Are Subjective Evaluations Biased by Social Factors or Connections? An Econometric Analysis of Soccer Referee Decisions

Scoppa, Vincenzo

Dipartimento di Economia e Statistica - Università della Calabria

2008

Are Subjective Evaluations Biased by Social Factors or Connections? An Econometric Analysis of Soccer Referee Decisions

Vincenzo Scoppa*

Working Paper Version

Forthcoming in Empirical Economics, 2008, vol. 35, Number 1, August, pp. 123-140.

Many incentive contracts are based on subjective evaluations and contractual disputes depend on judges' decisions. However, subjective evaluations raise risks of favouritism and distortions. Sport contests are a fruitful field for testing empirically theories of incentives. In this paper the behaviour of the referees in the Italian soccer (football) league ("Serie A") is analyzed. Using data on injury (or extra) time subjectively assigned by the referee at the end of the match and controlling for factors which may influence it (players substitutions, yellow and red cards, penalty kicks, etc.), we show that referees are biased in favour of home team, in that injury time is significantly greater if home teams are losing. The refereeing bias increases greatly when there is no running track in the stadium and the crowd is close to the pitch. Following the 2006 "Serie A" scandal we test whether favouritism emerges towards teams suspected of connections with referees finding that these teams obtain favourable decisions. Social pressure by the crowd attending the match however appears to be the main cause of favouritism.

***JEL Classification**:* M50; L83; Z13.
Keywords: Favouritism, Subjective evaluation; Sport economics.

1. Introduction

Many incentive mechanisms are based on subjective evaluations since it is often difficult to find verifiable objective measures of performance (Prendergast, 1999). Often these evaluations are made by supervisors, placed between the principal and the agent, who are delegated by the principal to observe and evaluate agent performances and are not residual claimants of output. Incentive theories have emphasized that subjective evaluations are plagued by a host of problems, from influence costs and rent-seeking to collusion and corruption, from compression of ratings to favouritism (Milgrom, 1988; Tirole, 1992; Prendergast and Topel, 1996). Such problems may be caused by exchange of money or favours between the agent and the supervisor (side-payments) to obtain a better evaluation or may be related to psychological and social considerations (Akerlof, 1980; Becker and Murphy, 2000). In fact, the social environment

* Dipartimento di Economia e Statistica, Università della Calabria, 87036 Arcavacata di Rende (Cosenza), Italy. E-mail address: v.scoppa@unical.it. I would like to thank Francesco Aiello, Martin Brimble, Maria De Paola, Filippo Domma, Nicola Meccheri, Antonio Nicita, Michela Ponzo and two anonymous referees for useful comments. The usual caveats apply.

created inside an organization by daily personal relationships between agents and supervisors influences judgments and, at the same time, evaluations might be based on a supervisor's personal preferences.

Contractual enforcement realized by a Court may be subject to the same problems, because it often depends on a judge's subjective evaluation. As pointed out by Tirole (1999, p. 761) "contracts are enforced by human beings. Judges are subject to both moral hazard and adverse selection; they may also have their own preferences […] Judges and arbitrators may also collude with the parties".[1]

In the light of the great relevance of judges' and supervisors' subjective evaluations, it would be useful to understand empirically to what extent economic or social factors can influence evaluators' behaviour in real-world contracts. However, in normal business contracts it is hard to assess judges' decisions or to obtain data on supervisor misconduct and empirical evidence on these aspects is scant.[2] Data from sport contests can be very useful in this respect.

The aim of this paper is to provide empirical evidence on the working of subjective evaluation systems using data on the behaviour of referees of professional soccer (football) matches. Referees act as agents of the Soccer League (the national or international Association of Teams) to impartially control and enforce the rules of the game. Their behaviour may be observed and some aspects of their decisions can be measured.

Economists are increasingly interested in the analysis of sport contests. The great availability of data on individual and team performances makes sport contests a fertile terrain for testing theories of incentives, labour markets and game theory (for a review of the economics of sports, see Szymanski, 2003a, 2003b). A number of papers have pursued this strategy. Ehrenberg and Bognanno (1990) test predictions of tournament theory using data from European and American golf tournaments, finding that higher prizes elicit better player performances. Duggan and Levitt (2002) study corruption by analyzing the performances of Japanese professional sumo wrestlers and find evidence that decisive matches are rigged and wrestlers exchange favours with each other. Chiappori, Levitt and Groceclose (2002) and Walker and Wooders (2001) analyze players' behaviours in using mixed strategies, respectively, in kicking penalties in soccer and in serving in tennis matches.

We empirically test whether soccer referees take decisions impartially (as they are delegated to do) or if they are subject to favouritism, because of social pressure, corruption or other forms of influence. We have detailed data for each match from two recent seasons of the Italian professional soccer league "Serie A": 2003-2004 and 2004-2005. We analyse in particular refereeing decisions regarding injury time (or extra time) to be added at the end of a match. In principle, this time should be given to compensate for time lost for player injuries,

[1] For a similar view see Posner (1993).

[2] A notable exception is the recent analysis by Ichino, Polo and Rettore (2003) who verify how social considerations influence judges' decisions on firing litigations in Italian regional labour markets.

player substitutions, yellow and red cards, protests, and other wastes of time.

The existence of a problem with the injury time decision is confirmed by the ruling by FIFA ("Fédération Internationale de Football Association") which, in 1998, in order to limit referees' discretion, established that referees have to announce publicly at the end of normal time the minutes of extra time to be played.

The suspicion exists that referees are not impartial between home and visiting teams or in games between "big" and "small" teams.

Favouritism towards a home team is due probably to social pressure from the crowd in the stadium noisily supporting the home team.[3] Referees are human beings who – as shown by a series of psychological studies (see, amongst others, Nevill, Balmer and Williams, 1999 and 2002) – can be subconsciously influenced by the noise of a large crowd in the stadium who react to adverse decisions by yelling and ferociously protesting.

