

IZA DP No. 1595

**Social Pressure Influences Decisions of
Individuals: Evidence from the Behavior
of Football Referees**

Thomas J. Dohmen

May 2005

Social Pressure Influences Decisions of Individuals: Evidence from the Behavior of Football Referees

Thomas J. Dohmen

IZA Bonn

Discussion Paper No. 1595

May 2005

(substantially revised version of IZA Discussion Paper No. 755)

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Social Pressure Influences Decisions of Individuals: Evidence from the Behavior of Football Referees*

Analyzing the neutrality of referees during twelve German premier league (1st Bundesliga) football seasons, this paper documents evidence that social forces influence agents' preferences and decisions. Those, who are appointed to be impartial, tend to favor the home team as they systematically award more injury time in close matches when the home team is behind. Referees also tend to favor the home team in decisions to award goals and penalty kicks. The composition of the crowd affects the size and the direction of the bias. The intensity of social pressure as measured by the crowd's proximity to the field determines how strongly referees' decisions are influenced. Not all agents are, however, affected to the same degree by social pressure.

JEL Classification: J00, M50

Keywords: favoritism, principal-agent relationship, personnel economics

Corresponding author:

Thomas J. Dohmen
IZA
P.O. Box 7240
D-53072 Bonn
Germany
Email: dohmen@iza.org

* I am grateful to IMP AG for providing the data. I would like to thank especially Holger Rahlfs and Jörn Wendland for their cooperation. I am also indebted to the DFB, particularly to Klaus Löw, for providing detailed information on financial rewards for referees and their evaluation. I thank Armin Falk, Luis Garicano, Ben Kriechel, Winfried Koeniger, Steven Levitt, Canice Prendergast, two anonymous referees, seminar participants at Maastricht University, the University of Chicago, IZA in Bonn, participants of the 17th Annual Conference of the European Society for Population Economics in New York and of the 58th European Meeting of the Econometric Society in Stockholm for comments and suggestions on an earlier draft of the paper that was circulated under the title "In Support of the Supporters? Do Social Forces Shape Decisions of the Impartial?". All errors are mine.

1 Introduction

The question of whether and how social pressure affects preferences or actions of individual agents is of central importance for many situations of decision making in some social context. Theoretical work in different fields of the economic literature has dealt with the impact of social forces on behavior in various economic contexts (e.g., Akerlof 1980, Austen-Smith and Fryer 2003, Becker and Murphy 2000, or Bernheim 1994). Prendergast and Topel (1996) have theoretically studied the consequences of favoritism on compensation, the extent of authority and the use of bureaucratic rules in an agency model. Viewing the firm as a social institution, they argue that supervisors distort their subjective performance evaluations of subordinates when supervisors derive utility from being biased. Favoritism is harmful because inefficient assignments result from the distorted information, so that the firm deemphasizes incentive pay and relies on bureaucratic rules in pay and promotion decisions in order to thwart the adverse effects of favoritism.

This paper demonstrates empirically that social forces can induce agents who are appointed by the principal to be impartial to make biased, wrong and disputable decisions, even if such bias arguably does not provide value for them. Such empirical evidence for the relevance of social pressure and for the existence of favoritism is rare. A notable exception is the work by Garicano, Palacios-Huerta and Prendergast (2005) (hereafter termed GPP) who provided evidence based on data from two Spanish football seasons that referees systematically favor home teams by shortening matches in which the home team is one goal ahead and lengthening matches in which the home

team is one goal behind at the end of regular time.

The scarcity of hard evidence of favoritism in economic relations can probably be ascribed to the problem of obtaining the required sensitive data and difficulty for researchers to undeniably identify favoritism that leads to an inefficient allocation of resources. Economic situations are usually complex so that it is often difficult, if not impossible, to argue convincingly that favoritism results in inefficient outcomes even if evidence for favoritism is found. For example, a promotion of a worker that is based purely on favoritism might be an optimal allocation if the good work relations between the favored worker and his supervisor outweigh any productivity advantage of contestants.

This empirical study circumvents these problems by switching to the analysis of sports data, and it shows that referees in German professional football favor the home team by lengthening games in which the home team is behind, and by awarding more disputable and incorrect goals or penalties. More importantly, the analysis provides compelling evidence that social forces induce these biased decisions. I argue that the type of referee bias observed is not likely to be desired by the principal, because home teams are not favored in the same way. In fact, the data suggest that the size of referee bias depends on the composition of the crowd and on the architecture of the stadium. The crowd's physical proximity to the field, which depends on the existence of a track that separates the stands from the field, has an impact on referee decisions. Referee decisions concerning penalty decisions are also more likely not to be correct when the match takes place in a stadium without a running track.

This striking finding is both novel and fundamental, for it rules out a potential objection to the interpretation of GPP's results. A sceptic of the proposed dependency of individual decisions or preferences on social factors might argue that the type of favoritism, which is also unveiled in the German data, is provoked by the football association, the principal who appoints the referee (agent), because of a revenue maximizing motive (see also GPP). In fact, even a football association that has strong preferences for nondiscriminatory competition, but wants to maximize gate revenues, should instruct its referees to favor home teams if more spectators attend matches when the home team is more likely to win and if home bias would balance out over the season so that it would not affect the outcome of the championship.¹ The observed home bias, which is desired by the principal, would then be the result of rational and intentional decisions, but no team would be systematically favored as all home teams would be treated in the same way.

This paper reveals, however, that home teams are systematically treated differently. Since physical conditions and crowd attributes vary across home teams but differentially affect the magnitude of the home bias, it is hard to defend that the German football association (DFB) has a desire for kind of favoritism detected in the data, unless it is biased itself. By documenting the inequality in referee bias and showing that it is related to architectural conditions of the match location the paper supplements the work by GPP (2005), and it augments the literature on the relationship between social

¹In principle, both conditions are met in German professional football. First, more supporters of the home team typically attend matches and supporters are *ceteris paribus* more likely to go to the stadium when their team is more likely to win. Second, teams play each other twice a season, once as the home team and once as the visiting team.

pressure and individual decision making in a crucial way.

The paper does not allege that referees are intentionally biased because it is in their own private interest. Instead, a likely explanation for the observed behavior is that referees are emotionally influenced by the atmosphere in the stadium. Since referees have to judge a situation quickly and have little time for deliberation, their decision making process might be heavily influenced by cues in the environment. Nevill, Balmer, and Williams (2002) have provided experimental evidence that football referees are affected by the crowd's noise. They showed video-taped tackles from an English Premier League football game to qualified referees who had to decide whether or not to award a foul. One group watched the videotape without the noise of the crowd being played while the other group heard the noise. The latter group called 15.5 percent fewer fouls against the home team.

The analysis in this paper is based on data from 3519 games of 12 seasons of the German premier football league (1st Bundesliga). It first verifies that referees in German professional football, like Spanish professional football referees (see GPP), also favor the home team by lengthening the match when the home team is slightly behind at the end of the match. The key differences to GPP's findings are that the amount of home bias as measured by differential average injury time is much smaller (about 20 seconds compared to almost 2 minutes in the Spanish data) and that there is no evidence that referees shorten matches in which the home team is slightly ahead compared to drawn matches. Beyond being founded on more data, which allows to control for referee fixed effects and assess whether biased and unbiased referees coexist, this part of the analysis offers additional enrichment of the

estimation as put forward by GPP (2005).²

Some of these improvements are possible thanks to the detailed available information. Having information on the frequency of match interruptions for injury treatments on the field I can control more convincingly for the “appropriate” amount of injury time. According to Law 7 of the *Laws of the Game* as authorized by the International Football Association Board and established by the Fédération Internationale de Football Association (FIFA) “[a]llowance is made in either period for all time lost through: substitution(s), assessment of injury to players, removal of injured players from the field of play for treatment, wasting time, any other cause.”³ The allowance for time lost is at the discretion of the referee.” (FIFA, 2002, p. 19). Since all information is recorded separately for the two 45 minute periods of a match, I can effectively compare referees’ behavior at the end of both halves. Moreover, the enormous detail of the data makes it feasible to control for relative team performance during the match. But I also extend GPP’s work in several other dimensions. I investigate, for example, injury time decisions in matches that are drawn at the end of regular playing time in order to assess whether referees award more injury time in exciting matches and whether more allowance for time lost is made when the home team is more likely to

²Recently, Sutter and Kocher (2004) provided a mere replication of GPP’s work using data on only a single season (2000/2001) of German football. It seems that the authors were not aware of the existence of this work, especially the drafts that were presented earlier in various places. Sutter and Kocher constructed their data set from written information provided by a German sports magazine. A major shortcoming of their analysis is the lack of information on the number of actual injury treatments, which should be a key determinant of injury time according to the FIFA rules.

