal difference is expected to be negatively (positively) correlated with aggression by the home (away) team and, therefore, with that team awarded yellow cards (Dawson et al., 2007). The following within-game influences involve previously awarded yellow and red cards. On the one hand, a yellow or red card given to a team can reduce the chance of another yellow or red card given to the same team, known as the deterrence effect (Buraimo et al., 2010; Buraimo et al., 2012). On the other hand, referees tend to compensate with the consequence that a yellow or red card given to a team may increase the chance of a yellow or red card awarded to its opponent (Buraimo et al., 2010; Buraimo et al., 2012). Specifically, and like Buraimo et al. (2010) and Buraimo et al. (2012), we included the following variables to capture these effects: (i) yellow card received by the home (away) team during the preceding three minutes of the game, (ii) number of yellow cards for the home (away) team during the game and (iii) red card received by the home (away) team during the game.

Finally, we included several within-game influences not presented in Table 1. These all control for the moment in the game. First, we included ‘minute’ and ‘minute squared’ because Buraimo et al. (2010) and Buraimo et al. (2012) found that the likelihood of a sanction increases as the game proceeds, at a decreasing rate. Second, ‘minute 45’ and ‘minute 90’ were incorporated as well. The objective of these dummies is to capture the events that happened during injury time at both half time and end of the game. We did not include these variables in Table 1 because their descriptive statistics are the same for every game.

II.2 Regression model

The data presented in Subsection II.1 are analysed by means of a bivariate probit model in line with Buraimo et al. (2010) and Buraimo et al. (2012). We jointly estimate the probability of a yellow card awarded in a certain minute to either the home or away team because the award of sanctions to one team cannot be assumed to be uncorrelated to the award of sanctions to the other team. In all models, standard errors are clustered at the game level.

III. Results

III.1 Benchmark analysis

Table 2 presents the results of our benchmark analysis. In Model (1), yellow cards for home and away teams are regressed on the control variables and one independent variable, i.e. the track variable. As noted, based on Buraimo et al. (2010), Buraimo et al. (2012) and Dawson et al. (2007), we expect that home teams are less favoured by referees concerning the award of yellow cards when a track separates the crowd from the field. Indeed, the sign of the track variable is the expected sign, i.e. positive for home teams and negative for away teams. However, in contrast to Buraimo et al. (2010), Buraimo et al. (2012) and Dawson et al. (2007), this variable is not statistically significantly different from 0. Thus, the evidence for referee bias in the sense that home (away) teams are awarded more (fewer) yellow cards when a track separates the crowd from the pitch is not replicated based on our more recent data. This can possibly be explained by the fact that most games (i.e. 69.9 percent) in our dataset are games with additional assistant referees that were not yet present in the data used in the earlier literature.

In Model (2), we add two variables to Model (1), the variable capturing the presence of additional referees and the interaction between the latter variable and the track variable. Model (2) confirms our expectation of less referee bias after the adoption of a fifth and sixth referee. First, after inclusion of the additional variables, the track variable becomes statistically significant with respect to both dependent variables. Thus, in the reference situation in which the additional assistant referees are not present, the home team has a statistically significantly higher probability to receive a yellow card ($b = 0.072$; $p = 0.032$) and the away team has a statistically significantly lower probability to receive a yellow card ($b = -0.063$; $p = 0.038$) in each minute if there is a track. In other words, we find evidence for referee bias in the sense of Buraimo et al. (2010), Buraimo et al. (2012) and Dawson et al. (2007) as long as no additional referees are present.

Second, this pattern disappears when additional referees are added. That is, the sum of the track variable and the interaction between this variable and the additional referees variable is economically and statistically insignificant with respect to the award of a yellow card as a home team ($b = 0.013 = 0.072 - 0.059$; $p = 0.549$) and as an away team ($b = 0.008 = -0.063 + 0.071$; $p = 0.663$). Indeed, the fact that we do not find evidence for referee bias as did Buraimo et al. (2010), Buraimo et al. (2012) and Dawson et al. (2007) at the level of the complete dataset is driven by most games being played with additional referees.

Third, the interaction between the track variable and the additional referees variable is statistically significant with respect to yellow cards for the away team ($b = 0.071$; $p = 0.046$) and close to weak statistical significance with respect to yellow cards for the home team ($b = -0.059$; $p = 0.129$). This variable has the expected sign in both equations: tracks are less adverse (beneficial) for home (away) teams after the introduction of additional referees.