In addition to home bias, referees have been often accused of favouring "big" teams, who have economic, political and media power off the field. As a partial confirmation of these suspicions, in May 2006 an enormous scandal exploded involving some of the top teams in "Serie A" (Juventus, Milan, Fiorentina, Lazio, Reggina) which were accused of trying to appoint favourable referees for their matches and put pressure on them.

Our simple identification strategy is based on the assumption that, once controlled for a set of observable factors, the amount of extra time should be independent of the identity of the team who is winning the match at the end of normal time.

Therefore, we first insert a dummy variable (*Home Behind*) equal to one if the home team is behind by one goal and equal to zero if the home team is ahead by one goal or if the teams are drawing. We focus on close matches because – given the low probability of scoring a goal in the few minutes of injury time – only the final result of these matches could possibly be changed.

We find that the coefficient of the dummy variable *Home Behind* – in many different specifications – is positive and significantly different from zero, implying that the time added by the referee is greater if the home team is behind, giving the latter a greater chance of equalizing. Results show the great importance of physical proximity of a crowd to the pitch in influencing the referee, since the bias is much higher when there is no running track in the stadium.

A similar analysis is carried out for the teams suspected of connections with referees, including in the regressions a dummy *Suspected Behind* indicating when a suspected team is losing. We find that these teams tend to obtain favourable decisions, in that, when one of these teams is behind, extra time is significantly longer.

In both situations, we conclude that referees do not behave impartially, favouring both home teams and "big" teams.

Our analysis replicates for Italy – one of the major European soccer championships – the

[3] In soccer and other team sports, supporters usually attach themselves to local teams.

analysis of Garicano, Palacios-Huerta and Prendergast (2005) who study soccer referee performance in the Spanish "Primera Liga" and Sutter and Kocher (2004) and Dohmen (2005) who conduct a similar analysis for the German "Bundesliga". All these works find evidence of favouritism towards home teams.

Our paper confirms their findings with regards to Italy, and extends the analysis to provide evidence that "big" teams and, in particular, the suspected teams in the 2006 scandal – who allegedly had established connections with referees – are favoured.

We consider this kind of biased decisions to be only the "tip of the iceberg" of favouritism. There are, of course, other discretionary decisions taken by the referee during a match which could influence the result in a much more definite way: decisions on penalties and free-kicks, red cards, off-sides and so on. However, the merit of these decisions cannot be judged objectively and it is necessary to have a subjective evaluation of the phase of the game and of player behaviour. Moreover, when an evaluation is made on TV it is often different from that of the referee, which is immediate and without the support of technological devices. Even when decisions appear numerically to be in favour of a particular team, this might be the result, for example, of the fact that that team takes the attacks more often than the opposition, leading to a higher probability of winning a penalty or a free-kick, inducing the opponents to commit fouls.

Besides being useful for verifying economic theory, sport is interesting in its own respect. Soccer is an important and growing industry in Italy and other European countries. According to Baroncelli and Lago (2006), the total revenues generated by soccer in Italy are 4.2 billion euro; moreover, about 25 million Italians follow soccer on television and 8 million watch matches at stadiums. Hot debates on television and newspapers, as well as between opposing supporters, are triggered by crucial referee decisions.

One of the most interesting product characteristics that soccer teams sell is suspense, that is, the interest of spectators (and, as a consequence, attendance at matches and willingness to pay for TV coverage) is not only a function of the ability and effort of players but is directly related to the uncertainty of outcome (Neale, 1964). It follows that the aspect we examine is crucial for the economic health of soccer, because if referees are not impartial, suspense is undermined and demand might suffer.

The paper is organized in the following way. In Section 2 we describe the data and provide preliminary evidence. In Section 3 we conduct an econometric analysis to estimate whether referee favouritism exists towards a home team and provide some robustness checks. In Section 4 we aim to evaluate whether "big" teams are favoured. Section 5 concludes.

2. Data and First Evidence on Referee Bias

Data for the 2003-2004 and 2004-2005 seasons of the Italian soccer league "Serie A" were

collected from the website of *RAI* Sport (http://www.raisport.rai.it/).[4]

The Italian "Serie A" had 18 teams in 2003-2004 and 20 teams in 2004-2005. In each season, teams played each other twice (both as the home and visiting team). Therefore, there were 306 and 380 matches respectively, totalling 686 observations. For each match (which is divided into two periods of 45 minutes), there are data available on teams, goals scored, injury time added at the end of each half, penalties, player substitutions, disciplinary sanctions (yellow and red cards), serious injuries, and attendance[5]. We have the time at which each episode occurred in the game (except yellow cards) and thus we can impute each incident to the first and second half or to injury time. Score difference is equal to *Goals Home* minus *Goals Visitor* (both measured after 90 minutes). Descriptive statistics are reported in Table 1.

Table 1. Descriptive statistics

Variables	**Mean**	**SD**	**Min.**	**Max.**	**Obs.**
Goals Home♣	1.423	1.229	0	6	686
Goals Visitor♣	1.058	1.019	0	6	686
Score Difference♣	0.364	1.481	-4	6	686
Injury Time Second Half (minutes)	3.625	1.113	0	8	686
Injury Time First Half (minutes)	1.590	1.034	0	8	686
Player Substitutions (Second Half)	4.564	0.991	1	6	686
Player Substitutions (First Half)	0.275	0.539	0	3	686
Red Cards (Second Half)	0.287	0.595	0	6	686
Red Cards (First Half)	0.058	0.247	0	2	686
Yellow Cards	3.933	1.865	0	9	686
Penalties (Second Half)	0.169	0.408	0	2	686
Penalties (First Half)	0.114	0.344	0	2	686
Goals Home (First Half)	0.627	0.755	0	3	686
Goals Visitor (First Half)	0.456	0.642	0	3	686
Attendance (1000s)	25.860	17.137	3.8	81	681
Track	0.475	0.499	0	0	686

♣: Measured at the end of normal time in the second half. Goals scored after 46 minutes are counted as being in injury time. 74 goals were scored in injury time.

Injury time ranges between 0 and 8 minutes. Average injury time is 3.62 minutes. A total of 74 goals were scored in injury time (34 in the 2003-2004 season and 40 in the 2004-2005).

