Società italiana di economia pubblica

XX CONFERENZA

ECONOMIA DELLA TASSAZIONE
Sistemi tributari, pressione fiscale, crescita

Pavia, Università, 25 - 26 settembre 2008

CORRUPTION AND REFEREE BIAS IN FOOTBALL: THE CASE OF CALCIOPOLI

WALTER DISTASO, LEONE LEONIDA, DARIO MAIMONE ANSALDO PATTI AND PIETRO NAVARRA¶

pubblicazione internet realizzata con contributo della COMPAGNIA di San Paolo

Corruption and Referee Bias in Football: the Case of *Calciopoli*[*]

Walter Distaso[†] Leone Leonida[‡] Dario Maimone Ansaldo Patti[§]

Pietro Navarra[¶]

11 January 2008

Abstract

We test the hypothesis that the impartiality of referees in football matches suffers from two types of biases. First, the unconscious bias, a result of the pressure exerted by suppoters benefiting the teams playing at home. Second, the conscious bias, the outcome of a voluntary action taken by the referee to determine of influence match results. Using data from the Italian 2004/05 *Serie A* season, we show that our hypothesis is empirically supported; our results help also to shed light on different aspects of the *Calciopoli* corruption scandal that exploded in Italy in the summer of 2006.

Keywords: referee bias, contest success function, economics of sport.

JEL: C73, C35, L83.

[*]We are in debt with Valentina Corradi, Abhinay Muthoo, Stefan Szymanski and Tommaso Valletti for helpful suggestions. Usual disclamer applies. Distaso gratefully acknowledges financial support from the ESRC, grant RES-062-23-0311.

[†]*Corresponding Author*. Tanaka Business School, Imperial College London, Exhibition Road, SW7 2AZ, London, UK. Tel.: +44 207 594 3293. Fax: +44 207 594 9189. email: `w.distaso@imperial.ac.uk`.

[‡]Queen Mary University of London and Università di Messina.

[§]University of Essex and Università di Messina.

[¶]CPNSS London School of Economics and Università di Messina.

1 Introduction

In Summer 2006 the nationwide enthusiasm generated by the Italian team's victory of the World Cup in Germany has been overshadowed by a game-fixing investigation that uncovered widespread corruption in football. The scandal, commonly referred to as *Calciopoli*, resolved around twenty months of wiretapped telephone conversations involving key figures of Italian football. The magistrates of the Italian Football Federation (FIGC) formally investigated a total of 41 people and examined 19 matches played in *Serie A* (the top Italian league) during the season 2004/05. The prosecutors believed that there was an organization aimed at influencing the results of the matches.

The FIGC judicial investigation formulated allegations of corruption and fraud against team owners, managers and league officials. The President of FIGC and his deputy resigned soon after the scandal erupted. Two of the top members of the referees' Association stepped down, allegedly thought of being part of the organization. The football clubs involved in the scandal were *AC Milan*, *Fiorentina*, *Juventus* and *Lazio*. All of them have been punished: the clubs were sentenced with points deductions, and *Juventus* was relegated to play the 2006/07 season in a lower division (*Serie B*). Table 1 summarizes the final sanctions inflicted by the FIGC to the four teams.

[Table 1 about here]

Corruption in professional sport is not an isolated phenomenon. For example, in summer 2007 the US sport has been shaken by "...the worst that could happen to a professional sport league", according to the National Basketball Association (NBA) commissioner David Stern. An NBA referee, Timothy Donaghy, was accused of two felonies by a district court in Brooklyn. He admitted to betting on several basketball matches, including the ones he was going to direct. Also, he confessed trying to manipulate results as well as to passing sensitive information to professional gamblers. The investigation is still ongoing to check whether other referees were also involved.

The two phenomena share the fact that two popular sports in team context seems to be plauged with illegal practices; in turn, this is a signal of the increasing diffusion of a misbehavior that needs to be further investigated. In this paper we focus on the Italian case and empirically evaluate whether, during the 2004/05 *Serie A* season, the performance of the four teams involved in the corruption scandal mentioned above benefited from the decisions taken by the referees suspected of fraud (henceforth called biased referees). More specifically, our study's objective is to answer the following questions: did referees' bias affect the performance of the teams involved in the scandal? Is it possible to identify specific interaction effects among individual teams and single biased referees? Can we draw some general conclusions about the overall outcome of the championship?

Research in academic journals has extensively examined the impartiality of referees in football matches. This literature has mainly focused on the analysis of the so called "home team bias", i.e. the extent to which referees tend to be biased in favour of home teams. Several papers document that home pressure tends to influence referees' decisions. This type of advantage has been measured in terms of adding extra time, awarding penalties, allocating yellow and red cards and awarding free-kicks and off-sides. Empirical papers examining the existence of the home team bias seem to be unequivocal and consistent across countries. As far as football is concerned, Garicano et al. (2005) examined the Spanish *Primera Division*, Dohmen (2005) studied the home advantage in the German *Bundesliga* and, finally, Goddard (2005) and Dawson et al. (2007) investigated the English *Premier League*. All these papers support the home team bias hypothesis. Conscious illegal behavior has also been partially studied. Duggan and Levitt (2002) analyse the effects of the existence of illegal practices in sumo wrestling competitions, where the wrestling stables implement game-fixing activities.

In this respect, our framework is different from Dougan and Levitt (2002), where the colusive behaviour is between wrestlers and their stables. In our paper, we adapt standard conflict theory and construct a simple model where the probability of victory in a football match is affected by either a conscious or an unconscious behavior on the part of the referee. Our theory is then tested using data from the 2004/05 Italian Football Season. Given the outcome of the investigations, this represents an unique experimental case study to examine the simultaneous existence of the two forms (conscious and unconscious) of referee bias. An important aspect of the empirical analysis is that, by controlling for the unconscious referee bias, it highlights the full effect of the conscious component.