Fourth, the additional referees variable is highly significantly associated with yellow cards for the home team ($b = 0.074$; $p = 0.000$) and weakly significantly associated with yellow cards for the away team (*b = 0.028*; *p = 0.062*). Thus, after the introduction of a fifth and a sixth referee, more yellow cards are awarded both to home and away teams. This might be explained by a more complete observation of violations on the pitch after these additional referees were introduced. From a broader perspective, it is in line with Heckelman and Yates (2003) and Depken and Wilson (2004) showing that the introduction of a second referee led to a greater enforcement of the rules in ice hockey. In a robustness check discussed below, we elaborate on potential bias in this pattern by a general tendency towards more yellow cards over time.

Before proceeding to the results for our alternative indicator of referee bias, we discuss some secondary insights gained from the estimates of our control variables. First, if the crowd increases, the probability that the home team receives a yellow card significantly declines, whereas the probability that the away team receives a yellow card significantly increases. Second, as the game proceeds towards the end, the probability that both teams receive a yellow card increases, albeit at a decreasing rate. Third, in minute 45 and minute 90, the probability of a sanction for both teams is significantly higher, probably as a result of the injury time that is added to these minutes. Fourth, as the goal difference increases, the probability that the home team receives a yellow card decreases. Fifth, if the home team received a yellow card in the previous three minutes, the probability that the away team receives a yellow card increases significantly. Interestingly, if the away team received a yellow card in the previous three minutes, the chance that the home team receives a yellow card does not increase significantly. Sixth, an extra yellow card previously received by the home team is associated with a reduction of the home team's chance to receive another yellow card and an increase in the away team's chance to receive a yellow card. Conversely, an extra yellow card previously received by the away team reduces the away team's chance to receive another yellow card and increases the home team's probability to receive a yellow card. Finally, a red card previously received by the home team is associated with a significant increase in the away team's probability to receive a yellow card, while a red card previously received by the away team is not associated with a significant increase in the home team's chance to receive a yellow card.

III.2 Alternative analysis

Table 3 presents the results of our alternative analysis. Here, we inspect referee bias in terms of favourable treatment of teams from the big five European soccer competitions, in the spirit of Dawson and Dobson (2010). Again, we test whether this operationalisation of referee bias is heterogeneous by the adoption of additional assistant referees.

Model (1) of Table 3 is an extended version of Model (1) in Table 2. That is, indicators for home and away teams from the big five competitions are added in both blocks of the regression framework. As in Dawson and Dobson (2010), the likelihood of a yellow card awarded to the home team in a certain minute is higher if the away team is from the big five competitions ($b = 0.031$; $p = 0.031$). Consistently, the likelihood of a yellow card for the away team is (weakly significantly) higher if the home team is from the big five competitions ($b = 0.025$; $p = 0.069$).

We now turn to Model (2), which includes, compared with Model (1), also the variable capturing the presence of a fifth and sixth referee as well as interaction variables between the latter variable and the indicators of home and away teams from the big five competitions. As in our benchmark analysis, we find that the overall referee bias can be decomposed in a statistically significant bias in games without additional assistant referees and an insignificant bias when such additional referees are added. On the one hand, in the reference situation of no additional referees, the coefficient of away teams from the big five competitions with respect to yellow cards for the home team ($b = 0.070$; $p = 0.010$) and the coefficient of home teams from the big five competitions with respect to yellow cards for the away team ($b = 0.028$; $p = 0.026$) are more substantial than the corresponding coefficients based on all games in Model

(1). On the other hand, in games with additional referees, away teams from the big five competitions are no longer associated with more yellow cards for the home team ($b = 0.023 = 0.070 - 0.047$; $p = 0.161$) and home teams from the big five competitions are no longer associated with more yellow cards for the away team ($b = 0.018 = 0.053 - 0.035$; $p = 0.260$). Moreover, the introduction of additional assistant referees is again associated with additional yellow cards for both home teams ($b = 0.081$; $p = 0.004$) and away teams ($b = 0.060$; $p = 0.016$).

III.3 Robustness checks

To examine the robustness of these results, we conducted numerous additional analyses. First, we re-estimated Model (1) in Table 2 and Table 3 for the subsample of games with and without additional assistant referees separately. Consistent with the analyses in the previous subsections, we find significant evidence for referee bias in line with Buraimo et al. (2010), Buraimo et al. (2012) and Dawson and Dobson (2010) only for the games without a fifth and sixth referee.

Second, we checked whether the analyses in Model (2) of Table 2 and Table 3 were not just picking up an evolution towards less referee bias over time. Therefore, we added a continuous indicator of the season both as such and in interaction with the track variable (or with the home and away team from the big five competitions variables). However, these additional variables turned out to be insignificant.