Figure 1 is a first preliminary evidence of referee favouritism towards the home team. It shows injury time as a function of the score difference between the home and visiting teams after 90 minutes. In the situations in which the score is close – and hence extra time might be used by the teams to influence the final score of the match – there is clear evidence that referees tend to award significantly more time when the home team is behind by one goal (4.19 minutes, about half a minute more than average). On the other hand, there is a tendency for referees to reduce injury time when the score is not close, independently of which team is leading the game (in our sample 9 percent of matches involved a difference of 3 or more goals).

[4] RAI is the national public television broadcasting system.
[5] There are a few missing observations as regards attendance.

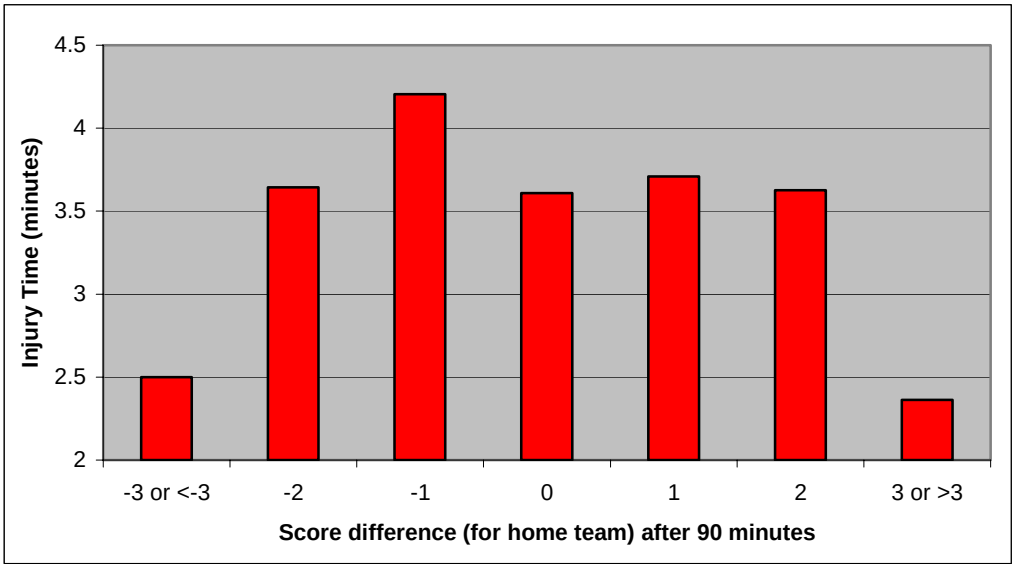

Figure 1. Injury time (in the second half) as a function of score difference in favour of the home team (at the end of normal time).

Another kind of refereeing favouritism towards a home team that one might expect to see is the shortening of the match when the home team is ahead. As can be seen in Figure 1, this form of favouritism does not emerge in our sample. This is similar to the findings of Sutter and Kocher (2004) and Dohmen (2005) for Germany but it is in contrast to Garicano *et al.* (2005) who find that Spanish referees also favour home teams when they are ahead. A possible explanation could be that referees get less pressure because the risk of the visiting team's scoring is significantly lower than the possibility of the home team's scoring (see Table 1). Moreover, referee behaviour might reflect the preferences of the home team's supporters who might be more averse to the risk of a defeat at home than to the risk of not winning the match.

3. An Econometric Analysis of Referee Favouritism toward Home Teams

Since extra time is influenced by a series of observable (and unobservable) factors, in this and in the next Section, we proceed with an econometric analysis in order to control for these factors, focusing on close score margins (485 matches) and ignoring matches with two or more goals difference in score. We do not consider these matches because, given the low probability of scoring per minute, only in close matches can a goal in extra time have an impact on the final score.

We define a dummy variable "*Home Behind*" equal to one if, at the end of normal time in the second half, the home team is behind by one goal (N=111) and equal to zero if the home team is ahead of one goal (N=153) or if the game is a draw (N=221).

Our main point is that if referees act impartially this dummy should have no role in

explaining extra time added at the end of a match and should be statistically insignificant.

In column (1) of Table 2, a simple OLS regression is run with *Injury Time* as a dependent variable on a constant and our variable of interest *Home Behind*, with no other right-hand-side variables. The results show that if a home team is behind by 1 goal, referees tend to add 0.564 minutes to the game (extra time is therefore 15% percent longer than the average of 3.652 minutes), allowing the home team more time to equalize. The coefficient is statistically strongly different from zero (t-ratio is 4.91; p-value=0.00).[6] In this and all other specifications, standard errors (reported in parentheses) are corrected for heteroskedasticity.

[6] Considering all the observations (686) instead of just the close games, and defining *Home Behind* as equal to zero except when the home team is behind by 1 goal, the coefficient would be 0.69 with a t-ratio=6.26.

Table 2. OLS regression estimates for injury time. Dependent variable *Injury Time*.

Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.652***	2.085***	2.238***	2.306***	2.263***	2.069***
	(0.050)	(0.238)	(0.245)	(0.313)	(0.295)	(0.239)
Home Behind	0.564***	0.485***	0.509***	0.522***	0.489***	0.745***
	(0.115)	(0.111)	(0.108)	(0.105)	(0.129)	(0.166)
Player Substitutions [a]		0.280***	0.261***	0.262***	0.283***	0.274***
		(0.043)	(0.041)	(0.043)	(0.050)	(0.043)
Red Cards [a]		0.135**	0.138**	0.163**	0.113*	0.136**
		(0.064)	(0.063)	(0.072)	(0.064)	(0.062)
Yellow Cards		0.062**	0.071***	0.067***	0.040	0.062**
		(0.027)	(0.027)	(0.024)	(0.032)	(0.027)
Penalties [a]		0.094	0.067	0.036	0.147	0.087
		(0.112)	(0.116)	(0.125)	(0.143)	(0.111)
Difference in ranks			-0.001	-0.005	0.004	
			(0.010)	(0.010)	(0.013)	
Attendance (1000s)			0.001	-0.000	-0.003	
			(0.003)	(0.007)	(0.003)	
Other Stoppages			1.503***	1.571***	1.871**	
			(0.547)	(0.478)	(0.806)	
Goals scored			0.044	0.037	0.086	
			(0.043)	(0.039)	(0.050)	
Season 2005 dummy			-0.310***	-0.304***	-0.269***	
			(0.086)	(0.106)	(0.108)	
Team Fixed effects				YES: home[b]		
Interaction Attendance					0.008	
					(0.008)	
Budget difference					-0.0005	
					(0.0008)	
Interaction Budget					0.0005	
					(0.001)	
Track						0.099
						(0.095)
Track*Home Behind						-0.535**
						(0.217)
R^2	0.054	0.155	0.209	0.235	0.221	0.167
Observations	485	485	481	481	332	485