³The rules regarding the number of substitutions have changed during the observation period. Until 1994/95, teams could only substitute 2 players during a match. In 1994/95, each team could replace the goalie in addition to the two players. Since 1995/96, teams can substitute 3 players (including the goalie).

score next.

Since home bias can have many forms, I study whether referees also favor home teams in decisions concerning penalty kicks or goals. Such referee decisions have much more immediate and severe, often even decisive, consequences for the outcome of a match. Penalty kick decisions can be directed to favor a particular team; and penalty kicks lead to goals in roughly three out of four cases while only 0.05 goals are scored on average by both teams in each minute of injury time. I provide compelling evidence for the existence of biased referee decisions on penalty kicks and goals, which has, to my knowledge, not been done to date.⁴ Moreover, referees tend to make worse decisions in stadiums without a track separating the crowd and the field, which reinforces the conjecture that social pressure affects referees' decisions.

The most important contribution of this paper from a social scientist's point of view is, however, the investigation of the role of the spectators' physical proximity to the field. The evidence that referees tend to make more mistakes when no running track separates the stands from the field suggests that social forces influence individual decision making and that the intensity of social pressure is a determinant of the extent of such influence. The finding

⁴Sutter and Kocher's (2004) tentative analysis of penalty kick decisions is based on data that deviates from the official statistics. Their data is collected from a German sports magazine for the season 2000/2001 and seems to be incomplete as it contains only 55 of all 57 penalties for the home team and 21 of all 24 penalties for the visitor. The authors claim that the home team was awarded a legitimate penalty kick in 81% of cases (50 out of 62) while the visiting team was awarded a legitimate penalty kick in 51% of cases (20 out of 39). However, the detailed data from the official data base of the DFL Deutsche Fußball Liga GmbH suggests that 43 penalty kicks for the home team and 19 for the visiting team were correctly awarded, and the home team was denied 11 and the visitor 8 legitimate penalty kicks.

that the social environment matters when individuals make decisions should stimulate theoretical work in various areas of economics. The results are, for example, of great consequence for agency theory. The existence of influencing by others implies, for instance, that it can be optimal for the principal to deprive the agent of his discretion when social forces manipulate the agent to take actions that result in undesired outcomes. This should renew the interest in modelling favoritism as has previously been done by Prendergast and Topel (1996).

The remainder of the paper is organized in six sections. Section 2 describes the data in more detail. Section 3 explains the relation between the German football association and its referees. Section 4 studies injury time decisions in close and in drawn matches. Section 5 analyzes the impact of the crowd's proximity to the field and of its composition on injury time decisions of referees. Section 6 examines whether referee bias is also prevalent in goal and penalty kick decisions. Section 7 summarizes and concludes.

2 The Data

I have data on 3519 matches from 12 seasons (from 1992/93 until 2003/2004) of the German premier football league (1st Bundesliga), in which 18 teams determine the national football champion, by far the most important trophy in German professional football. A season consists of 34 rounds with 9 matches each (306 matches per season) so that teams play each other twice a season, once in the first 17 rounds and once in the last 17 rounds when the

status of home and visiting team is reversed.⁵ Typically, one round is played per week. Usually six to seven matches took place on Saturdays during the observation periods with the remaining two or three being scheduled for play on Fridays or Sundays.

The winner of a match was awarded two points until the 1994/1995 season, but has received three points since the 1995/1996 season. No points are awarded to the losing team, while both teams receive one point if the match ends in a draw. The league ranking is based on the number of points. Among teams with the same number of points the one with the biggest difference between goals scored and goals received is higher ranked. If teams are identical in this respect as well, the team that has scored most goals is higher ranked in the league table. The team ranking highest at the end of the season wins the championship. The three lowest ranked teams are relegated to the 2nd Bundesliga being replaced by the three top ranked teams of this second division. There are also incentives to finish 2nd to 5th, or even lower, as these ranks qualify for participation in various European Cups, of which the UEFA Champions League is the most lucrative and prestigious.⁶

The data were made available to me for scientific research by IMP – Innovative Medientechnik und Planung AG, a company that maintains the official football data base of the DFL Deutsche Fußball Liga GmbH.⁷ IMP AG sends several observers to each league match who record about 2000 actions per match — including all goals, shots on goal, tackles, passes, corner kicks,

⁵I have only data for the first half of the last season.

⁶See Dohmen (2003) for details.

⁷All statistics that are published on the website of the DFL Deutsche Fußball Liga GmbH are provided by IMP AG.

every single ball contact, yellow and red cards. These match statistics are provided separately for both periods of the match together with the allowance for time lost at the end of the first period and at the end of the second period.

I also examine data on referees' decisions that concern (potential) goals and penalty kicks. Decisions concerning important events such as penalty kicks or goals are re-evaluated after each match using video recordings. The specialists of IMP AG consider about 30 pieces of detailed information when they judge whether the referees' decisions to award goals or penalty kicks were correct, disputable or wrong. Decisions are called correct if they agree that the decision was correct; they are labeled as wrong if they consent that the decision was wrong, and as disputable if the evaluators' verdict is not unanimous.

Data on the correctness of referees' decisions are available for all but three of the 10166 goals that have been awarded from the start of the season 1992/1993 until the end of the first half of the season 2003/2004 and for all 857 penalties that have been called since 1993/1994, when IMP started to collect this information, until the end of the first half of the season 2003/2004. Data on the correctness of decisions not to award a goal are collected since 1993 and cover 463 critical situations in which a goal was not accepted.⁸ Information on the decision not to award a penalty kick is only recorded since the start of the season 1998/1999. Until January 2004, 892 critical situations, in which suspicious action took place in one of the penalty areas, had been evaluated. Additional data used in this study comprise the date,

⁸Information concerning the correctness of the decision is missing for one of the 463 goals that have not been awarded since 1993/1994 and for three of the 10166 goals that have been awarded since 1992/1993.

destination and outcome of the match, the number of spectators, the referee's name, as well as the information on various events and their timing during the game, including goals scored, or yellow cards and red cards called.

3 The Referees and the German Football Association

German premier football league referees are appointed by the German Football Association (DFB), which is organized in 5 regional associations (Regionalverbände) consisting of 21 different smaller associations (Landesverbände). Referees must be a member of a club that is organized in a regional association. Requirements for becoming a referee include the completion of a course and passing a written and physical test. Referees usually start in the lowest division and after having been promoted to referee in the Landesliga, the sixth division, they can be promoted at most one division each year if judged as qualified by official observers. Since 1995, the number of Bundesliga referees is limited to 22. Before that, the DFB appointed up to 36 referees each season. In total, 73 umpires have refereed at least one of the 3519 1st Bundesliga games for which data is available.

Financial incentives for refereeing a Bundesliga match are large and have risen substantially over time. From July 1992 (i.e. when the observation period begins) until July 1997 referees received 2500 DM per game, from August 1997 until July 2000 they were paid 4000 DM per game, and since August the reward is 6000 DM or 3067.75 Euro per game. In addition,

travel expenses (including hotel and transportation) have been covered since 1992.⁹ The agent, i.e. the referee, clearly has strong pecuniary incentives to be reappointed by the principal, the DFB, who commands referees to be impartial. Being biased is not in the referee's private interest, because it leads, if detected, the DFB to sack a referee.¹⁰

The performance of referees is monitored and judged by an official observer of the DFB referee committee, who attends the match in the stadium and fills a performance evaluation form afterwards. The first aspect addressed in this evaluation is the referee's appearance during the game, in particular whether the referee was decided, secure, and had the courage to take unpopular decisions, or whether he was impressed by complaining players. As all these points concern how referees deal with the social environment on the field, performance in a social environment seems to be a critical issue for the evaluation and appointment of referees. Surprisingly, these evaluations pay no explicit attention to whether the referee's preferences are affected by the social environment created by the crowd.¹¹ This is at the center of attention in this paper.