Some interesting results emerge from our empirical investigation. In line with the literature on the economics of football, the home team bias exerted relevant effects in the Italian national championship: referees decisions, on average, tend to favour the team playing at home. In addition, there are different individual referees' effects that affect the performance of the four teams involved in the scandal. Such effects emerge since we control for all the factors associated by the literature with unconscious referee bias. Hence, the unexplained results seem to be driven by conscious referee bias. These results may indicate the existence of two levels of competition. First, the official level (the *legal championship*), in which competition takes place on the pitch. Second, the unofficial level (the *illegal championship*), in which teams' owners and managers compete in order to corrupt referees and affect the results of the games. The overall outcome of the tournament depends on each team's performance in both the legal and illegal championship.

The paper is organized as follows. In Section 2 we discuss the theoretical and empirical literature on referee bias in sport contests, and motivate our study. In Section 3 we describe our model. In Section 4 we present the data and the empirical framework. In Section 5 we comment on the main results. In Section 6 we discuss the implications of our findings. In Section 7 we draw some concluding remarks.

2 Performance in sport contests and the role of referees

2.1 The related literature

The economics of sport has received large attention from the literature since Rottenberg (1956). Broadly speaking, scholars have focused on two main questions. The first focuses on the optimal design of sport contests. It deals with devising the most appropriate incentive scheme to ensure that all participants unfold their best effort to win the prize. Rosen (1986) models the incentive properties of prizes in sequential elimination tournaments in order to determine the best contestants and support the survival of the fittest. Ehrenberg and Bognanno (1990) empirically examine how the structure of prizes in the Professional Golfers' Association (PGA) influences players' performance and elicits their efforts. The importance of the optimal design of sport competitions is also clarified by Taylor and Trogdon (2002). They empirically test how rule changes for determining the draft order in NBA recruitment can have a negative incentive for teams, which may voluntarily underperform during the season. An extensive review of this branch of literature can be found in Szymanski (2003).

The second issue is more closely related to our study. It examines the impact of the factors that are expected to affect the probability of success in a sport competition. Two variables emerge as the main determinants of the probability of success. First, the strength of the participants/teams. In football, this variable has been typically measured in terms of the number of goals scored and conceded in a match. Moroney (1956) and Reep et al. (1971) adopted a Poisson and a negative binomial distribution, respectively, to model the distributions of the number of goals scored and conceded by the teams participating in a football tournament. Although the use of the relative attack and defense capabilities of the two competing teams has also been adopted in more recent studies (Maher, 1982; Dixon and Coles, 1997; Rue and Salvesen, 2000), Kuypers (2000) prefers to use the cumulative points gained by the competing teams to consider their relative strength.

The second determinant of the probability of success is the so-called *home bias*, defined as "the consistent finding that home teams in sport competitions win over 50% of the games played under a balanced home and away schedule" (Coruneya and Carron, 1992; p. 97). This effect has been extensively investigated by the literature with respect to both team and individual sport contests. Bray and Carron (1993), Holder and Nevill (1997), Nevill et al. (1997) and Koning (2005) find a significant impact of the home advantage on the athletes' performance in alpine skiing, tennis, golf and skating, respectively. However, some of the evidence is not robust, in the case of individual sport contests (Koning, 2005). In those cases the estimation of home bias is troublesome since the athletes generally play elimination tournaments, in which there is no guarantee of reaching the final stage of the competition. The home bias has been studied more effectively in team sport contests. Nevill and Holder (1999) document the home advantage significantly helping the French football team in the 1998 World Cup. Clarke and Norman (1995) and Carmichael and Thomas (2005) show that the home bias is statistically significant in English football matches played in the seasons going from 1981 to 1996. Bray (1999) reaches the same conclusion by examining the National Hockey League (NHL) matches from 1974 to 1993.

According to the literature, home bias is due to five different factors: crowd, learning/familiarity, travel, rule and referees. The crowd factor is defined as the support from the spectators to the home team which, normally, largely outweighs that to the visiting team. This unbalance is expected to affect the relative performance of the players of the two competing teams and, in turn, the outcome of the game. The learning/familiarity factor denotes the advantage for the home team stemming from the knowledge of the environment in which the match is played. Again, this unbalance is expected to alter the likelihood of success. The travel factor is related to the resources and time that the visiting team has to spend to reach the hosting field. Finally, the rule factor acknowledges the fact that, sometimes, the rules of the game may favor the home teams in some sports. For example, in ice hockey the home team is given the last opportunity to make a line change. This implies that the home teams can observe the players chosen by their opponents and perhaps catch a glimpse of their strategy. The team playing away has no such opportunity, and must simply change players without knowledge of what their opponent is doing. Another example of a rule benefiting the home team occurs in baseball, where the home team always bats last to end the game. If the home team is losing going into the last inning, they still have one last chance to win the game.

Given the focus of our study, we are more interested in the role that referees may play in affecting result of a match. Prior literature stressed that the surrounding environment, apart from influencing the outcome directly, may also have an indirect effect by affecting referees' decisions. In other words, because of the pressure the environment is able to exert on the referee, he is likely to treat more favorably the home team. Mohr and Larsen (1998) analyze the results of the Australian football championship and support the existence of the home referee bias, measured as the number of free kicks awarded to the home team as compared to those given to the teams playing away. In an interesting paper, Nevill et al. (2002) conduct an experiment using two groups of qualified referees. Referees are asked to award free kicks while watching the same matches. A first group of referees watched the matches hearing the crowd noise, while the other evaluated the same episodes without sound. The results of the experiment strongly supports the existence of a home referee bias, since the first group of officials sanctioned fewer fouls to home teams, whereas the number of fouls called against the visiting team was unaffected.

Other more recent contributions offer further support to the home referee bias. Sutter and Kocher (2004) find that referees are more inclined to add extra time when the home team is not leading the match; Garicano et al. (2005) show that the referee bias operates in favor of home teams as an outcome of the support of the crowd, measured by the number of spectators watching the match. Finally, Dawson et al. (2007), analyzing football matches in the English Premier League from 1996 to 2003, find that the team playing away is more likely to face higher sanctions from the referees than the home team. It is important to highlight that these results may be used in order to separate the more general home bias from the home referee bias and argue that the former is a necessary condition for the presence of the latter.