Notes: The dependent variable is injury time added at the end of the second half (in minutes). *Home Behind* is coded one if the home team is behind by one goal, zero if it is ahead by one goal or in the case of a draw. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

[a]: Counted only for the second half.

[b]: F test that all Team Fixed Effects $\mu_i = 0$: F(23, 447) =0.67; Prob > F = 0.8774

In column (2) of Table 2 a more complete model of determination of injury time is estimated. According to the rules of soccer[7], referees can add – discretionarily – extra time for time lost for player substitutions, player injuries, and other wastes of time. Therefore, we insert as explanatory variables *Player Substitutions*, *Penalties*, *Red Cards* and *Yellow Cards* (these variables are counted only if the episodes take place in the second half, with the exception of *Yellow Cards*, for which timing is not available in the data). Disciplinary sanctions are often related to foul play which causes injuries to opponent players. Moreover, penalties and disciplinary sanctions usually give rise to protests among players which lead to waste of time. The estimates reported in column (2) confirm that player substitutions and red and yellow cards appear to increase extra time significantly (respectively by 0.280, 0.135 and 0.062 minutes), while *Penalties* are not significant. The percentage of variability explained increases substantially $\left(R^2 = 0.155\right)$. As regards to the dummy aimed at capturing favouritism towards the home team, the value of *Home Behind* is only slightly reduced (0.485) but its strong significance does not change (t-ratio=4.37; p-value=0.00) when controlling for these factors: the evidence thus shows that the greater time awarded by referees cannot be imputed to player substitutions or disciplinary sanctions.

In model (3) we consider additional control variables which could influence injury time. A match of great intensity (a hard-fought match) is likely to cause more player injuries and protests and, therefore, it should lead to more extra time being added by the referee. As a proxy for the intensity of the match we include the absolute difference between teams' rankings (given by their position at the end of the season) and the attendance at the game. Moreover, in some cases, matches are interrupted for serious injuries to players or for disorders among the supporters: we consider a dummy variable (*Other Stoppages*) equal to one when these episodes occur and zero otherwise.[8] Finally, another two control variables which we take into account are the number of goals scored and a year dummy.

The variables included to gauge the intensity of the game do not seem very significant and neither is the number of goals scored. *Other stoppages* is significant and the season 2005 dummy is also significant and negative. Our dummy of interest *Home Behind* remains substantially of the same size (0.509) and highly significant (t-ratio is 4.37). R-squared is increased to 0.209 when these controls are added.

We next estimate a Least Square Dummy Variables model to verify whether home team fixed effects are present in explaining extra time. Results are reported in column (4). Team fixed effects are not significant implying that there are no substantial differences between teams in

[7] See the link "Laws of the game" on the FIFA website: www.FIFA.com.
[8] Excluding these (few) observations from the data as outliers does not change our results.

extra time (more on this aspect below).[9] The dummy *Home Behind* is always largely significant.

In column (5) we test the existence of some interaction effects. First, we interact the dummy *Home Behind* with home team dummies to check whether there are differences in the extent of favouritism between teams: none of the interaction variables were individually statistically significant (results are not reported), implying that referees favour home teams in adding extra time more or less to the same extent. However, this result must be interpreted with caution, because the variability of many interaction dummies in the sample is very low.

Further, we control for the teams budgets (revenues) as a proxy of the power of the clubs.[10] We insert, as explanatory variables, both the difference in revenues between the home and visiting team and this variable interacted with the *Home Behind* dummy, to check if bigger teams obtain more favours.[11] These two variables are not significant[12].

We also test, in column (5), whether the number of spectators influences the extent of favouritism. In principle, if referees are subject to social pressure, when the roaring crowd is larger their decisions should be more distorted in favour of the home team. We define an interaction dummy "*Interaction Attendance*" equal to *Home Behind**(*Attendance-24.876*), where 24.876 is the average attendance in the considered sample.

According to our estimates, the coefficient is positive, that is, higher than average attendance increases referee favouritism towards the home team, but it is not statistically different from zero. Therefore, even if we find strong evidence that referees favour home teams, we do not find that referee behaviour is influenced by the sheer number of spectators. A reasonable explanation is that we cannot assume that a larger number of spectators necessarily support the home team. In fact, matches are usually attended by larger crowds when they are played by teams from nearby towns or when very popular teams are visitors. In these cases, both the home and visiting supporters are likely to be more numerous: local supporters are more interested in watching matches against major teams or historically rival teams and, at the same time, visiting supporters are more numerous if the geographic distance is not high or because popular teams tend to have supporters who come from all geographic areas. Therefore, since there is no clear indication with regards the composition of a crowd which attends a big match, it is difficult to infer how the number of spectators influences the referee.