The second and third points of the evaluation form concern how well the referee interpreted the Laws of the Game and additional instructions. Item 3.3 of the evaluation form concerns the correct translation of instructions into action. This includes whether the referee's allowance for time lost is

⁹Before July 1992, referees received a monthly lump-sum payment of 520 DM, as well as a lump-sum payment of 100 DM per match to cover travel expenses. The information on referee remuneration was provided by the DFB.

¹⁰This was confirmed by the DFB upon request.

¹¹See Becker and Murphy (2000) for a treatise on how social forces shape the preferences of agents.

adequate, as is clearly stated by Linn (2003). Linn (2003) also interprets “wasting time” as including delays in carrying out throw-ins, corner kicks, goal kicks, or free kicks, as well as impeding a quick restart of play. Referees can lose their position due to poor performance. According to a DFB official, referees in lower level leagues are regularly relegated or dismissed. Dismissal or relegation is less common for Bundesliga referees. The DFB official attributed this to selection and incentives arguing that Bundesliga referees are “simply better referees as they have been promoted from lower leagues in several rounds upon very positive judgements of DFB observers” and that “much more is at stake for Bundesliga referees”. The DFB official was reluctant to state why referees were not reappointed in the past. Unfortunately, the data do not allow me to distinguish between voluntary quits and dismissals.

4 Empirical Analysis of Decisions on Injury Time

4.1 Evidence of Favoritism in the Raw Data

In order to judge whether there is evidence for systematic bias in the allowance for time lost in the German data and whether the social environment created by the crowd affects this bias, I follow the idea of GPP and classify matches according to the score difference at the end of regular playing time in each of the two halves of the match, i.e. after 45 and 90 minutes. This score difference is defined as the number of goals scored by the home team

minus the number of goals scored by the visitor.¹² Figure 1 plots kernel density estimates of the distributions of allowance for time lost in the second half conditional on the score. The Figure reveals some interesting results and suggests evidence of favoritism towards the home team in the raw data. Most additional time is awarded when the home team is one goal behind, and second most when the home team is one goal ahead. The distribution of injury time awarded when the home team is one goal behind in score stochastically dominates the densities of injury time in all other situations. Interestingly, distributions of allowance for time lost in the first half — i.e. when there is less at stake with another halftime to score — do not differ so markedly by score.¹³

The magnitude of German referees’ favoritism in these close matches as measured by the size of the referees’ bias — a mean difference of 21 additional seconds awarded when the home team is one goal behind rather than one goal ahead — is much smaller than that reported by GPP for Spanish football referees, who find a bias of about 113 seconds. Contrary to GPP’s findings, I find no evidence that referees shorten close matches in which the home team is ahead relative to drawn matches.¹⁴ Referees, instead, award more injury time when the home team is ahead by one goal than in drawn matches. This

¹²Duggan and Levitt (2002) propose a similar approach of looking at the “relevant margin” in their study of corruption in Sumo wrestling tournaments which commonly involve 66 wrestlers participating in 15 bouts each. There are strong incentives for achieving eight wins because a winning record guarantees a rise in the rankings. Duggan and Levitt examine whether winning probabilities are exceptionally high for wrestlers who are close to an eighth victory, e.g. those having a record of seven wins prior to the 15th bout.

¹³The estimated distributions are not presented here, but are available from the author upon request.

¹⁴The average injury time differential between drawn matches and matches in which the home team is one goal behind is similar to that reported by GPP.

result might be explained by the fact that more supporters accompany their team to away matches in Germany, where home cities of opponents are less far apart than in Spain. This issue is addressed below.

The Figure makes it also obvious that games last longer the closer the score at the end of regular playing time is. One possible explanation for this finding is that referees lengthen exciting games in order to serve the preferences of the spectators who like suspense (see Chan, Courty and Li, 2003 for a theoretical model). Matches in which one of the teams is ahead in score by two goals and drawn matches last strikingly longer than decided games in which one of the teams is leading by three or more goals. The much smaller variance of the injury time distribution in decided matches is also suggestive of the fact that referees' decisions on injury time depend on the excitement of matches.

4.2 A Closer Inspection of Favoritism in Close Games

In order to rule out that differences in injury time at distinct scores, which are observed in the raw data, do not result from differences in biased discretion of referees but stem from differences in factors that determine the amount of allowance for time lost according to the rule, I control such objectively influencing variables.

Table 1 documents that referees allow 20 seconds more additional time on average when the home team is 1 goal behind compared to games in which the home team is leading by 1 goal (see column [1]). This difference in mean duration of matches is extremely robust and does not disappear even if I

control for factors that should affect, according to the *Laws of the Game*, the amount of extra time awarded at the end of each half. The size and statistical significance of the home bias as measured by the coefficient on the “Home Ahead” dummy, which is 1 if the home team is one goal ahead and 0 if the home team is one goal behind, remains virtually unchanged when more controls for the number of treatments, substitutions and cards, as well as other potentially confounding factors are added to the regression model (see columns [2] - [7]). Although the number of actual injury treatments on the field has the expected positive effect on the duration of injury time, the effect is small (see column [2]) and it becomes weaker and statistically insignificant when more controls are added (columns [3] - [7]). A similar observation holds for the number of substitutions. The number of cards, which can be viewed as a proxy for the intensity of the match and is thus expected to be positively correlated with appropriate allowance for time lost, has the expected sign. A yellow card leads to about 6 seconds of extra-time. This result is also robust. Controls for the teams’ relative strength and performance in the match do not affect the estimates in any notable way.

Allowance for time lost in the second half is on average bigger after 1997. This suggests that the revision of the Laws by the International Football Association Board (IFAB) in 1997 induces referees to prolong the allowance for time lost, even though the DFB neither introduces a Fourth Official, nor changed its instructions for referees.¹⁵ The size of the home bias is, however,

¹⁵The revised Laws permit the appointment of a Fourth Official, an off-field observer who assists the referee. In international games, the Fourth Official indicates the number of the minimum minutes of injury time in the last minute of regular playing time. However, the DFB has not introduced a Fourth Official, an off-field observer who assists the referee, until the year 2003.

not affected by controls for season-specific effects.

Separate univariate regressions of extra time on a score-difference dummy (i.e. as in column [1] of Table 1) for each season indicate that games, in which the home team is behind, last significantly ($p\text{-value} < 0.05$) longer in seven out of the 12 seasons.¹⁶ In these seven seasons the estimated difference in additional time when the home team is behind ranges from 43.3 seconds in the 1994/95 season to 21.0 seconds in the 1997/1998 season. The coefficient estimates for the 2000/2001 and 2002/2003 are comparable in magnitude (25.3 and 17.3 seconds) but insignificant at the 5 percent level ($p\text{-values}$ of 0.164 and 0.053). A small (8 seconds) but insignificant effect is found for 1998/1999. In 1992/1993 and 1995/1996 games in which the home team is ahead by one goal last about 8 seconds longer, but this estimate is insignificant ($p\text{-value} > 0.4$).¹⁷

The home bias in awarding injury time is generally much smaller (and marginally significant) at the end of the first half, while various explanatory variables including actual treatments and the number of yellow cards have

¹⁶It is also notable that the pattern observed in Figure 1 is found in the season-specific distributions. The injury time distribution of matches in which the home team is one goal behind, lies to the right of all other distributions in all but the 1992/1993 and the 1995/1996 seasons.

¹⁷This is remarkable since the 1995/96 season marks the change in the rewards for winning, as a victory yields 3 points rather than 2 points from that season onwards. If referees favor the home team, they should become more likely to shorten matches in which the home team is 1 goal ahead, because the home crowd is expected to put more intense pressure on the referee. This should on average increase the difference in second-half injury time in close games. GPP report evidence for this conjecture finding a negative coefficient for the interaction between the score difference (which takes the value -1 when the home team is one goal behind and 1 if it is one goal ahead) and a dummy which is one in the season in which the 3-points rule applies. Replicating their analysis using the 1994/95 season (the one just preceding the change in rewards) and the 1995/96 season, I find a significant positive (about 22 seconds) effect of the interaction term. It should be noted, however, that if I had picked other pairs of pre- and post-change seasons, I would have found exactly the result that GPP report.

similar effects (see Table 2). This result supports the conjecture that the referee is influenced by social pressure exerted by the crowd, which is probably less intense after the first half when fans have still hope and faith that their team at least ties the score in the second half.