While in the literature there is a large support to the hypothesis that referees' decisions are biased in favour of the home team, few contributions argue against it. An example of such a literature is Ridder et al. (1994). They provide empirical evidence of the fact that referees in Netherland professional football divisions do not award higher numbers of red cards against visiting teams.

2.2 Motivation

As noted above, the referee bias is generally viewed by the literature as the effect of external factors, such as the support of spectators for the home team, on the impartiality of the referee's decisions.

In this paper, we have the opportunity of exploiting the information taken from the 2004/05 *Serie A* championship, which formed the evidence for the *Calciopoli* scandal. The wiretapped telephone conversations that uncovered the game-fixing activities involving both referees and team owners and managers highlight a different kind of referee bias, which we call consciuos bias. In this respect, the *Calciopoli* scandal may be considered as a unique experimental case study to evaluate the extent of both conscious and unconscious referee bias and to examine the impact of each of them on the outcome of the championship. More specifically, our empirical analysis will allow us to identify and highlight the effect of the conscious referee bias by controlling for the other factors related to the unconscious type of bias.

It is important to underline the fact that we are evaluating the *ex-post* realized effect of any form of referee bias. Therefore, results that seem to be against the conclusions of the official investigation should not be interpreted as evidence denying the existence of game-fixing activities. Rather, they should be interpreted as a sign of counterbalancing activities or of their ineffectiveness.

3 A simple model

When engaged in a game, the two teams are contenders fighting over a valuable prize (Szymanski, 2003). This section presents a theoretical model describing the determinants of the outcome of a football match.

It is important to note from the outset that several elements may influence the strategic behavior of the teams playing a football match. For example, recall that not all the matches take place simultaneously. Therefore, the results of earlier matches definitely influence the behavior of teams playing later in the day (or season). In addition, consider the fact that some of the teams are contemporarily involved in national and UEFA football matches. These teams, in order to optimally exploit their performing capabilities, have to distribute strategically their effort over national and international games.

Rather than modelling such a complex scenario, which is beyond the objectives of this paper, in this Section we construct a simple theory by relying on some simplifying assumptions. This allows us to sketch the determinants of the probability of victory in football matches. More specifically, we assume that all matches are played at the same time and that each team aims at winning all the matches. These two assumptions imply that the competing teams play every match in isolation, without any concern for the results of other matches already played and/or to be played. The high reward given to victories (3 points) makes it plausible to assume that every team's objective is to win games. Within this framework, we examine the team's optimal behavior. We assume the existence of two teams and define

$$M_i = p_i(e_i, e_j)V_i - c_i e_i, \tag{1}$$

as the payoff for team i from playing a football match. Here, p_i is the probability of victory for team i, with $i \in \{1, 2\}$; $e_i \in (0, V_i)$ is the the effort that team i unfolds in the game; and c_i is an inverse measure of the quality of each team's players. The higher the value of c_i, the weaker the team and *viceversa* (Moldovanu and Sela, 2001). Finally, $V_i \in (0, \infty)$ is the value that team i assigns to victory in the game. This value can be interpreted as the sum of all the gains associated with winning a football match. Both e_i and V_i are expressed in terms of financial resources, physical assets (infrastructure such as the stadium, training centres, etc.) and human and reputational capital (players and prestige) that a team is expected to lose and/or gain as a result of playing the football match.

Victory in a given match may or may not be valued differently by the two teams. The different evaluation can be determined by several factors which include: the different position of the teams in the table, the degree of the competition characterizing the championship, whether the match is played at the beginning or at a different stage of the season and so on. Therefore, we assume:

$$V_1 = \beta V$$

and:

$$V_2 = V,$$

where β is a positive parameter. If $\beta = 1$, both teams assign the same value to victory; if $\beta > 1$, victory for team 1 is more valuable than for team 2 and if $\beta < 1$ the opposite applies.

The probability of winning a game for each team depends on its own effort as well as on the effort spent by the opponent team. The literature suggests the use of the so-called contest success function (CSF) to specify this probability (Szymanski, 2003):

$$p_i(e_i, e_j) = \frac{\alpha_i e_i}{\alpha_i e_i + \alpha_j e_j}, \text{ where } j \neq i, \quad j, i \in \{1, 2\}, \tag{2}$$

and $\alpha_i \in (0, 1)$ are parameters capturing the factors affecting the result of a match, which are not dependent on the teams' effort.

In line with the existing literature, α captures, for example, the home team bias and home referee bias. But it is also possible that α_i and α_j may differ from team to team because the impartiality of the referee is affected by his/her voluntary actions (in this case, his/her bias is conscious). Therefore, α_i and α_j should be considered as aggregate (or reduced form) parameters, able to capture several different factors, including unconscious and conscious referee bias.

Competing teams choose the optimal amount of effort to maximize (1). Substituting (2) into (1), differentiating with respect to e_i and solving, yields the optimal amount of effort for each team:

$$e_1 = \frac{\alpha_1 \alpha_2 c_2}{(\alpha_1 c_2 \beta + \alpha_2 c_1)^2} \beta^2 V \tag{3}$$

and

$$e_2 = \frac{\alpha_1 \alpha_2 c_1}{(\alpha_1 c_2 \beta + \alpha_2 c_1)^2} \beta V. \tag{4}$$

Hence, in equilibrium the optimal level of effort is different for the two teams, since it depends on the relative quality of their players and on the value that they attach to victory. The level of effort would be therefore equalized if $c_1 = c_2 = \beta = 1$.

Substituting back the optimal levels of efforts in the contest function for both teams, we obtain the probability of winning as follows:

$$p_1 = \frac{\beta \alpha_1 c_2}{\alpha_1 \beta c_2 + \alpha_2 c_1} \tag{5}$$

and

$$p_2 = \frac{\alpha_2 c_1}{\alpha_1 \beta c_2 + \alpha_2 c_1}, \tag{6}$$

where we note that the probability of success in a game is increasing in β for team 1 and decreasing in the same argument for team 2. This follows from the fact that, if team 1 assigns a higher value to victory, its effort will be greater and, consequently, its chances of winning the game will increase. The opposite holds for team 2. Furthermore, p_1 is also increasing in c_2 and decreasing in c_1, since a lower quality of the players playing for team 2 increases the chances of success for team 1. Similarly, the lower the quality of players playing for team 1, the weaker the team in the football field and the smaller its probability of winning the game. *Ceteris paribus*, the same holds for team 2.