Third, as mentioned in Section II.1, the introduction of the additional assistant referees in the UEFA Europa League coincided with a new format of the group stage of this league--we are not aware of any (other) changes in UEFA instructions to referees at that moment though. Therefore, we re-estimated our analyses using only the games of the UEFA Champions League. In addition, we re-estimated our regression models after excluding (i) games without competitive value and (ii) return games in the knock-out stage (in which additional time may be added). These analyses on a restricted dataset led to similar conclusions as those described in the previous subsections.

Fourth, we re-estimated our models with team and referee fixed effects. Unlike Buraimo et al. (2010) and Buraimo et al. (2012), we did not incorporate these fixed effects in our main analyses because they result in less efficient estimates. In particular, coefficients for home and away teams from the big five competitions cannot be identified when one controls for team fixed effects. In addition, these team fixed effects remove most of the variation in our track variable. Moreover, in general, combining nonlinear models with fixed effects may yield a substantial incidental parameter problem (Lancaster, 2000; Greene, 2004). After including these fixed effects, standard errors are indeed somewhat higher compared with those in Table 2 and Table 3—this is marginally the case when only referee fixed effects are added. However, coefficient estimates are highly comparable, so our conclusions remain valid when opting for fixed effects regression models.

IV. Conclusion

In this study, we investigated whether the introduction of additional assistant referees in professional soccer is associated with lower referee bias. We examined whether the various measures of referee bias discussed in Buraimo et al. (2010), Buraimo et al. (2012) and Dawson and Dobson (2010) were homogeneous by the introduction of a fifth and a sixth referee in the UEFA Europa League (in season 2009-2010) and the UEFA Champions League (in season 2010-2011). While we found significant referee bias before the introduction of additional referees, no such evidence was found later. In addition, the introduction of additional assistants led to more yellow cards for both home and away teams. We showed that these findings are not just picking up a general time evolution towards less bias or the effect of other, parallel reforms.

Our results suggest that the investments by the UEFA in two additional assistant referees were effective to obtain better control over the game. However, our findings might not be generalised to other investments in more referees in other contexts, let alone in better trained or equipped referees. Therefore, we are in favour of future work evaluating the effectiveness of other investments by the UEFA or other governing bodies in this respect (such as recent investments in goal line technology or video assistant referees).

References

Baert, S., & Amez, S. (2018). No better moment to score a goal than just before half time? A soccer myth statistically tested. *PLoS ONE, 13*, e0194255.

BBC (2013). Pellegrini criticises referee after defeat. Retrieved 10 March 2018 from http://www.bbc.com/sport/football/22091155.

Boyko, R., Boyko, A., & Boyko, M. (2007). Referee Bias contributes to home advantage in English Premiership Football. *Journal of Sports Sciences, 25*, 1185-1194.

Brown, L. (2017). Barcelona vs PSG referee “facing UEFA demotion” after controversial performance. Retrieved 10 March 2018 from https://www.independent.co.uk/sport/football/european/barcelona-6-psg-1-reaction-referee-demotion-a7624266.html.

Buraimo, B., Forrest, D., & Simmons, R. (2010). The twelfth man? Refereeing bias in English and German soccer. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 173*, 431-449.

Buraimo, B., Simmons, R., & Maciaszszyk, M. (2012). Favoritism and referee bias in European soccer: Evidence from the Spanish League and the UEFA Champions League. *Contemporary Economic Policy, 30*, 329-343.

Dawson, P., & Dobson, S. (2010). The Influence of social pressure and nationality on individual decisions: Evidence from the behaviour of referees. *Journal of Economic Psychology*, *31*, 181-191.

Dawson, P., Dobson, S., Goddard, J., & Wilson, J. (2007). Are football referees really biased and inconsistent? Evidence on the incidence of disciplinary sanction in the English Premier League. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 170*, 231-250.

Depken, C. A., & Wilson, D. P. (2004). Wherein Lies the Benefit of the Second Referee in the NHL? *Review of Industrial Organization*, *24*, 51-72.

Dohmen, T. (2008). The Influence of social forces: Evidence from the behaviour of football referees. *Economic Inquiry*, *46*, 411-424.

Dohmen, T., & Sauermann, J. (2016). Referee bias. *Journal of Economic Surveys*, *30*, 679-695.

FIFA (2017). Laws of the game 2017/2018. Zurich: FIFA.

Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism under social pressure. *Review of Economics and Statistics*, *87*, 208-216.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal*, *7*, 98-119.

Heckelman, J. C., & Yates, A. J. (2003). And a Hockey Game Broke Out: Crime and Punishment in the NHL. *Economic Inquiry*, *41*, 705-712.