Social pressure exerted by the crowd is likely to increase not only towards the end of the game, but also towards the end of the season, when more is at stake so that the difference in injury time awarded in close matches is presumed to be larger towards the end of the season. While I do find that this difference widens by about 10 seconds in the last 5 rounds compared to earlier rounds, the effect is statistically insignificant.

In order to assess whether agents are influenced by the social environment to different degrees, I compared the size of the average individual home bias in close matches across referees. I defined this average individual home bias as the difference between a referee's average injury time awarded when the home or the visiting team is one goal ahead at the end of regular time that can be attributed to the referee's discretion. This discretionary part of awarded injury time is defined as the residual from regressions of second-half injury time on controls for the amount of injury time that should objectively have been awarded (number of treatments, cards and substitutions in the second half) as well as on season dummies. I restricted the comparison to referees who have umpired at least five matches in which the home team was one goal ahead at the end of regular time and five matches in which the visitor was one goal ahead. Thus I only considered fairly experienced referees. The point estimates of the individual biases give rise to the conjecture that biased and unbiased referees coexist, but strong statistical support is lacking as the

standard deviations of the amount of injury time awarded are too large to make strong statements at sensible significance levels.

4.3 Evidence from Drawn Matches

If referees favor the home team, we expect them to lengthen drawn matches when the home team is more likely to score next. Table 3 finds some evidence for this conjecture. Allowance for time lost in drawn matches is larger when the home team has had relatively more shots on goal than the visiting team and smaller if the visiting team has had more shots on goal and was presumably superior during the match (see column [1]). Other measures like the relative number of crosses and the relative number of tackles won are also supportive having a positive sign, but these measures are not statistically significant. The weakness of measures such as the relative number of shots on goal or tackles won is that they can only proxy for average relative strength during the entire game. To the extent that these relative measures are still a good indicator of a team’s superiority during injury time, the estimates indicate that the match is likely to last longer when the home team is relatively stronger. In addition, drawn matches in which the home team tied the match, i.e. when the home team is likely to be stronger towards the end of the game also tend to last longer (column [2]). This effect disappears, however, if the sample is restricted to matches in which both teams scored at least once when simultaneously controlling for relative strength.

An intriguing result is documented in columns [3] and [4] of Table 3: referees seem to lengthen more exciting drawn matches — just as is true for

exciting matches in which one of the teams is leading (cf. Figure 1). Games in which neither team has scored (0:0 Score) end about 9 to 11 seconds earlier on average than other matches that end in a draw at the end of regular time. Moreover, matches with more shots on goal, more tackles and more crosses last longer (column [4]).

5 Social Pressure from the Crowd?

So far, it has been taken for granted that the crowd in the stadium prefers the home team to win and likes exciting matches to last longer. I have conjectured that the crowd creates a particular atmosphere, i.e. social environment, in the stadium (e.g. by shouting, complaining, etc.) that pressures the referee to adjust his decisions according to the crowd's preferences. These preferences and the social pressure that is created by the crowd depend presumably on the relative numbers of home team and visiting team supporters, i.e. on relative preferences for a home team victory. It would also seem likely that the social atmosphere also depends on how vigorously supporters can express their preferences. Such intensity of pressure is likely related to the size of the crowd, the attendance-to-capacity ratio, and the proximity of supporters to the field.

In order to explore how the social environment affects referees, I augment the regression models from Table 1, columns [6] and [7], by controls for the absolute number of spectators, the attendance-to-capacity ratio, and interactions of these variables with an indicator for the score difference ("Score Difference") that equals 1 if the home team is one goal ahead and -1 if it is

one goal behind. At the same time, exploiting the fact that a running track separates the field and the ranks in some stadiums, I estimate the regression models separately for two sub-samples: matches that were played in stadiums with a running track (columns [1] and [2] of Table 4) and matches that took place in stadiums without a running track (columns [3] and [4] of Table 4), because the existence of a track might affect to what extent the crowds’ articulated preferences “reach” the referee.

It is intriguing that the estimated difference in injury time awarded in close matches is statistically insignificant in stadiums with a running track, but much larger and significant in stadiums in which the crowd is physically closer to the referee. The estimated home bias amounts to almost 1 minute (twice the coefficient on “Score Difference”) of additional time when the home team is one goal behind rather than ahead. Neither the number of spectators, nor the attendance-to-capacity ratio seem to have a significant effect on the length of injury time. More important yet, both factors do not seem to affect the home bias significantly, as the estimated coefficients for the interaction terms with the score difference suggest.¹⁸

Alternative specifications, in which controls and interaction terms for attendance and attendance-to-capacity ratio are dropped consecutively, reveal that a larger crowd and attendance-to-capacity ratio is required in stadiums with a running track to induce referee bias.¹⁹ When no running track sepa-

¹⁸However, a higher attendance-to-capacity ratio tends to reduce the home bias slightly in stadiums without a running track but to increase it in stadiums with a running track.

¹⁹In fact, the apparent difference in the size of the home bias in stadiums with and without a running track disappears when both interaction terms are dropped from the regression. This also corresponds with an inspection of the raw data, because separate kernel density estimates of injury time in close matches for stadiums with and without a track suggest a similar magnitude of home bias in both situations.

rates the stands from the field, more spectators tend to reduce the otherwise large home bias. The finding that the number of spectators has different effects on referees' home bias across stadiums might stem from differences in crowd composition. For example, we would expect the home bias to fall with a higher attendance when it is predominately caused by more supporters of the visiting team being present in the stadium.

Unfortunately, the data do not contain information about crowd composition. It is generally known, however, that supporters of teams are usually concentrated in the proximity of the team's home city. This is also reflected in the geographical distribution of fan clubs of German Bundesliga teams. It is also commonly accepted that fans are more likely to attend their team's away matches the shorter is the travelling distance to the opponent's stadium, so that we would expect more visiting team supporters to attend the match the closer the travelling distance between the home towns of the opponents.²⁰ Based on coordinates I calculated the distance between home cities of the opponents of a game.²¹ On the other hand, the number of visiting team supporters is also likely to depend on the team's nationwide popularity. I proxy for a team's general popularity by the number of fans it attracts when playing away from home. Therefore, I calculate for each team the average attendance-to-capacity ratio in the stadium when the team is the visitor.²² This proxy is expected to be correlated with the number of fans that support

²⁰GPP also exploit this fact in their analysis.

²¹This was done using the program on the webpage http://www.koordinaten.de/online/dist_wel.shtml.

²²An inspection of the popularity ranking of teams according to this measure suggests that it is a surprisingly good measure. The highest four ranked teams are FC Bayern München, Borussia Dortmund, Schalke 04 and Borussia Mönchengladbach.

the team when it plays away from home and should capture especially supporters of the visiting team that live in the vicinity of the match location.²³ It is reassuring that the constructed variable is highly correlated with the number of fan clubs of the different teams.

If these proxies really affected the number of visiting team supporters, they should also affect total attendance, i.e. if the visiting team is a very popular team or if the visiting team's home city is not too far away, the number of visiting team supporters and hence the total number of fans in the stadium should rise. I regress absolute deviations from average attendance of the home team's home matches on the proxy for the visiting teams' nationwide popularity and the distance between the opponents' home cities.²⁴ Since the effect of distance on the willingness to travel might not be linear, I estimate a piecewise constant step function to capture the effect of distance on attendance. The results in Table 5 indicate that both the popularity of the visiting team (as approximated by its ability to fill stadiums in away matches) and the distance between cities of opposing teams have a significant impact on deviations from average attendance. This implies that popularity and proximity to the stadium affect attendance and hence the number of supporters of the visiting team. The effect of distance is non-linear. Compared to the reference group in which opponents' home towns are 300–450 km apart, attendance rises to 4638 (865) above regular attendance if the

²³For example, this proxy is expected to cover Bayern München supporters who live in the North of Germany rather than close to Munich and support Bayern München when they play in Hamburg.

²⁴To avoid the introduction of a correlation between the error term of the regression and the explanatory variables, I do not include observations with the respective home team and visiting team when calculating average attendance and the attractiveness of the visiting team.

visiting team’s hometown is located less than 150 km (150–300 km) away, while there is no statistically significant difference in attendance when teams’ home towns are even further apart.