Finally, following Dohmen (2005) we consider whether the presence of a running track in the stadium influences the extent of favouritism. In stadiums in which there is no track the crowd is much closer to the field and is able to put more pressure on the referee. About half of

[9] Moreover, we have tried to consider visiting team fixed effects (not reported) without appreciable modifications.
[10] Data on budgets are from Baroncelli and Lago (2006).
[11] Moreover, we also considered *Home Budget* and *Visitor Budget* separately and the interactions of these variables with *Home Behind* but the results (not reported) do not change significantly.
[12] This is probably due to the fact that not all big teams tend to obtain favours (see Section 4 for more details).

the stadiums of teams in "Serie A" have been built with a running track for athletic competitions. The dummy *Track* is coded equal to one if a track is present in the stadium and zero otherwise. Much more important is the interaction variable between the dummy *Home Behind* and the dummy *Track*. Estimates are presented in column (6). This interaction variable has a dramatic effect. It is negative (-0.535) and statistically significant (p-value=0.014). When the match is played in a stadium in which there is no track, the increase in injury time if the home team is behind is much higher (0.745) (about 45 seconds). Instead, when there is a track favouritism towards the home team is lower and equal to 0.21 (which is only marginally significantly different from zero, p-value=0.101). This constitutes strong evidence that psychological pressure in a particular environment leads to biased refereeing decisions.

3.1. Alternative Measures of Favouritism

In Table 3, in order to check the robustness of our results, we follow Garicano *et al.* (2005), in defining an alternative dummy variable to gauge a team's need for injury time (*Home Behind GPP*) coded as one if the home team is behind by one goal, zero if it is ahead by one goal, that is, excluding drawn games. In columns (1), (2) and (3) we replicate estimations of, respectively, regressions (1), (2) and (6) of Table 2 by substituting *Home Behind* with this new variable. The coefficient of "*Home Behind GPP*" results as nearly of the same size as our previous results (around 0.5 minutes) and it is again significantly different from zero. When no running track is present in the stadium referees add 0.67 minutes on average.

Table 3. OLS regression estimates. Dependent variable *Injury Time*.
Home Behind as defined by Garicano *et al.* (2005).

Explanatory Variables	(1)	(2)	(3)
Constant	3.712*** (0.070)	2.659*** (0.299)	2.709*** (0.301)
Home Behind GPP	0.504*** (0.126)	0.438*** (0.125)	0.675*** (0.181)
Player Substitutions [a]		0.201*** (0.059)	0.189*** (0.059)
Red Cards [a]		0.220** (0.099)	0.229** (0.097)
Yellow Cards		0.020 (0.031)	0.019 (0.031)
Penalties [a]		0.079 (0.125)	0.074 (0.127)
Track			0.011 (0.141)
Track*Home Behind GPP			-0.474** (0.241)

R^2	0.062	0.131	0.153
Observations	264	264	264

Notes: The dependent variable is injury time added at the end of second half (in minutes). *Home Behind GPP* is defined one if the home team is behind by 1 goal, zero if it is ahead by 1 goal (that is, not considering drawn games). Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.
[a]: Counted only for the second half.

A further check of favouritism towards home team consists in verifying whether home teams are awarded penalties more often than visiting teams. Given that there is a high probability penalties will be transformed into goals with high probability, penalties have a much greater impact on the final outcome than the lengthening of injury time. 123 penalties were awarded to home teams and 75 penalties to visiting teams in the two seasons under consideration.

Since we cannot consider that the probability of being awarded a penalty is the same for the home and visiting team, because the home team tends to attack more often, we take the ratio between penalties and goals scored. This proportion is equal to 0.135 for home teams and to 0.101 for visiting teams. A t-test for the difference between two averages gives a t=1.71 and p-value 0.046, leading us to reject the hypothesis that there is equal possibility that penalties will be awarded to the home and visiting teams in favour of the hypothesis that home teams obtain more penalties.[13]

3.2. What Causes Favouritism towards Home Teams?

The existence of favouritism is therefore strong and clear in the data. What is the reason for referees favouring home teams? The most likely factor appears to be the influence of the crowd in the stadium supporting the home team loudly: referees subconsciously succumb to this pressure, avoiding decisions which would disappoint the spectators in the stadium. This is strongly confirmed by the different impact that the crowd has in stadiums with and without a track.

It has been widely shown by several psychological studies that social pressure by the crowd is one of the factors that induces a home advantage in team sports. In an experiment, Nevill, Balmer and Williams (2002) have compared professional referees decisions (taken watching a videotaped recording of a match) when they hear the reactions of the crowd with their behaviour when they watch the match in silence. They show that referees hearing the noise of the crowd were significantly more acquiescent to the home team and in line with the effective

[13] Unfortunately we do not have available data on refused penalties in order to examine the relationship between awarded and refused penalties for home and visiting teams (as made in Sutter and Kocher, 2004).

decisions taken by the referee on the field. As explained by Wallsten and Barton (1982) one of the reasons which leads referees to take biased decisions is that they have to make instantaneous decisions, and they tend to focus on the most salient cues, one of which may be the crowd noise. Moreover, they tend to avoid potential displeasure for the crowd. Carmichael and Thomas (2005, p. 266) report that indoor sports – in which there is great proximity between the crowd, players and referees – are usually characterized by greater home advantage than outdoor sports.

Akerlof (1980), Bernheim (1994) and Becker and Murphy (2000) have shown how individual preferences can be influenced by social interactions, by the human need to be approved or to avoid displeasure for others, even if this can be costly to individuals. Prendergast and Topel (1996) present a model in which supervisors in organizations trade-off the satisfying of personal preferences towards subordinates with a wage penalty deriving from biased evaluations. Garicano *et al.* (2001) explicitly model the trade-off referees face between material incentives deriving from a high probability of being re-appointed, if impartial, and their desires to satisfy the crowd.

3.3. Falsification checks

In this Section we undertake some further robustness checks in order to verify previous results. If the idea of the paper is correct, referees would be less influenced by the pressure of the crowd in the stadium (or by some other factor) if their decisions were not so pivotal, and therefore in these contexts we should observe less favouritism.

Two circumstances in which referees take decisions which are not crucial are when they add extra time at the end of the first half – in fact even if the home team is behind there are more than 45 minutes left to equalize in the second half – or when, at the end of regular time in the second half, the game is not close (that is, there is a difference of two or more goals).