Substituting the previous values in the payoff function, we obtain the following expected payoffs:

$$M_1 = \frac{\alpha_1^2 \beta^2 c_2^2}{(\alpha_1 \beta c_2 + \alpha_2 c_1)^2} V \tag{7}$$

and

$$M_2 = \frac{\alpha_2^2 c_1^2}{(\alpha_1 \beta c_2 + \alpha_2 c_1)^2} V. \tag{8}$$

Equations (7) and (8) show that the expected payoffs of the two competing teams differ. This is because their probabilities of success are different as highlighted in equations (5) and (6). More specifically, these probabilities depend on the values of the parameters α_1 and α_2, the quality of each team's players and the importance for each team of winning the game:

$$\frac{p_1}{p_2} = \frac{\alpha_1 \beta c_2}{\alpha_2 c_1}. \tag{9}$$

As noted above, the two parameters α_1 and α_2 capture many things unrelated with teams' efforts, including the effect of referees' bias. From equation (9) we derive that, if the two teams have the same quality (i.e. $c_1 = c_2$) and assign the same value to victory (i.e., $\beta = 1$), the result of the match is affected by the ratio α_1/α_2. This ratio indicates the extent of the relative referee bias between the two competing teams.

As noted above, referee bias can be conscious or unconscious. It is very interesting to disentangle and separately identify the effects of the two sources of bias. We do so in our empirical investigation where, by examining the case of the Italian *Calciopoli* scandal, we control for unconscious referee bias and therefore highlight the full extent of the conscious bias component.

4 Data and empirical framework

We base our analysis on the 2004/05 *Serie A* season, in which 20 teams competed for the title. For each match, we collected data on the host team, the result of the game and the name of the referee. Our data consists of 760 team-matches representing 380 total games played during the season. Following Ferrall and Smith (1999) and Duggan and Levitt (2002) we consider each team in a match as observation units. This choice allows us to evaluate whether a single match might be differently played by the two teams according to whether they are involved in game-fixing activities.

The analysis above suggests that the probability of winning a match for one of the two teams playing the game depends on three factors: the value that each of the teams attaches to victory (β), the quality of each team's players (c) and some factors that do not depend on each team's effort (α) and capture the effect of both conscious and unconscious referee bias. We model this probability as follows:

$$\begin{aligned} \Pr(Win)_i &= \mu_1 + \mu_2 Strength_i + \mu_3 Referee_i^{(s)} + \mu_4 Team_i^{(j)} + \\ &\quad + \mu_5 Team_i^{(j)} \times Referee_i^{(s)} + \mu_6 Home_i + \varepsilon_i, \qquad (10) \end{aligned}$$

where the dependent variable $\Pr(Win)$ is the probability of winning a football match. We use binary choice models for estimating parameters, our binary variable being win/not win. However, the possible outcomes of a football match are three: victory, draw and loss. We have decided to merge draw and loss into one outcome (no victory) because of the high reward given to victory (3 points), relative to draw (1 point) or loss (0 points). We checked our results for robustness by estimating ordered logit models, taking into account all the possible outcomes of a match. Results are qualitatively the same.[1]

In Model (10), $Strength_i$ is a measure of the quality of the challenger. This variable is measured as the difference between the points accumulated in the table by the opponent team until the match is played and the average number of points collected by all other teams. Therefore, it takes a positive value if the challenger is stronger than average and a negative value otherwise. Since (on average), the stronger the challenger, the lower the probability of winning, we expect the sign of the associated parameter μ_2 to be negative.

$Referee_i^{(s)}$, for $s = 1, \ldots, 8$ (since there are eight referees involved in *Calciopoli*), is a dummy variable taking value 1 if the referee s is involved in the scandal and 0 otherwise. This variable captures the direct effect exerted by each investigated referee on the likelihood of winning a football match. This analysis might reveal whether the probabilities of winning matches are different according to whether matches are or are not refereed by a biased official.

$Team_i^{(j)}$, for $j = 1, \ldots, 4$ (since four teams have been involved in *Calciopoli*), is a dummy variable assuming value 1 if the team j is involved in the scandal and 0 otherwise. This variable captures the direct effect of involved teams on the probability of winning.

In order to evaluate whether one (or more) of the eight biased referees favoured one (or more) of the four teams involved in the scandal, we constructed the interaction variable $Team_i^{(j)} \times Referee_i^{(s)}$. If the coefficient μ_5 is statistically significant and positive, then the likelihood of winning a match increases when the biased referee is associated with a team involved in *Calciopoli*. This would provide evidence of a positive effect of a biased referee on the performance of a corrupt team.

$Home_i$ is a dummy variable assuming value 1 if the team hosts the match and 0 otherwise. This variable captures, among other effects, the home team bias, and it is a proxy of the unconscious referee bias. Therefore, its associated coefficient, μ_6, is expected to be positive. Since in Model (10) we control for unconscious referee bias through the dummy variable $Home_i$, the variable $Referee_i^{(s)}$ should capture the effect

[1] For more details see Section 5.4.

of conscious bias. Model (10) then tells us how much of the α in (2) can be attributed to conscious and unconscious referees' behavior.

It is important to notice that in Model (10) we make the assumption that each team attaches to victory the same value, i.e. $\beta = 1$. This choice is determined by the unavailability of a suitable criterion to construct an appropriate measure of team-varying β.[2]

5 Empirical results

5.1 Preliminary analysis

In Table 2, we report the results from a preliminary analysis that builds on some descriptive statistics. Column A reports the set of teams we focus on, and Column B the total number of matches for each set of teams, disaggregated by referees according to whether they are (they are not) biased (B_1 and B_2, respectively). Column C displays the average points that each set of teams gained at the end of the season, disaggregated by referee, again according to whether they are (they are not) biased (C_1 and C_2, respectively). Finally, Column D reports the t-statistics to test the null hypothesis that the difference between the two means in C_1 and C_2 is zero.