In order to assess whether the size of the home bias depends on the composition of the crowd, I augment the regression model from Table 1, column [7], by an interaction term between the score difference and a dummy variable that equals 1 if the visiting team’s relative popularity is in the upper third of the distribution. Relative popularity is calculated as the ratio of the team’s general popularity (i.e. the average attendance in matches played away from home, see above) to the general attendance of matches on the home team’s turf. The latter is calculated as the average attendance-to-capacity ratio in home games. To account for the impact of distance between the cities of the opponents, I estimate the regression model separately for matches in which both teams come from cities that are less than 150 km apart, and cities which are further apart. In the former case, it is suspected that a large fraction of the crowd supports the visiting team.

The regression results in Table 6 show that the home bias is negligible and statistically not different from zero if the distance between the cities is less than 150 km, unless the visiting team is relatively popular (see column [1]). When the cities are further than 150 km apart, the home bias is significant and large, especially when the visiting team’s general popularity is not particularly large (about 24 seconds, see column [2]). If the visiting team is popular, meaning that it generally attracts many fans in matches played away from home relative to the average attendance in the home team’s stadium, the bias is slightly, but insignificantly, smaller.

6 More Evidence from Decisions to Award Goals and Penalty Kicks

6.1 Goals

Biased referee decisions concerning goals have the most severe and immediate consequences for the outcome of a match. It is therefore intriguing to assess, whether referees also tend to favor the home team on this dimension.

Table 7 summarizes the correctness of decisions when referees have awarded goals for home teams and for visiting teams. The raw data on decisions concerning goals that have been awarded suggest that referee decisions are more likely to be wrong or disputable when the goal is awarded for the home team: 95.99% of goals scored by the visiting team but only 95.05% of goals scored by the home team are rightly awarded. This 0.94 percentage point gap is statistically significant at the 5 percent significance level and although it might seem small, it is far from being negligible since home teams are awarded 6026 goals in total during the period under consideration.

A simple probit model confirms that the probability of a correct referee decision as opposed to a wrong or disputable one is significantly less likely if the home team scored. An augmented model also reflects that referee decisions concerning goals that have been awarded are significantly less likely to be correct when the match takes place in a stadium without a track that separates the stands from the field. There is no evidence, however, that home teams profit significantly from worse refereeing in stadiums with a track. A multinomial logit model for the probability that the decision is

correct, disputable or wrong indicates that home teams are more likely to be granted disputable goals and that decisions in stadiums with a track are significantly less likely to be wrong.

Strikingly and consistent with the evidence for biased referee decisions concerning the allowance for time lost, awarded goals are less likely to be correctly awarded when a team is behind in score. This is especially true for the home team.²⁵ Probit estimates prove that a goal is particularly likely to be awarded incorrectly for the home team when it is behind by one or two goals; and results from multinomial logit regressions show that decisions in these cases are likely to be disputable rather than plainly wrong.²⁶

Clearly, evidence of home bias that comes only from the correctness of referee decisions when a goal was awarded is imperfect since situations in which a goal was scored but not awarded are neglected. Information on the second type of decisions is available only since the start of the season 1993/1994. Considering all situations in which a goal has been scored, goals can be classified as legitimate, disputable or illegitimate in the following way: legitimate goals are those goals that should have been awarded; illegitimate goals are those goals that should not have been awarded; and disputable goals are those goals for which there is no consensus classification as either legitimate or illegitimate. Table 8 displays the number of goals that referees have actually awarded for home and visiting teams in each of these categories. The first row of the table shows that there is no indication that referees award

²⁵This finding is interesting in another respect: If supporters like suspense they are likely to prefer close scores and hence have a preference for the score to narrow.

²⁶The latter result squares with the fact that blatantly wrong decisions might have too strong negative repercussions for referees.

the visitor fewer legitimate goals. There is tentative, but not statistically significant, evidence that the visiting team is awarded fewer disputable goals. Home teams are, however, awarded significantly (at 10 percent level; p -value = 0.07) more illegitimate goals than visitors. There is evidence that home teams are most likely to be awarded a disputable or illegitimate goal when being behind in score (see Table 9).

6.2 Penalties

Home teams are also favored in penalty kick decisions. A first piece of evidence is the fact that a smaller percentage of penalties awarded for the home team is rightly awarded (65.20% vs. 72.57%, see Table 10).²⁷ Observed differences in the frequencies of wrong, correct, and disputable decisions are statistically significant.²⁸ Remarkably, the fraction of wrong or disputable decisions in favor of the home team is largest when the home team is behind in score. However, it must be noted that referees also take more disputable decisions in favor of the visitor, when the visitor is just one goal behind.

A remarkable finding is also that more penalties are awarded in stadiums without a running track: From August 1993 until December 2003, 0.281 penalties per match in matches that took place in stadiums without a running

²⁷Home teams were awarded roughly twice as many penalty kicks as visitors (569 vs. 288) during the seasons from 1993/1994 until the first half of the season 2003/2004. This difference largely results from different strategies. There is simply more play in the visiting team's penalty area, which is confirmed by the fact that home teams are also more often involved in critical situations in which no penalty kick was awarded.

²⁸The null hypothesis that observed differences are purely due to random chance is rejected by a *chi-square* test at the 10% significance level. The $\chi^2(2)$ statistic equals 4.75. Grouping wrong and disputable decisions into one category and testing for the significance of differences in the observed frequencies of correct and not correct decisions for the home and visiting team yields a $\chi^2(1)$ test-statistic of 4.41 rejecting the null hypothesis that referees decide in the same way for both teams at the 5% significance level.

track as opposed to 0.254 per match in matches played in stadiums with a running track. Strikingly, given that a penalty was awarded, the decision was more likely to be correct when the game took place in a stadium in which a track separates the stands from the field.²⁹

Table 11 discloses that visiting teams are also more likely to be denied a legitimate or a disputable penalty kick. When a penalty kick should have been awarded but was not — data for this kind of analysis is only available from 1998 onwards — referees were more likely to penalize the visiting team: the visiting team was wrongly denied a legitimate penalty kick in 35.75% of cases, but the home team only in 29.59% of cases, an unequal treatment that is statistically significant (see 1st row of table 11). In case of disputable penalty kick decisions the evidence for home team favoritism is even clearer: home teams are awarded 28.67% of disputable penalties but visiting teams only 20.27%.

Estimates in Table 12 based on the data from August 1998 until December 2003, which include information on critical situations in which no penalty kick was awarded, provide additional, although somewhat weaker, evidence for the existence of a home team effect and an impact of physical distance between the crowd and the field: Home teams are significantly more likely to be awarded a penalty when it should be awarded and when the situation is disputable. In addition, a penalty kick is significantly more likely to be awarded when it should objectively be awarded in stadiums with a track.

²⁹A probit model for the probability that a penalty awarded between August 1993 and January 2003 was correctly awarded indicates that referees are 10% less likely to decide correctly when the game takes place in a stadium without a track. This effect is statistically significant at the 1%-level.

Under these conditions, referees also tend to award fewer penalty kicks in situations that are disputable or in which a penalty kick is not justified.

7 Conclusion

I have shown that social forces can influence agents to make decisions that are not in their private interest. In particular, I have provided empirical evidence that referees' decisions in German professional football are influenced by the preferences of the crowd. In line with the preferences of the crowd, referees lengthen exciting games and favor the home team by allowing most additional time when the home team is behind by one goal, especially when the crowd largely consists of home team supporters. When relatively many supporters of the visiting team are attending the game, the referee's bias in favor of the home team becomes smaller. Referees also prolong a drawn match when the home team is more likely to score next, and favor the home team in matters of penalty kick and goal decisions. Referees' reaction to the preferences of the crowd are more pronounced the smaller the physical distance between the referee and the crowd is, so that it can be concluded that the social atmosphere in the stadium leads referees into favoritism although being impartial is optimal for them in order to maximize their re-appointment probability.

These findings have important implications for agency theory. The evidence of the existence of favoritism and the fact that agents' decisions can be influenced by social pressure, which might lead to undesirable outcomes, can, for example, explain the existence and importance of formal rule struc-

tures in large organizations. How agents decide under social pressure must therefore be a major concern of economists.