Firstly, in columns (1) and (2) of Table 4 we test whether there is favouritism in deciding injury time at the end of the first half, replicating the regressions (1) and (6) of Table 2. Even if the coefficient of *Home Behind* is positive, that is, referees tend to add extra time if the home team is behind by one goal, it is not significantly different from zero (p-value=0.30). In column (2) the usual explanatory variables (player substitutions, red and yellow cards, penalties) are added and they all result as being significant, but *Home Behind* is still not significant.

Table 4. OLS estimations for injury time in the first half and in games which are not close.

Dependent variable	(1) Injury Time at the end of	(2) Injury Time at the end of	(3) Injury Time at the end of	(4) Injury Time at the end of
Explanatory Variables	First Half	First Half	Second Half	Second Half
Constant	1.555***	1.169***	3.647***	1.964***
	(0.048)	(0.110)	(0.045)	(0.218)
Home Behind	0.109	0.030		
	(0.100)	(0.137)		
Home Behind Two Goals			0.027	-0.123
			(0.152)	(0.195)
Player Substitutions [a]		0.213***		0.267***
		(0.083)		(0.039)
Red Cards [a]		0.474**		0.154**
		(0.226)		(0.066)
Yellow Cards		0.057**		0.085***
		(0.024)		(0.026)
Penalties [a]		0.304**		0.094
		(0.124)		(0.113)
Track		0.094		0.155
		(0.093)		(0.086)
Track*Home Behind		0.087		0.317
		(0.197)		(0.299)
R^2	0.002	0.062	0.000	0.119
Observations	614	614	511	511

Notes: In columns (1) and (2), the dependent variable is injury time added at the end of first half (in minutes). Home Behind Two Goals is coded one if the home team is behind by 2 goals, zero if it is ahead by 1 or 2 goals or in case of a draw. Robust standard errors are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level. [a]: in regressions (1) and (2) these variables refer to first half events, whereas in regressions (3) and (4) they refer to the second half.

Secondly, we do not expect referees to show favouritism when the score is not close at the end of the second half. In order to test for this possibility, we consider matches with two goal differences and define a dummy variable (*Home Behind Two Goals*) equal to one if the home team is behind by two goals and zero if the home team is ahead by one or two goals or it is in a draw. The results are reported in columns (3) and (4) of Table 4. In both regressions, there is almost no effect on the allowance of injury time due to the fact that the home team is behind by two goals. The coefficients are not significant at any conventional level.

These further tests confirm that referees add more injury time only when it is crucial for home teams, implying that the aim is just to favour them.

4. Referee Favouritism towards "Big" Teams

"Big" teams have been often accused of being favoured by referees.[14] As a confirmation of these suspicions, in May 2006 five Italian football clubs from "Serie A" (Juventus (Turin), AC Milan, Lazio (Rome), Fiorentina (Florence) and Reggina (Reggio Calabria)) were involved in a major scandal regarding match fixing. Police investigations uncovered – through intercepted telephone conversations – a network of relations between team managers and officials responsible for the assigning of referees to matches. Allegedly, this network was aimed at guaranteeing the designation of compliant referees for certain matches[15] and to put pressure on referees to favour these teams.

Juventus and Milan are the "biggest" clubs in "Serie A". They have won most of the championships ("scudetti") and the most important European club competitions. They are related to big economic empires (Juventus to FIAT and the Agnelli family; Milan to Mediaset and Silvio Berlusconi). Fiorentina and Lazio are also important teams from Florence and Roma respectively.[16],[17]

Sporting tribunals found these teams guilty of misbehaviour and decided to strip Juventus of two titles ("scudetti") won in 2004-05 and 2005-06 and to relegate it to the second division (and with 9 points deduction); the other teams began the 2006-007 season with several points deduction (Milan: -8; Fiorentina: -15; Lazio -3; Reggina -11).

In this Section we test for the existence of favouritism towards these teams (and then towards other "big" teams) verifying whether referees are likely to be influenced by connections, by the economic power of the teams or by their vast popularity. Our aim is to test whether referees show favouritism towards major teams regardless whether they play at home or away.

We conduct the same analysis as in Section 3 but, instead of the dummy "*Home Behind*", we define a dummy variable "*Suspected Behind*" equal to one if a suspected team is behind by one goal and equal to zero otherwise.[18] We consider all the close games played by all the teams in the first three columns.

Column (1) in Table 5 shows the estimation results of a regression of *Injury Time* on *Suspected Behind* as the only right-hand-side variable. The coefficient is positive and statistically significant at a 5 percent level (p-value=0.03) implying that when a suspected team is behind, referees add on average 0.38 minutes to injury time (this is substantially lower than

[14] For an analysis of this aspect for the German Bundesliga see Sutter and Kocher (2002).
[15] According to the rule of "Serie A", referees should be randomly selected, once being allocated to certain groups to take into account referee abilities.
[16] Another two teams which – on the basis of past success, of the size of their support and budget – can be considered as "big" teams are Inter Milan and AS Roma. These two clubs were not involved in the scandal.
[17] On the other hand, Reggina is a "small" club from a Southern Italian town.
[18] We do not consider matches in which two suspected teams play one against the other.

the estimate of about 0.50 for home teams).

Column (2) includes in this regression some other explanatory variables (player substitutions, disciplinary sanctions and penalties). These variables, similarly to regressions in Table 2, results significant. *Suspected Behind* is again statistically significant (p-value=0.013). Therefore, our estimations show clear evidence of referee favouritism towards suspected teams.