[Table 2 about here]

The 380 matches of the season lead to a sample of 760 observations (since a single match involves two teams). Out of the whole sample, 274 matches have been directed by a biased referee (35%) and 486 by another referee (65%). The average points in the two different cases is similar (1.33 *vs* 1.34), and in fact the associated test statistic does not reject the null hypothesis of equality of means.

Interestingly, this is also the case when only the subsample of the teams involved in *Calciopoli* is considered. The average points are still similar (1.76 *vs* 1.57), and the

[2]It should be also noted that all the variables employed in the empirical investigation do not show any statistically correlation at 1% level (correlation matrix is available upon request to the authors).

outcome of the test is the same. Therefore, this simple analysis seems to suggest that the biased referees did not, as a group, have a systematic effect on the outcome of matches.

Results partially differ when the analysis is conducted at the team level. *Fiorentina* was directed 13 times by a biased referee and 25 times by other referees. In the first case, *Fiorentina* realized an average of 1 point per game and in the second case 1.16 points. However, the test does not reject the hypothesis that the two means are equal.

Juventus has been directed 20 times by a biased referee and 18 times by other referees. This team had an average of 1.90 points in the first case, and 2.67 in the second case. Moreover, the test strongly rejects the hypothesis that the difference between the two means is zero. Hence, the evidence suggests that *Juventus* has fared remarkably worse when refereed by biased officials.

Lazio has been refereed 13 times by a biased referee and 25 times by other referees. This team realized an average of 1.54 points per game in the first case and 0.96 points in the second case. The associated test, however, does not reject the hypothesis that the two means are equal. This is also the case for *AC Milan*. Its average points per game when refereed by a biased official is higher (2.35 *vs* 1.86), but the test does not reject the null hypothesis of equality.

Summarizing, results from this preliminary analysis suggest that *Fiorentina* achieved a lower (but statistically not significant) average of points per game when refereed by biased officials; *Juventus* achieved a lower (and statistically significant) average of points of per game when refereed by biased officials; *Lazio* and *AC Milan* had a higher than normal (but statistically not significant) average of points per game when refereed by biased officials.

These results, coupled with the evidence that the average number of points per game realized by all the teams playing in *Serie A* does not seem to be influenced by the type of referee directing the match (i.e., those under investigation and those not investigated), suggest that the conscious bias did not have a general effect. Rather (the case of *Juventus*), it seemed to influence the results of football matches in a counterintuitive manner: biased referees damaged the team they were supposed to help.

5.2 Were the investigated referees acting as a group?

The analysis in the previous section misses some important elements. Among these, the most important is the strength of the opponent team. If one really wants to determine the effect of a referee on the outcome of a match, then the strength of the opponent team should be properly controlled for. Take the case of *Juventus*, for example. From the previous section, one concludes that its performance has been worse when refereed by biased officials. But, it could be very well be the case that the matches refereed by biased officials have been played against very strong teams. So, the outcome of the previous test could be driven by the strength of the opponent teams, rather then by a biased referee. Therefore, it is important to disentangle these two effects.

To this aim, we estimate model (10) using a logit specification. Results are reported in Table 3. We estimate six models, according to the level of aggregation of the variable $Team$.

In Column (A), $Team$ is a dummy variable assuming value 1 if the team belongs to one of the four involved in the scandal, and 0 otherwise. Results suggest that the variable $Referee$ has a negative but not statistically significant effect on the probability of winning the match. Therefore, the evidence suggests that the subset of biased referees did not change, on average, the likelihood of winning a match. As expected, the variable measuring the strength of the opposite team is highly significant, and displays the expected sign. Moreover, in line with the empirical evidence emerging from the existing literature (Garicano et al., 2001; Goddard, 2005; Dohmen, 2005; Dawson et al., 2007), results confirm the existence of home team bias. The variable $Home$ is highly significant and displays a positive sign. On average, therefore, playing at home provides hosting teams with significant advantages. The two previous results are robust across all the estimated models.

[Table 3 about here]

The variable $Team$ has a positive sign and it is statistically significant. This possibly reflects the fact that two of the teams involved in the investigation (*AC Milan* and *Juventus*) scored a very high number of wins during the season. The interaction term, $Referee \times Team$, is not statistically different from zero. This provides evidence against the hypothesis that biased referees had an effect on the outcome of football matches.

Of course, these results may be due to some counterbalancing effects. If referees are biased in favour of a particular team but against another one, the effects of their actions on match results might not be statistically significant, if they are taken in aggregate. For this reason, we have conducted the analysis at a more disaggregated level. Results of this exercise are reported in Columns (B) to (F).

Starting from Column (B), *Fiorentina* has a negative (but not significant) sign, reflecting the low number of matches won during the season; the interaction term with $Referee$ is also not significant, suggesting no systematic influence of biased referees on the probability of winning a match.

The variable $Juventus$ is statistically significant and positive. This is not surprising, since the team won a high number of matches in the season. The interaction of this variable with $Referee$ is negative and statistically significant. Therefore, biased referees had a negative impact on the probability of winning games for *Juventus*. This result is very surprising, given that prosecutors argued that the management of this team had established a system of connections able to illegally determine and influence the outcome of the matches and indeed of the entire season to its own advantage.

Column (D) shows that $AC\ Milan$ is statistically significant and positive, again reflecting the good performance of the team in the season. The interaction with the variable $Referee$ is not statistically significant. This result is against the hypothesis that biased referees were associated with a higher probability of winning games.

Column (E) shows for *Lazio* similar results to those already observed for *Fiorentina*. The poor performance of the team in the season determines a negative (but not significant) sign of the variable $Team$; and the interaction term is also not significant. Finally, Column (F) reports results obtained including all the individual teams involved in the investigation. Qualitatively, it is possible to draw the same conclusions as those coming from models (B) to (E).

The evidence seems to suggest that all the teams investigated for trying to illegally influence the results of the matches did not in fact fairly (statistically) better when refereed by biased officials. Furthermore, quite surprisingly our results suggest that *Juventus*, the team accused of masterminding the whole system of corruption, has been the only one which consistently underperformed when directed by those referees.