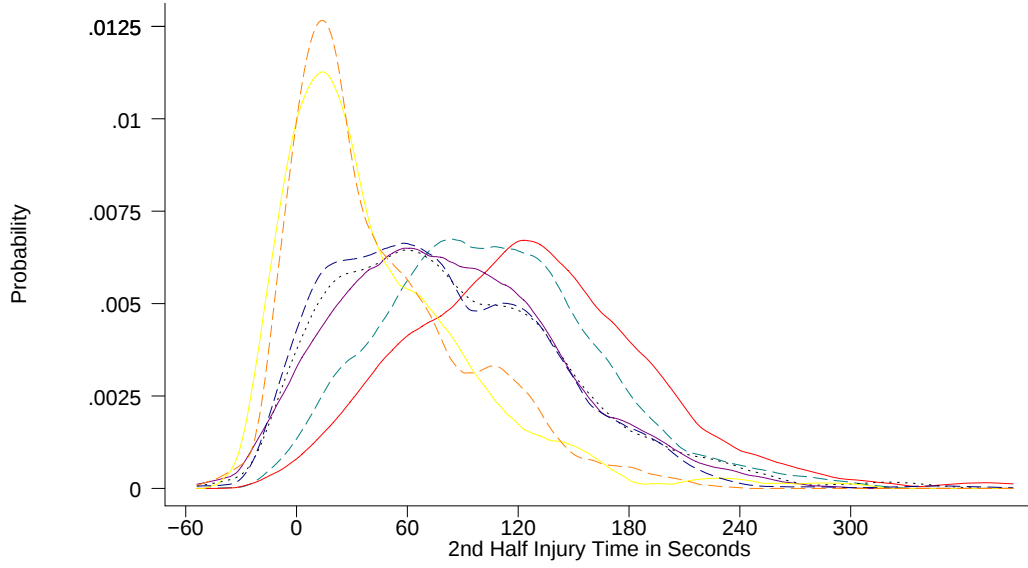
References

- AKERLOF, G. A. (1980): “A Theory of Social Custom, of Which Unemployment May Be One Consequence,” *Quarterly Journal of Economics*, 94(4), 749–775.
- AUSTEN-SMITH, D., AND R. G. FRYER (2003): “The Economics of ‘Acting White’,” unpublished manuscript, Harvard University.
- BECKER, G. S., AND K. M. MURPHY (2000): *Social Economics. Market Behavior in a Social Environment*. Harvard University Press, Cambridge, MA.
- BERNHEIM, D. B. (1994): “A Theory of Conformity,” *Journal of Political Economy*, 102(5), 841–877.
- CHAN, W., P. COURTY, AND H. LI (2003): “Suspense,” Working Paper, London Business School.
- DOHMEN, T. J. (2003): “In Support of the Supporters? Do Social Forces Shape the Decisions of the Impartial,” IZA Discussion Paper No. 755.
- DUGGAN, M., AND S. D. LEVITT (2002): “Winning Isn’t Everything: Corruption in Sumo Wrestling,” *The American Economic Review*, 92(5), 1594 – 1605.
- FÉDÉRATION INTERNATIONALE DE FOOTBALL ASSOCIATION (FIFA) (2002): *Laws of the Game*. Fédération Internationale de Football Association, Zurich, Switzerland.
- GARICANO, L., I. PALACIOS-HUERTA, AND C. PRENDERGAST (2005): “Favoritism Under Social Pressure,” *Review of Economics and Statistics*, 87(2).
- LINN, G. (2003): “Ziel und Ablauf von Schiedsrichterbeobachtungen,” *DFB Schiedsrichter-Zeitung*, 1.

- NEVILL, A., N. BALMER, AND A. WILLIAMS (2002): “The influence of crowd noise and experience upon refereeing decisions in football,” *Psychology of Sport and Exercise*, 3, 261–272.
- PRENDERGAST, C., AND R. TOPEL (1996): “Favoritism in Organizations,” *Journal of Political Economy*, 104(5), 958–978.
- SUTTER, M., AND M. G. KOCHER (2004): “Favoritism of Agents — The Case of Referees’ Home Bias,” *Journal of Economic Psychology*, 25(4), 461–469.

Figures and Tables

Figure 1: KERNEL DENSITY ESTIMATES OF THE SECOND-HALF INJURY TIME FOR VARIOUS SCORE DIFFERENCES



Lines are dashed for the kernel density estimates of the injury time distributions in matches in which the home team is ahead in score and solid when the home team is behind. The kernel density estimates of the injury time distribution in drawn matches is plotted with a dotted line. The lines are light (yellow and orange) if the score difference exceeds 2 goals, dark (purple and dark blue) if the score difference is 2 goals, and medium light (red and green) if the absolute value of the score difference is 1.

Table 1: LENGTH OF SECOND-HALF INJURY TIME IN CLOSE GAMES

The dependent variable is the length of second-half injury time (in seconds) in matches in which one team was leading by one goal after regular time. “Home Ahead” takes the value 1 if the home team is one goal ahead and zero if the home team is one goal behind.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Home Ahead	-20.947*** [4.293]	-21.229*** [4.393]	-22.142*** [4.168]	-21.129*** [4.250]	-20.601*** [4.501]	-21.896*** [4.493]	-21.66*** [4.478]
No. Treatments		-0.13 [0.230]	-0.187 [0.219]	1.256 [0.490]	1.364 [0.460]	1.197 [0.477]	1.262 [0.447]
No. Substitution		2.217** [1.089]	1.915 [1.168]	1.334 [1.540]	1.562 [1.602]	1.267 [1.649]	2.172 [1.421]
No. Fouls		0.52** [0.252]	0.284 [0.251]	-0.754** [0.363]	-0.819** [0.360]	-0.756** [0.350]	-0.558* [0.311]
No. Yellow Cards			6.667*** [0.998]	6.064*** [0.924]	6.041*** [0.929]	6.592*** [1.037]	6.118*** [1.162]
No. 2nd Yellow			3.694 [3.952]	3.053 [3.716]	2.871 [3.899]	3.404 [4.210]	1.056 [4.231]
No. Red Cards			11.238* [6.527]	11.923** [5.920]	12.703** [5.749]	14.483** [5.835]	12.519** [5.176]
Season Dummies	No	No	No	Yes	Yes	Yes	Yes
controls	No	No	No	No	Yes	Yes	Yes
rel. strength							
Home Team	No	No	No	No	No	Yes	Yes
Fixed Eff.							
Visitor Team	No	No	No	No	No	Yes	Yes
Fixed Eff.							
Referee	No	No	No	No	No	No	Yes
Fixed Eff.							
Constant	126.753*** [4.001]	97.439*** [11.131]	91.178*** [11.264]	93.389*** [11.779]	74.724*** [14.831]	57.568*** [18.688]	73.615*** [16.947]
Observations	1166	1166	1166	1166	1117	1117	1117
R-squared	0.03	0.04	0.07	0.14	0.15	0.21	0.31

Notes:

1. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5, or 1 percent level respectively. Standard errors allow for correlation between observations of the same referee.
2. The control variables “No. Treatments”, “No. Substitution”, “No. Yellow Cards”, “No. 2nd Yellow”, “No. Red Cards”, sum the total number of respective events in the second half. The variable “No. Fouls” sum the total number of fouls during the game. Information on the number of fouls is not available for the two halves separately. When the number of any such event for the home and visiting team is controlled for separately, differences in coefficient estimates are very small and always statistically insignificant. Moreover, all other parameter estimates are robust to such specification changes.
3. Controls for relative strength include the relative number of tackles won, shots on goal, fouls, corner kicks and crosses as well as the absolute difference in rankings before the match and the ranking of the home team. None of these variables are significant. The hypothesis that controls for relative strength are jointly significant is strongly rejected by a Wald test at any reasonable significance level (p-value exceeds 0.7 in all specifications). Measures on which controls of relative strengths are based are missing for the majority of matches in the season 1992/93, which explains differences in observations used across columns.