Table 5. Referee Favouritism towards suspected and other major teams. OLS regressions. Dependent variable: Injury Time (second half)

Explanatory Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Constant	3.752*** (0.048)	2.045*** (0.237)	2.041*** (0.239)	3.75*** (0.086)	2.195*** (0.436)	2.668*** (0.416)
Suspected Behind	0.380** (0.174)	0.422** (0.169)	0.381* (0.194)	0.382** (0.189)	0.398** (0.185)	
Player Substitutions [a]		0.298*** (0.042)	0.285*** (0.043)		0.228*** (0.073)	0.176*** (0.066)
Red Cards [a]		0.150** (0.063)	0.137** (0.064)		0.161* (0.083)	0.115 (0.092)
Yellow Cards		0.069** (0.027)	0.061** (0.027)		0.113** (0.047)	0.082 (0.052)
Penalties [a]		0.099 (0.118)	0.108 (0.113)		0.062 (0.182)	0.146 (0.229)
Home Behind			0.472*** (0.117)			
(Home Behind)*(Suspected Behind)			-0.179 (0.351)			
Big Behind						0.210 (0.181)
R^2	0.010	0.128	0.161	0.022	0.131	0.064
Observations	485	485	485	182	182	192

Notes: *Suspected Behind* is coded one if a suspected team (Juventus, Milan, Fiorentina, Lazio and Reggina) is behind and zero otherwise. *Big Behind* is coded one if a big team (Juventus, Milan, Fiorentina, Lazio, Inter and Roma) is behind and zero otherwise. Columns (4), (5) and (6) only consider matches played by one of the teams involved. Robust standard errors are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

[a]: Counted only for the second half.

In column (3) of Table 5 we analyze jointly the effect of favouritism towards home teams and towards suspected teams. We include both the *Home Behind* and *Suspected Behind* dummies and an interaction variable (*Home Behind*)*(*Suspected Behind*). Results show that if a home team (but not from among the suspected teams) is behind, referees add 0.472 minutes (p-value=0.000) to extra time. If a suspected team is behind (not at home) referees add 0.381 minutes. *Suspected Behind* is in this case marginally significant at 5% (p-value=0.050). Finally, if a suspected team is behind while playing at home, referees add 0.673 minutes (0.472+0.381–0.179). However, the additional effect with respect to home team favouritism is

not statistically different from zero (the joint significance of *Suspected Behind* and (*Home Behind*)*(*Suspected Behind*) is rejected through a F-Fisher test (F=2.17; p-value=0.115).

We conclude that even if suspected teams appear to receive favours, these favours seem much less strong when compared with home team bias. Psychological pressure by the crowd can be more influential on referees than connections.

In column (4) and (5) we focus exclusively on close matches in which a suspected team plays (*N=182*). The results are substantially the same as column (1) and (2): suspected teams appear to be favoured by referees who add an extra time of around 0.40 minutes (p-value=0.033) when they are behind.

Finally, in order to test the robustness of our results, we checked whether – beyond suspected teams – big teams in general are favoured by referees. We consider as big teams other than four of the five suspected (excluding Reggina) also Inter (Milan) and AS Roma. Results are shown in column (6). The dummy *Big Behind* is positive but now it is not statistically significant (p-value=0.249). Economic or political power by these teams does not appear, *per se*, to influence referees.

5. Concluding Remarks

In this paper we have provided strong empirical evidence showing that soccer referees are not impartial between home and visiting teams. In fact, when home teams are behind in close games, we show that referees tend to add significantly more extra time (around half a minute), giving the home team more chances of equalizing.

This additional injury time is not related to genuine waste of time because when we control for a set of variables which determine injury time (player substitutions, disciplinary sanctions, intensity of the game, other spurious wastes of time) this favouritism remains unchanged in size and statistically significant. We show that a decisive factor in determining favouritism is the absence of a running track which implies physical proximity of the crowd to the referee.

In order to substantiate that this additional time is due to favouritism we conduct some falsification tests: we show that referees do not add more time when this is not crucial for the home team, that is, at the end of the first half or when the home team is behind by two goals. Our results also appear robust to the definition of the *Home Behind* dummy.

In the second part of the paper, taking into account the recent Scandal in Serie A in which some big teams have been accused of trying to select favourable referees and to influence them, we test whether the teams suspected of connections have been favoured by referees: we find evidence that these teams have been helped by referees' lengthening of injury time when they were behind.

However, when we consider jointly the effects of favouritism towards home teams and towards suspected teams, we deduce that favouritism towards "big" teams is much less pronounced than favouritism towards home teams is.

We conclude that social pressure from the crowd in the stadium noisily supporting the home team (due to referees sub-consciously succumbing to pressure or to their desire to conform to the crowd) is the main cause for the biased behaviour of referees.

It is likely that this form of favouritism has had only a limited impact on final results and on championship rankings. 74 goals have been scored in injury time. 42 goals were scored by home teams and 32 by visiting teams. On average, 0.108 goals (74/686) are scored in injury time in a game or 0.030 per minute (0.108/3.625). Since favouritism has lead to longer matches of about half a minute, on average 0.015 goals for each match can be imputed to the home bias referee. The expected number of matches altered are therefore 10.2. Attributing to the referee bias the goals scored after 3.62 minutes, we are able to identify 4 matches in our sample which could be considered as being altered because of refereeing bias in 2003-04[19] season and 6 games in the 2004-05 season[20]. Some teams were effectively helped and others were penalized, but the final rankings would not have been significantly different without referee bias in deciding on the length of extra-time.

However, even if the addition of extra time has probably had minor effects on final results, we consider the allowance of additional extra time as just a form of a more general favouritism towards home teams, albeit the only form measurable, to some extent, objectively. Other forms are more difficult to assess as confirmed by the fact that decisions on penalties, off-sides, etc. often cause endless disputes between experts or supporters after the match, despite watching the episodes on TV, repeatedly, backward and forward and in slow motion.

Our results on referee favouritism are qualitatively similar to those found by Garicano *et al.* (2005) and Sutter and Kocher (2004) and Dohmen (2005), respectively, for the Spanish "Primera Liga" and the German "Bundesliga". However, even though there is rather strong evidence of favouritism towards home teams, Italian and German referees can be judged as being comparatively better than Spanish referees: Garicano *et al.* (2005) find that injury time added at the end of the match is about two minutes greater if the home team is behind. Results for Italian "Serie A" show that "only" half a minute is added by referees. A similar figure is found by Sutter and Kocher (2004) and Dohmen (2005) for German referees. Moreover, similarly to the latter authors we do not find evidence of referees also favouring home teams who are ahead by one goal by shortening the game, while, Garicano *et al.* (2005) on the other

[19] Due to refereeing bias the points in the final ranking for the 2003-04 season were affected for the following teams as follows: Lazio: +2; Lecce +2; Milan +1; Perugia -1; Chievo -1; Bologna -2; Sampdoria -2.