5.3 A parallel championship?

The rather surprising results from previous sections seem to be consistent with two alternative hypothesis. Either:

1. the alleged organization of team owners, managers and referees created to control the outcome of 2004/05 *Serie A* did not exist or deliver the expected outcome; or:

2. the game-fixing activities are not the result of illegal practices implemented by an organization, but rather the outcome of illegal behavior exerted by individual teams and single referees to manipulate the outcomes of the matches.

The first hypothesis contrasts with the evidence of wiretapped telephone conversations and other game-fixing practices collected by prosecutors and with the final outcome of the trial, involving key powerful figures of the Italian football system. It is also contradicted by the empirical results related to *Juventus*.

What the results in Table 3 might point to is the existence of a parallel championship played outside the football pitch, where managers, team owners and referees compete and interact to determine the outcome of the game. This is what is implied by the second hypothesis. It postulates the contemporary existence of two parallel championships: the first is the official/legal championship in which teams competition takes place in the football field, and the second is the unofficial/illegal championship in which teams' owners and managers compete to corrupt referees.

In the parallel championship each team undertakes a collusive relationship with a given referee or group of referees in order to manipulate the match results. Therefore, each team's overall performance in the official championship is affected by the net effect of the single interactions it entertains in the unofficial championship. In this perspective, the fact that (on aggregate) the performance of three out of the four teams involved in the scandal is not associated with biased referees might be explained by possible individual effects of biased referees compensating each other. Similarly, the fact that the effect of the referee bias on the performance of *Juventus* is, on aggregate, negative, could be a result of the positively biased referees underperforming with respect to the negatively biased ones.

In a context in which two parallel championships are played by a set of teams and referees, the overall outcome of the football tournament depends on each team's performance in both its official/legal and unofficial/illegal component. Therefore, we check whether any systematic relationship exists between each one of the four teams involved in *Calciopoli* and one or some of the eight biased referees. Table 4 reports the results of a preliminary analysis linking individual biased referees and the performance of the teams involved in *Calciopoli*. An inspection of the Table reveals that *Juventus* played the highest number of matches refereed by biased officials, compared to the other three teams. We also note that, in line with the results of difference in means tests shown in Table 2, there is no significant difference between the outcome of the matches refereed by biased and other officials for *AC Milan*, *Fiorentina* and *Lazio*. However, *Juventus* seems to have performed substantially worse when refereed by biased referees. *Juventus* lost only four matches during the 2004/05 *Serie A* season and all of them were directed by biased referees.

[Table 4 about here]

In order to evaluate the effect of individual biased referees, we estimate Model (10) with the disaggregated variable $Referee$ and report the results in Table 5.

[Table 5 about here]

We note that, for example, Ref 2 and Ref 4 are associated with a significantly worse performance of *Juventus*. Ref 5 and Ref 7 are associated with a significantly better performance of *Lazio* and *AC Milan*, respectively. *Fiorentina*'s performance does not seem to be positively or negatively associated with any individual biased referee.

[Table 6 about here]

These results seem to lend some support to hypothesis 2, stated at the beginning of the section. As shown in Table 6, which summarizes the findings of Table 5, some referees' behavior damaged the performance of *Juventus* whereas other referees' behavior benefited *Lazio* and *AC Milan*.

5.4 Robustness checks

A first concern on the results of the analysis we have conducted so far stems from the level of aggregation of the dependent variable. Indeed, our dependent variable is the likelihood of winning a match. Further, our sample consists of all the matches played in the season. Therefore, some of the relationships we are trying to uncover may be hidden. For this reason, we undertake a completely different approach, and focus only on the matches played by the team(s) involved in the scandal.

Results are reported in Table 7, where we present six models. Column (A) is obtained by using observations from all the teams involved in the scandal. In line with our previous conclusions, results suggest that the stronger the opponent team, the lower the likelihood of winning a match. Moreover, the variable capturing the home team bias is statistically significant and positive. Finally, the evidence rejects the hypothesis that biased referees, when considered as a group, have a different impact on the probability of victory with respect to the other referees.

Column (B) reports the results obtained by using observations related to teams untouched by the scandal. Results are very similar to previous ones. Column (C) is obtained by estimating model (10) for *Fiorentina*. The same holds for Columns (D), (E) and (F), where only observations for *Juventus*, *Lazio* and *AC Milan* are, respectively, used. Results generally confirm our analysis.

[Table 7 about here]

The variable $Strength$ exerts a negative impact on the likelihood of winning a match for *Lazio*, whereas in the case of *Fiorentina*, *Juventus* and *AC Milan* it is not significant. This finding can be explained by the fact that these teams have won (or lost, in the case of *Fiorentina*) a large number of matches regardless of the strength of the team they played against. Interestingly, the only team taking advantage of the home team bias is *Fiorentina*. Also, *Juventus* had, again, a significantly lower performance when refereed by biased referees. The opposite holds for *Lazio*.

In Table 8 we report the results obtained with individual biased referees. Results corroborate our conclusions: *Juventus* has consistently under-performed when refereed by Ref 1, Ref 2 and Ref 4, while *Lazio* and *AC Milan* over-performed when refereed by Ref 5 and Ref 7, respectively.

[Table 8 about here]

A second concern on our results is related to the definition of our dependent variable. We recall that we used as dependent variable the probability of winning *vs* the probability of not winning a match. However, even a draw may be the result of the referee conscious bias if, for example, the team he is supposed to favour deserves to loose the match. Therefore, we repeated the entire analysis by estimating an ordered logit model to account for the three possible outcomes of a match.

The findings of this empirical exercise are reported in Table 9. They generally confirm our previous findings. As before, the strength of the challenger and the home team bias are significant determinants of the probability of carrying out a better performance in a match. Their statistical significance and their signs emerge as expected. As before, the variable $Referee \times Team$ is not statistically significant. However, a more disaggregated investigation suggests that *Juventus* obtained a worse performance when the match was refereed by biased officials. On the other hand, *Lazio* over-performed. Finally, both *AC Milan* and *Fiorentina* do not appear to have been significantly benefited or damaged by biased referees. Results on individual referees are maintained: signs and statistical significance of the interaction variables do not change.