Table 2: LENGTH OF FIRST-HALF INJURY TIME IN CLOSE GAMES

The dependent variable is the length of first-half injury time (in seconds) in matches in which one team was leading by one goal after regular time. “Home Ahead” takes the value 1 if the home team is one goal ahead and zero if the home team is one goal behind.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Home Ahead	-6.5* [3.839]	-5.296 [3.620]	-6.11* [3.651]	-6.718* [3.523]	-7.285* [3.828]	-7.108* [3.810]	-7.833** [3.786]
No. Treatments		0.933*** [0.163]	0.781*** [0.156]	1.08*** [0.343]	1.201*** [0.330]	1.275*** [0.335]	1.317*** [0.369]
No. Substitution		20.118*** [2.430]	18.396*** [2.406]	18.659*** [2.300]	18.131*** [2.197]	18.458*** [2.243]	17.89*** [2.214]
No. Fouls		0.227 [0.168]	0.058 [0.172]	-0.038 [0.260]	-0.066 [0.258]	-0.067 [0.283]	0.046 [0.260]
No. Yellow Cards			5.754*** [1.160]	5.78*** [1.137]	5.926*** [1.129]	5.685*** [1.233]	5.818*** [1.176]
No. 2nd Yellow			15.433*** [5.816]	15.196** [5.874]	13.989** [5.822]	12.69** [6.119]	11.154* [6.459]
No. Red Cards			27.724*** [6.943]	26.88*** [7.333]	31.182*** [7.129]	31.02*** [6.935]	33.767*** [7.315]
Season Dummies	No	No	No	Yes	Yes	Yes	Yes
controls	No	No	No	No	Yes	Yes	Yes
rel. strength	No	No	No	No	No	Yes	Yes
Home Team	No	No	No	No	No	Yes	Yes
Fixed Eff.	No	No	No	No	No	Yes	Yes
Visitor Team	No	No	No	No	No	Yes	Yes
Fixed Eff.	No	No	No	No	No	No	Yes
Referee	No	No	No	No	No	No	Yes
Fixed Eff.	No	No	No	No	No	No	Yes
Constant	69.078*** [3.280]	36.408*** [6.717]	34.52*** [7.020]	41.761*** [8.567]	52.914*** [11.366]	46.028*** [14.502]	63.567*** [16.483]
Observations	1495	1495	1495	1495	1403	1403	1403
R-squared	0	0.07	0.1	0.13	0.13	0.17	0.26

Notes:

1. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5, or 1 percent level respectively. Standard errors allow for correlation between observations of the same referee.
2. The control variables “No. Treatments”, “No. Substitution”, “No. Yellow Cards”, “No. 2nd Yellow”, “No. Red Cards”, sum the total number of respective events in the second half. The variable “No. Fouls” sum the total number of fouls during the game. Information on the number of fouls is not available for the two halves separately. When the number of any such event for the home and visiting team is controlled for separately, differences in coefficient estimates are very small and always statistically insignificant. Moreover, all other parameter estimates are robust to such specification changes.
3. Controls for relative strength include the relative number of tackles won, shots on goal, fouls, corner kicks and crosses as well as the absolute difference in rankings before the match and the ranking of the home team. Only the rank of the home team is marginally significant (p-value=.1) in one specification. The hypothesis that controls for relative strength are jointly significant is strongly rejected by a Wald test at any reasonable significance level (p-value exceeds .1 in all specifications). Measures on which controls of relative strengths are based are missing for the majority of matches in the season 1992/93, which explains differences in observations used across columns.

Table 3: LENGTH OF SECOND-HALF INJURY TIME IN DRAWN MATCHES

The dependent variable is the length of second-half injury time in drawn matches.

	[1]	[2]	[3]	[4]
Home Tied		7.872*		
		[4.421]		
0:0 Score			-11.05**	-8.987*
			[4.585]	[4.694]
No. Treatments	0.654	0.604	0.683	0.719
	[0.622]	[0.622]	[0.620]	[0.618]
No. Substitutions	6.019***	5.966***	5.926***	6.021***
	[1.635]	[1.632]	[1.630]	[1.625]
No. Yellow Cards	5.608***	5.558***	5.438***	5.772***
	[1.429]	[1.428]	[1.426]	[1.427]
No. 2nd Yellow	8.533	8.051	7.326	8.878
	[6.081]	[6.078]	[6.082]	[6.117]
No. Red Cards	10.451	9.801	10.121	10.335
	[6.629]	[6.630]	[6.609]	[6.591]
No. Fouls	0.438	0.462	0.42	0.457
	[0.397]	[0.397]	[0.396]	[0.412]
Rank Difference	0.682	0.659	0.638	0.587
	[0.520]	[0.519]	[0.519]	[0.518]
Rank of Home Team	0.546	0.548	0.593	0.458
	[0.429]	[0.429]	[0.428]	[0.430]
Rel. Tackles Won	-60.156	-66.244	-64.345	-65.433
	[62.254]	[62.257]	[62.077]	[61.951]
Rel. Shots on Goal	34.207*	34.039	41.457**	35.884*
	[20.726]	[20.696]	[20.877]	[20.888]
Rel. Fouls	-39.29	-35.232	-40.1	-41.428
	[26.793]	[26.851]	[26.709]	[26.663]
Rel. Corners	15.589	12.957	13.104	14.932
	[12.988]	[13.053]	[12.987]	[12.956]
Rel. Crosses	19.179	16.396	16.848	14.665
	[17.103]	[17.149]	[17.075]	[17.035]
No. Tackles				0.105
				[0.078]
No. Shots on goal				0.63*
				[0.349]
No. Crosses				0.503
				[0.335]
Constant	-20.54	-20.994	-14.514	-61.487**
	[25.140]	[25.105]	[25.184]	[30.875]
Observations	886	886	886	886
R-squared	0.37	0.37	0.37	0.38

Notes:

1. All regressions include season dummies, team dummies and referee dummies.
2. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5 and 1 percent level respectively. Standard errors allow for correlation between observations of the same referee.
3. Relative (Rel.) measures for an event X are calculated as the ratio of the number of X for home team to the total number of X minus 0.5 and are therefore positive when the home team has had more of the events than the visiting team.

Table 4: THE STADIUM, THE CROWD, AND 2ND HALF INJURY TIME IN CLOSE GAMES

The dependent variable is the length of injury time awarded at the end of the match in matches where the home team is either one goal behind (Score Difference=-1) or one goal ahead (Score Difference=1). The sample is split into matches that took place in stadiums without a running track separating the field and the stands (columns 1 and 2) and matches in stadiums with a running track (columns 3 and 4).

	Stadiums without Track		Stadiums with Track	
	[1]	[2]	[3]	[4]
Score Difference	-30.603** [13.867]	-29.297** [13.938]	-4.805 [7.454]	-7.392 [7.569]
Attendance (1000s)	1.137* [0.654]	0.986 [0.680]	-0.417 [0.595]	-0.367 [0.612]
Att.*ScDiff.	0.012 [0.227]	0.126 [0.231]	0.011 [0.221]	-0.123 [0.223]
Att./Capacity	-37.857 [32.085]	-39.833 [33.191]	7.994 [33.688]	8.775 [34.761]
Att./Cap. *ScDiff.	22.943 [18.086]	18.447 [18.257]	-13.094 [14.214]	-2.129 [14.381]
No. Treatments	1.185** [0.601]	1.499** [0.618]	-0.663 [0.518]	-0.412 [0.527]
No. Substitutions	2.486 [2.355]	2.941 [2.390]	-0.112 [2.282]	1.25 [2.299]
No. Yellow Cards	7.086*** [1.876]	8.004*** [1.952]	6.528*** [1.668]	5.943*** [1.737]
No. 2nd Yellow	1.432 [8.867]	-3.135 [8.970]	3.112 [6.281]	1.419 [6.379]
No. Red Cards	21.01** [9.528]	15.202 [9.831]	11.527 [7.343]	9.21 [7.328]
Controls	Yes	Yes	Yes	Yes
rel. strength				
season	Yes	Yes	Yes	Yes
dummies				
Visitor Team	Yes	Yes	Yes	Yes
Fixed Eff.				
Home Team	Yes	Yes	Yes	Yes
Fixed Eff.				
Referee	No	Yes	No	Yes
Fixed Eff.				
Constant	10.306 [38.023]	45.464 [52.650]	44.815 [44.703]	76.008 [46.951]
Observations	554	554	563	563
R-squared	0.24	0.38	0.28	0.41

Notes:

1. Standard errors are in brackets. One (two) asterisk(s) denotes significance at the 5 percent level (1 percent level).
2. The effects of controls for relative strength are never significant.
3. The effects of season dummies are as in Table 1 for all specification.
4. The results concerning the home bias are not affected by the inclusion or exclusion of team fixed effects.