Johnston, R. (2008). On referee bias, crowd size, and home advantage in the English soccer premiership. *Journal of Sports Sciences*, *24*, 563-568.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, *95*, 391-413.

Pettersson-Lidbom, P., & Priks, M. (2010). Behavior under social pressure: Empty Italian stadiums and referee bias. *Economics Letters*, *108*, 212-214.

Rickman, N., & Witt, R. (2008). Favouritism and financial incentives: A natural experiment. *Economica*, *75*, 296-309.

Scoppa, V. (2008). Are subjective evaluations biased by social factors or connections? An econometric analysis of soccer referee decisions. *Empirical Economics*, *35*, 123-140.

The Guardian. (2017). Arturo Vidal rages at referee after Bayern Munich go out to Real Madrid. Retrieved 10 November 2017 from https://www.theguardian.com/football/2017/apr/18/arturo-vidal-bayern-munich-real-madrid.

UEFA (2014). Additional assistant referees. Retrieved 10 February 2017 from http://www.uefa.com/news/newsid=2048377.html.

UEFA (2017). 2017/18 UEFA Champions League revenue distribution. Retrieved 12 April 2018 from https://www.uefa.com/uefachampionsleague/news/newsid=2493261.html.

Van Damme, N., & Baert, S. (2018). Home advantage in European international soccer: Which dimension of distance matters? Unpublished manuscript, Faculty of Economics and Business Administration, Ghent University, Ghent, Belgium.

Table 1. Data: Summary statistics

	Mean	SD
A. Dependent variables		
Yellow card home	0.020	-
Yellow card away	0.025	-
B. Independent variables		
Track	0.174	-
Home big five	0.454	-
Away big five	0.454	-
Additional assistant referees	0.699	-
C. Control variables		
Attendance	30,083.220	19,430.880
Relative strength	5.158	14.863
Derby	0.017	-
Goal difference	0.200	1.164
Yellow card three minutes prior home	0.052	-
Yellow card three minutes prior away	0.066	-
Number of yellow cards prior home	0.644	-
Number of yellow cards prior away	0.839	-
Red card prior home	0.023	-
Red card prior away	0.030	-
N	195,120	

Note: A definition of these variables can be found in Section 2.1. No standard deviation (SD) is provided for binary variables.

Table 2. Regression results: Benchmark analysis

	(1)	(2)
Dependent variable: Yellow card home		
Additional assistant referees × Track		-0.059 (0.039)
Additional assistant referees		0.074*** (0.017)
Track	0.028 (0.018)	0.072** (0.033)
Log attendance	-0.022** (0.009)	-0.016* (0.009)
Relative strength	-0.002* (0.001)	-0.002* (0.001)
Relative strength squared	0.000 (0.000)	0.000 (0.000)
Derby	-0.010 (0.044)	-0.006 (0.043)
Minute	0.016*** (0.001)	0.016*** (0.001)
Minute squared	-0.000*** (0.000)	-0.000*** (0.000)
Minute 45	0.475*** (0.045)	0.475*** (0.045)
Minute 90	0.933*** (0.038)	0.934*** (0.038)
Goal difference	-0.025*** (0.006)	-0.025*** (0.006)
Goal difference squared	-0.010*** (0.003)	-0.010*** (0.003)
Yellow card three minutes prior home	-0.028 (0.031)	-0.028 (0.031)
Yellow card three minutes prior away	0.023 (0.026)	0.023 (0.026)
Number of yellow cards prior home	-0.016** (0.008)	-0.018** (0.008)
Number of yellow cards prior away	0.023*** (0.008)	0.023*** (0.008)
Red card prior home	0.045 (0.037)	0.042 (0.036)
Red card prior away	0.021 (0.035)	0.017 (0.034)
Intercept	-2.370*** (0.091)	-2.477*** (0.095)
Dependent variable: Yellow card away		
Additional assistant referees × Track		0.071** (0.036)
Additional assistant referees		0.028* (0.015)
Track	-0.012 (0.016)	-0.063** (0.030)
Log attendance	0.030*** (0.009)	0.034*** (0.009)
Relative strength	0.002* (0.001)	0.002 (0.001)
Relative strength squared	0.000 (0.000)	-0.000 (0.000)
Derby	0.044 (0.049)	0.049 (0.048)
Minute	0.020*** (0.001)	0.020*** (0.001)
Minute squared	-0.000*** (0.000)	-0.000*** (0.000)
Minute 45	0.493*** (0.042)	0.493*** (0.042)
Minute 90	0.962*** (0.036)	0.962*** (0.036)
Goal difference	0.009 (0.006)	0.009 (0.006)
Goal difference squared	-0.018*** (0.003)	-0.019*** (0.003)
Yellow card three minutes prior home	0.098*** (0.025)	0.098*** (0.025)
Yellow card three minutes prior away	-0.025 (0.026)	-0.025 (0.026)
Number of yellow cards prior home	0.023*** (0.008)	0.023*** (0.008)
Number of yellow cards prior away	-0.040*** (0.007)	-0.041*** (0.007)
Red card prior home	0.100*** (0.037)	0.096*** (0.037)
Red card prior away	0.052 (0.035)	0.048 (0.035)
Intercept	-2.849*** (0.092)	-2.912*** (0.096)
Log pseudo-likelihood	-39,752.151	-39,736.903
N	195,120	