[Table 9 about here]

A third and final concern is related to our estimator. If a team untouched by the scandal is particularly weak (or strong), then the reason for its poor (good) performance relies simply on its quality. If this is the case, it is likely that our results are inconsistent because of an omitted variable bias. We overcome this possible bias by employing a probit panel data estimator, which washes up all individual effects. Results are reported in Table 10.

[Table 10 about here]

Also this final check supports our previous results. Playing at home enhances the chances of success. While, on aggregate, referees do not seem to play a significant role in favour of the teams involved in *Calciopoli*, if we disaggregate our analysis taking into consideration teams and referees separately, *Juventus* underperformed with biased referees, while *Lazio*'s performance improved significantly.

6 Discussion

The empirical analysis carried out in the previous section highlights some interesting facts. The most surprising one is that the results seem to be partially in contrast with the official investigation. While the latter uncovered the role played by several key figures of Italian football system, the referee bias seems to exist only for two of the four teams sanctioned by FIGC magistrates. Moreover, estimates show that *Juventus* underperformed when directed by a biased referee. The contrast between the conclusion of the official investigation and the empirical results can be reconciled by considering counterbalancing individual effects of biased referees.

Generally speaking, the existence of referee bias has important implications not only for the fairness of a single match, but also for the overall fairness of the entire sport tournament. In Figure 1 we show how the behavior of referees may lead to different results according to whether they are or they are not biased and, if they are, according to whether their actions are conscious or unconscious.

[Figure 1 about here]

Consider the case of a national football championship. Since referees are always required to act as impartial agents, the team winning the title fairly deserves it if all the matches of the championship are conducted by unbiased referees. However, referees' decisions can be biased. For example, think of the well documented home bias effect. It may be argued that, since referees favour the team playing the game at home, the result of every single match of the national championship is likely to be biased. But, since each team plays the same number of matches at home during the season, the title at the end should still be fairly assigned. In this case, the net effect of the home team referee bias on the overall outcome of the championship is likely to be zero.

The same rationale does not apply to the case in which the referee bias is conscious, i.e., when it is the result of a voluntary behavior on the part of the referee. Since the very nature of referees is to act as impartial agents, the possibility of consciously biasing the result of a match is unfair. This implies that, when referees are consciously biased, the overall outcome of the championship ought to be considered unfair.

Conscious referee bias could be more or less effective in distorting results of matches. While in the latter case the final outcome of the championship might still be fair, in the former case the fairness of the championship may depend on the behavior of the other referees. It could be argued, in fact, that the case of a parallel championship *played* by all the team owners and managers in order to guarantee themselves friendly referees might still produce fair results. This is because referees' biased actions could counterbalance. The overall fairness of the championship is definitely jeopardized if only one or a few teams corrupt referees to manipulate match results (see Figure 1).

7 Concluding remarks

In the summer of 2006 a game-fixing scandal shook Italian football. Team owners and managers were suspected to corrupt referees in order to manipulate match results. In this paper, we used a standard model borrowed from conflict theory to describe how the outcome of a football match may be affected by factors directly associated with the relative quality of the competing teams, as well as by biased decisions of referees.

Typically, the literature on sport contests shows that referees may unconsciously affect the result of a match. This is the case of the home team bias in which referees' decisions benefit the team playing at home. Motivated by the *Calciopoli* scandal exploded in Italy in the summer 2006, we also allow referees to consciously affect the outcome of the games.

We then use the data on 2004/05 *Serie A* season to highlight the presence of these two types of bias, to quantify their effects on teams' performance and to examine and discuss the implications of our results on the overall fairness of the championsip. The findings of our paper are important to understand and evaluate growing episodes of suspected game-fixing and fraud in professional sports.

References

Bray, S. R. (1999). The home advantage from an individual team perspective. *Journal of Sport and Exercise Psychology*, 11, 116–125.

Bray, S.R. and A.V. Carron (1993). The home advantage in alpine skiing. *Australian Journal of Science and Medicine in Sport*, 25, 76–81.

Carmichael, F. and D. Thomas (2005). Home-Field effect and team performance: Evidence from English Premiership football. *Journal of Sport Economics*, 6, 264–281.

Clarke, S.R. and J.M. Norman (1995). Home ground advantage of individual clubs in English soccer. *The Statistician*, 44, 509–521.

Courneya, K.S. and A.V. Carron (1992). The home advantage in sport competitions: a literature review. *Journal of Sport and Exercise Psychology*, 14, 13–27.

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

Dixon, M.J. and S.C. Coles (1997). Modelling association football scores and inefficiencies in the football betting market. *Applied Statistics*, 46, 265–280.

Dobson, S. and J. Gobbard (2001). *The economics of football.* Cambridge, Cambridge University Press.

Dohmen, T.J. (2005). Social pressure influences decisions of individuals: evidence from the behavior of football referees. IZA Discussion Papers 1595.

Duggan, M. and S.D. Levitt (2002). Winning isn't everything: Corruption in sumo wrestling. *American Economic Review*, 92, 1594–1605.

Ehrenberg, R.G. and M.L. Bognanno (1990). Do tournaments have incentive effects? *Journal of Political Economy*, 98, 1307–1324.

Farrell, C. and A.A. Smith (1999). A sequential game model of sport championship series: Theory and estimation. *Review of Economics and Statistics*, 81, 704–719.

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

Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21, 331–340.

Holder, R.L. and A.M. Nevill (1997). Modelling performance at international tennis and golf tournaments: Is there a home advantage? *The Statistician*, 46, 551–559.

Koning, R.H. (2005). Home advantage in speed skating: evidence from individual data. *Journal of Sports Sciences*, 23, 417–427.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, 32, 1353–136.

Maher, M.J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36, 109–118.

Maroney, M.J. (1956). *Facts from figures.* London, Penguin Press.

Mohr, P.B. and K. Larsen (1998). Ingroup favoritism in umpiring decisions in Australian Football. *Journal of Social Psychology*, 138, 495–504.

Moldovanu, B. and A. Sela (2001). The optimal allocation of prizes in contests. *American Economic Review*, 91, 542–558.