[20] The points in the final ranking for the 2004-05 season were affected for the following teams as follows: Inter +2; Parma +2; Milan +4; Messina +1; Palermo +1; Lazio -1; Atalanta -1; Fiorentina -2; Lecce -1; Sampdoria -1; Reggina -1.

hand, also found this kind of favouritism in Spain. The considerable difference might be due to different financial incentives offered to referees in these championships. Italian and German referees have strong financial incentives to be impartial: they receive for each match a rather high fee (respectively 2250 and 3067) plus travel and accommodation expenses. Referees risk to lose these fees if they are judged partial and are not reappointed in following matches. Garicano et al. (2005) do not report of any financial incentives for Spanish referees.

The results of referee favouritism provide evidence that subjective evaluations can be affected by external factors and that judges can be biased by social factors. A similar result has been found recently, in a different context, by Ichino, Polo and Rettore (2003). They look at whether judgments about firing litigations are based only on a employee's misconduct and find instead that judges' decisions tend to be significantly more favourable to employees in Italian regions where unemployment is higher and vice versa.

This evidence shows that social pressure and psychological motivations are important factors which need to be taken into account when designing incentive contracts and enforcement mechanisms.

References

Akerlof, G. (1980), "A Theory of Social Custom, of Which Unemployment May Be One Consequence," Quarterly Journal of Economics, 94 (June 1980), 749–775.

Baroncelli, A., Lago, U., (2006), "Italian Football", Journal of Sports Economics, 7, 1, pp. 13-28.

Becker, G. and Murphy K. M. (2000), Social Economics: Market Behavior in a Social Environment (Cambridge, MA: Harvard University Press, 2000).

Bernheim, D. (1994), "A Theory of Conformity," Journal of Political Economy, 102, pp. 841–877.

Carmichael, F. and Thomas, D. (2005), "Home-Field Effect and Team Performance", Journal of Sports Economics, 6, 3, pp. 264-81.

Chiappori, P., Levitt, S., Groseclose, T. (2002), "Testing Mixed-Strategy Equilibria When Players Are Heterogeneous: The Case of Penalty Kicks in Soccer", American Economic Review, 92, 4, pp. 1138-1151.

Dohmen, T. (2005), "Social Pressure Influences Decisions of Individuals. Evidence from the Behavior of Football Referees", IZA Discussion Paper, No. 1595.

Duggan, M., Levitt, S. (2002), "Winning Isn't Everything: Corruption in Sumo Wrestling", American Economic Review, 92, 5, pp. 1594-1605.

Ehrenberg R. and Bognanno, M. (1990), "Do Tournaments Have Incentive Effects", Journal of Political Economy, 98, 6, pp. 1307-1324.

Garicano, L., Palacios-Huerta, I. and Prendergast, C. (2001), "Favouritism under Social Pressure", Working Paper NBER, No. 8376.

Garicano, L., Palacios-Huerta, I. and Prendergast, C. (2005), "Favouritism under Social Pressure", Review of Economics and Statistics, 87(2), pp. 208–216.

Ichino, A., Polo, M., Rettore, E. (2003), "Are Judges Biased by Labor Market Conditions", European Economic Review, 47, pp. 913-944.

Milgrom, P., (1988), "Employment Contracts, Influence Activities, and Efficient Organization Design", Journal of Political Economy, 96, pp. 42-60.

Neale, W. (1964), "The Peculiar Economics of Professional Sport", Quarterly Journal of

Economics, 78, 1, pp. 1–14.
Nevill, A., Balmer, N. and Williams, M. (1999), “Crowd influence on decisions in association football”, The Lancet 353, 1416.
Nevill, A.M., Balmer, N.J. and Williams, A.M. (2002), “The influence of crowd noise and experience upon refereeing decisions in football”, Psychology of Sport and Exercise, 3, 261-272.
Nevill, A.M., Newell S. and Gale, S. (1996), “Factors associated with home advantage in English and Scottish Soccer”, Journal of Sports Sciences, 14, 181-186.
Posner, R., (1993), “What do Judges and Justices Maximize? The Same Thing Everybody Else Does”, Supreme Court Economic Review, 3, pp. 1-41.
Prendergast, C. (1999), “The Provision of Incentives in Firms”, Journal of Economic Literature, 37, pp. 7–63.
Prendergast, C., and Topel, R., (1996), “Favouritism in Organizations”, Journal of Political Economy, 104, pp. 958–978.
Sutter, M., and Kocher, M. G. (2002), “Favoritism of agents – The case of referees’ home bias”. Papers on Strategic Interaction 28/2002, Max-Planck-Institute for Research in Economic Systems, Jena.
Sutter, M., and Kocher, M. (2004), “Favouritism of Agents: The Case of Referees' Home Bias”, Journal of Economic Psychology, 25, pp. 461–469.
Szymanski, S. (2003a), “The Assessment: The Economics of Sport”, Oxford Review of Economic Policy, 19, 4, pp. 467-477.
Szymanski, S. (2003b), “The Economic Design of Sporting Contests”, Journal of Economic Literature, 41, 4, pp. 1137-1187.
Tirole, J., (1992), “Collusion and the Theory of Organizations”, in Laffont, J. J., ed., Advances in Economic Theory, Cambridge University Press, Cambridge, pp. 151-206.
Tirole, J., (1999), “Incomplete Contracts: Where do we Stand?”, Econometrica, vol. 67, n. 4, pp. 741-81.
Walker, M., Wooders, J., (2001), “Minimax Play at Wimbledon”, American Economic Review, 91, 5, pp. 1521–38.
Wallsten, T. and Barton, C. (1982). “Processing probabilistic multidimensional information for decisions”, Journal of Experimental Psychology: Learning, Memory and Cognition, 8, pp. 361–384.