Note: The presented statistics are bivariate probit model estimates and standard errors, clustered at the game level, in parentheses. A definition of the variables used can be found in Section 2.1. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance level.

Table 3. Regression results: Alternative analysis

	(1)	**(2)**
Dependent variable: Yellow card home		
Additional assistant referees × Home big five		0.021 (0.031)
Additional assistant referees × Away big five		-0.047 (0.031)
Additional assistant referees		0.081*** (0.028)
Home big five	-0.022 (0.015)	-0.031 (0.027)
Away big five	0.031** (0.014)	0.070*** (0.027)
Track	0.023 (0.018)	0.023 (0.018)
Log attendance	-0.021** (0.010)	-0.017* (0.010)
Relative strength	-0.002* (0.001)	-0.002* (0.001)
Relative strength squared	0.000 (0.000)	0.000 (0.000)
Derby	-0.014 (0.044)	-0.010 (0.043)
Minute	0.016*** (0.001)	0.016*** (0.001)
Minute squared	-0.000*** (0.000)	-0.000*** (0.000)
Minute 45	0.474*** (0.045)	0.475*** (0.045)
Minute 90	0.933*** (0.038)	0.934*** (0.038)
Goal difference	-0.022*** (0.006)	-0.022*** (0.006)
Goal difference squared	-0.010*** (0.003)	-0.011*** (0.003)
Yellow card three minutes prior home	-0.028 (0.031)	-0.028 (0.031)
Yellow card three minutes prior away	0.023 (0.026)	0.023 (0.026)
Number of yellow cards prior home	-0.017** (0.008)	-0.019** (0.008)
Number of yellow cards prior away	0.024*** (0.008)	0.023*** (0.008)
Red card prior home	0.045 (0.037)	0.040 (0.037)
Red card prior away	0.024 (0.035)	0.019 (0.034)
Intercept	-2.376*** (0.098)	-2.483*** (0.102)
Dependent variable: Yellow card away		
Additional assistant referees × Home big five		-0.035 (0.027)
Additional assistant referees × Away big five		-0.001 (0.027)
Additional assistant referees		0.060** (0.025)
Home big five	0.025* (0.014)	0.053** (0.024)
Away big five	-0.008 (0.013)	-0.003 (0.023)
Track	-0.007 (0.016)	-0.005 (0.016)
Log attendance	0.025*** (0.010)	0.028*** (0.010)
Relative strength	0.002 (0.001)	0.002 (0.001)
Relative strength squared	0.000 (0.000)	-0.000 (0.000)
Derby	0.041 (0.050)	0.042 (0.049)
Minute	0.020*** (0.001)	0.020*** (0.001)
Minute squared	-0.000*** (0.000)	-0.000*** (0.000)
Minute 45	0.493*** (0.042)	0.493*** (0.042)
Minute 90	0.962*** (0.036)	0.962*** (0.036)
Goal difference	0.007 (0.006)	0.007 (0.006)
Goal difference squared	-0.018*** (0.003)	-0.018*** (0.003)
Yellow card three minutes prior home	0.098*** (0.025)	0.098*** (0.025)
Yellow card three minutes prior away	-0.025 (0.026)	-0.025 (0.026)
Number of yellow cards prior home	0.024*** (0.008)	0.023*** (0.008)
Number of yellow cards prior away	-0.040*** (0.007)	-0.041*** (0.007)
Red card prior home	0.100*** (0.037)	0.096*** (0.037)
Red card prior away	0.051 (0.035)	0.047 (0.035)
Intercept	-2.809*** (0.098)	-2.885*** (0.103)
Log pseudo-likelihood	-39,746.522	-39,731.298
N	195,120	

Note: The presented statistics are bivariate probit model estimates and standard errors, clustered at the game level, in parentheses. A definition of the variables used can be found in Section 2.1. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance level.