Table 5: ATTENDANCE OF GAMES

The dependent variable is the deviation from average attendance			
Distance:	less than 150 km	4.638	[0.423]***
	150 km - 300 km	0.865	[0.380]**
	more than 450 km	-0.610	[0.383]
Attractiveness of Visitor		48.874	[1.569]***
Constant		-36.190	[1.152]***
Observations		3519	
R-squared		0.24	

Notes:

1. The variable “Distance” measures the shortest distance between the home cities of opponents. The attractiveness of the visitor is calculated as the average attendance-to-capacity ratio in all away games except those that are played on the current opponent’s turf.
2. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5, and 1 percent level respectively).
3. The effects of controls for relative strength are never significant.
4. To avoid correlation between explanatory variables and the disturbance term of the regression, the attractiveness of the visitor is calculated as the average attendance-to-capacity ratio when the visitor plays away excluding the matches against the home team. Similarly, the average attendance has been calculated over all home matches except for those against the current opponent.

Table 6: CROWD COMPOSITION EFFECTS IN CLOSE GAMES

The dependent variable second-half injury time in close matches		
	[1]	[2]
	Distance $\leq$ 150 km	Distance $>$ 150 km
Score Difference	-4.199	-12.055***
	[8.237]	[2.590]
Attraction*Score Difference	-20.335*	3.596
	[11.704]	[4.170]
Observations	199	918
R-squared	0.58	0.34

Notes:

1. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5, and 1 percent level respectively.
2. The regressions include controls for the number of treatments, substitutions, fouls and cards, controls for relative strength, season dummies, referee dummies and team dummies.
3. The variable “Attraction*Score Difference” is the interaction between the “Score Difference” variable and a dummy that takes the value 1 if the visiting team is in the top third of the ranking according to the approximated relative composition of the crowd. The latter variable is the ratio of the average attractiveness of the visitor (see Table 5) to the average attendance-to-capacity ratio in home games of the home team.

Table 7: CORRECTNESS OF DECISIONS CONCERNING AWARDED GOALS

Score Difference	Goal for Home Team			Goal for Visiting Team		
	Decision			Decision		
	wrong	correct	disputable	wrong	correct	disputable
≤ -2	1.68	92.26	6.06	0.68	97.62	1.70
-1	1.43	93.69	4.88	0.15	97.51	2.34
0	1.56	95.04	3.39	0.88	96.08	3.04
1	1.38	95.48	3.13	1.62	93.73	4.65
≥ 2	1.01	96.64	2.35	0.95	96.76	2.29
Total	1.43	95.05	3.52	0.92	95.99	3.09

Notes:

1. The Table displays the percentages of goals for the home team and the visiting team that were wrongly, correctly or disputably awarded by score difference. The score difference gives the number of goals that had been scored by the home team minus the number of goals that had been scored by the visitor prior to the decision in question.
2. The Table is based on all 10,163 goals (6,025 for the home team and 4,138 for the visitor) that were awarded from the start of the season 1992/1993 until the end of the first half of the season 2003/2004 and for which the correctness indicator is available.

Table 8: DECISIONS TO AWARD GOALS

	Home Team		Visiting Team		M-W Test	
	Awarded	Not Aw.	Awarded	Not Aw.	z-value	p-value
legitimate	5199	68	3637	46	-0.175	0.569
disputable	194	57	120	43	0.852	0.197
illegitimate	79	125	37	84	1.480	0.070

Notes:

1. The Table displays the numbers of legitimate, disputable and illegitimate goals that were either granted or denied for home and visiting teams.
2. The Table is based on all goals that were awarded and all goals that were not accepted from the start of the season 1993/1994 until the end of the first half of the season 2003/2004. The indicator about the correctness of the referee's decision is missing for two goals that have been awarded in this period and for one goal that has not been awarded. These observations are not included in the calculations for this table.
3. The last column shows the p-value for test of the hypothesis that referees are unbiased against the one-sided alternative that they are more likely to award the home team a goal.

Table 9: DECISIONS TO AWARD GOALS

Dependent Variable: 1 if Goal Awarded

		Objective Assessment of Situation			
		disputable	disputable	illegitimate	illegitimate
<u>Visiting Team</u>					
behind:	2 goals		1.068 [0.708]		-0.431 [0.761]
	1 goal		0.994** [0.450]		-0.319 [0.658]
tied:		-0.238 [0.237]	0.572 [0.428]	-0.194 [0.257]	-0.507 [0.654]
ahead:		-0.411 [0.295]		-0.94** [0.424]	
	1 goal		0.712 [0.497]		
	2 goals		-0.431 [0.847]		-1.068 [0.858]
	more than 2 goals		-0.253 [0.692]		0 [1.085]
<u>Home Team</u>					
behind:		0.442 [0.282]		-0.062 [0.255]	
	more than 2 goals				-0.674 [0.926]
	2 goals		0.929* [0.533]		-0.253 [0.744]
	1 goal		1.398*** [0.496]		-0.377 [0.664]
tied:		-0.107 [0.216]	0.703* [0.417]	0.124 [0.221]	-0.188 [0.641]
ahead:		-0.264 [0.226]		-0.04 [0.234]	
	1 goal		0.674 [0.439]		-0.07 [0.661]
	2 goals		0.512 [0.482]		-0.674 [0.678]
	more than 2 goals		-0.18 [0.620]		-0.842 [0.895]
Constant		0.81*** [0.173]	0 [0.396]	-0.312* [0.175]	0 [0.627]
Observations		414	411	325	315

Notes:

1. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5, and 1 percent level respectively.

Table 10: CORRECTNESS OF PENALTY KICK DECISIONS FOR AWARDED PENALTY KICKS

Score Difference	Penalty Kick for Home Team			Penalty Kick for Visiting Team		
	Decision			Decision		
	wrong	correct	disputable	wrong	correct	disputable
≤ -2	8.57	45.71	45.71	11.76	70.59	17.65
-1	14.29	62.86	22.86	0.00	87.10	12.90
0	6.38	62.55	31.06	8.04	68.75	23.21
1	5.00	70.00	25.00	6.25	67.50	26.25
≥ 2	4.05	78.38	17.57	2.08	81.25	16.67
Total	7.38	65.20	27.42	5.90	72.57	21.53

Notes:

1. The Table displays the percentages of awarded penalties for the home team and the visiting team that were wrongly, correctly or disputably awarded. The score difference gives the number of goals that had been scored by the home team minus the number of goals that had been scored by the visitor prior to the penalty kick decision.
2. The Table is based on all penalties that were awarded from the start of the season 1993/1994 until the end of the first half of the season 2003/2004 and covers 857 penalty kick decisions, 569 for the home team and 288 for the visitor.

Table 11: DECISIONS TO AWARD PENALTY KICKS

	Home Team		Visiting Team		MW-Test	
	Awarded	Not Aw.	Awarded	Not Aw.	z-value	p-value
legitimate	207	87	115	64	1.393	0.082
disputable	86	214	30	118	1.906	0.028
illegitimate	23	192	13	87	-0.597	0.725

Notes:

1. The Table shows the numbers of legitimate, disputable and illegitimate penalty kicks that were either granted or denied for home and visiting teams.
2. The Table is based on all penalty kicks that were awarded and actions that could have potentially lead to a penalty kick but were not penalized from the start of the season 1998/1999 until the end of the first half of the season 2003/2004.
3. The last column shows the p-value for the test that referees are unbiased against the one-sided alternative that they are more likely to award a penalty for the home team.

Table 12: PROBABILITY OF AWARDING A PENALTY KICK CONDITIONAL ON JUSTIFICATION

	Dependent Variable: 1 if Penalty Kick Awarded		
	legitimate	disputable	illegitimate
Dummy, 1 if home	0.209* [0.125]	0.265* [0.140]	-0.123 [0.197]
Dummy, 1 if track	0.394*** [0.124]	-0.166 [0.134]	-0.205 [0.201]
Constant	0.175 [0.113]	-0.769*** [0.127]	-1.055*** [0.174]
Observations	473	448	315

Notes:

1. Standard errors are in brackets. One, two, and three asterisk(s) denote(s) significance at the 10, 5, or 1 percent level respectively.