Nevill, A.M., Balmer, N.J. and A.M. Williams (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, 3, 261–272.

Nevill, A.M. and R.L. Holder (1999). Home advantage in sport: an overview of studies on the advantage of playing at home. *Sports Medicine*, 28, 221–236.

Nevill, A.M., Holder, R.L., Bardsley, A., Calvert, H. and S. Jones (1997). Identifying home advantage in international tennis and golf tournaments. *Journal of Sports Sciences*, 15, 437–443.

Ridder, G., Cramer, J.S. and P. Hopstaken (1994). Down to ten: Estimating the effect of a red card in soccer. *Journal of the American Statistical Association*, 89, 1124–1127.

Reep, C., Pollard, R. and B. Benjamin (1971). Skills and chance in ball games. *Journal of Royal Statistical Society, Series A*, 131, 581–585.

Rosen, S. (1986). Prizes and incentives in elimination tournaments. *American Economic Review*, 76, 701–715.

Rottenberg, S. (1956). The baseball player's labor market. *Journal of Political Economy*, 64, 242–58.

Rue, H. and O. Salvesen (2000). Prediction and retrospective analysis of soccer matches in a league. *The Statistician*, 49, 399–418.

Sutter, M. and M.G., Kocher (2004). Favoritism of agents – The case of referees' home bias. *Journal of Economic Psychology*, 25, 461–469.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, 41, 1137–51.

Taylor, B.A. and J.G Trogdon (2002). Losing to win: Tournament incentives in the National Basketball Association. *Journal of Labor Economics*, 20, 23–41.

Table 1

Sanctions inflicted by the FIGC to the teams involved in the scandal

Team	Relegation	Points Deduction	Other Sanctions
Juventus	Serie B	9 points	Stripped of the 2005/06 title; Out of the 2006/07 UEFA Champions League; 3 home games behind closed doors
AC Milan	None	8 points	Deducted 30 points in 2005/06 season; 1 home game behind closed doors
Fiorentina	None	15 points	Out of the UEFA Champions League; 2 home games behind closed doors
Lazio	None	3 points	Out of the UEFA Cup; 2 home games behind closed doors

Table 2

Test for Difference in Means

Column A reports the sub set of team we use to test the null hypothesis that the difference in means is not statistically significant. Column B reports the total number of observations for each subset of teams, both when the match has (B_1) or has not (B_2) been refereed by a biased referee. Column C reports the average points for each subset of teams, both when the match has (C_1) or has not (C_2) been refereed by a biased referee. Finally, Column D reports the t-statistics to test the null hypothesis that the difference among the two means is not statistically different. *, ** and *** stand for rejection of the null hypothesis at 10%, 5% e 1% level, respectively.

Teams	Number of Observations			Average Points			t-statistic
A	B	B_1	B_2	C	C_1	C_2	D
1 All the Teams	760	274	486	1.34	1.33	1.34	-0.06
2 Teams Involved in *Calciopoli*	152	63	89	1.65	1.76	1.57	0.89
3 Other Teams	608	211	397	1.25	1.20	1.28	-0.77
4 Fiorentina	38	13	25	1.11	1.00	1.16	-0.40
5 Juventus	38	20	18	2.26	1.90	2.67	-2.19**
6 Lazio	38	13	25	1.16	1.54	0.96	1.35
7 Milan	38	17	21	2.08	2.35	1.86	1.28

Table 3
Determinants of the Probability of Winning a Match:
Teams Under Investigation

The dependent variable is the probability of winning a match. Number of observations: 760. Standard Errors are in parentheses. *, ** and *** stand for significance at 10%, 5% and 1% level, respectively.

Variables	(A)	(B)	(C)	(D)	(E)	(F)
Constant	-1.390***	-1.255***	-1.390***	-1.249***	-1.344***	-1.410***
	(0.152)	(0.142)	(0.146)	(0.142)	(0.144)	(0.154)
Strength	-0.074***	-0.075***	-0.074***	-0.074***	-0.075***	-0.074***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Home	0.971***	0.941***	0.972***	0.938***	0.967***	1.007***
	(0.166)	(0.164)	(0.167)	(0.164)	(0.166)	(0.171)
Referee	-0.092	0.062	0.077	0.012	0.000	-0.094
	3.000	(0.173)	(0.177)	(0.174)	(0.177)	(0.198)
Team	0.467*					
	(0.253)					
Referee × Team	0.532					
	(0.404)					
Fiorentina		-0.554				-0.444
		(0.491)				(0.494)
Fiorentina × Referee		0.045				0.220
		(0.877)				(0.884)
Juventus			2.561***			2.567***
			(0.698)			(0.701)
Juventus × Referee			-1.579*			-1.410*
			(0.851)			(0.856)
Lazio				-0.729		-0.610
				(0.536)		(0.539)
Lazio × Referee				1.219		1.329
				(0.804)		(0.812)
Milan					0.964**	1.011**
					(0.472)	(0.475)
Milan × Referee					0.791	0.892
					(0.757)	(0.763)

Table 4

Teams and Referees

The table reports the number and results of matches played by the teams involved in *Calciopoli*, disaggregated by referees. W stands for a victory, N stands for a draw, and L stands for a defeat. Note that the names of the referees are not spelled out to safeguard their privacy.

	Fiorentina				Juventus				Lazio				Milan			
Referee	W	N	L	T	W	N	L	T	W	N	L	T	W	N	L	T
Ref 1	0	1	2	3	1	2	0	3	0	0	0	0	4	2	0	6
Ref 2	1	1	1	3	2	1	2	5	0	0	0	0	2	0	1	3
Ref 3	1	0	2	3	2	0	0	2	1	0	1	2	0	0	0	0
Ref 4	0	0	0	0	2	1	1	4	2	1	1	4	1	0	0	1
Ref 5	1	1	0	2	2	0	1	3	1	0	2	3	3	0	0	3
Ref 6	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
Ref 7	0	1	1	2	2	1	0	3	0	0	0	0	2	2	0	4
Ref 8	0	0	0	0	0	0	0	0	2	0	1	3	0	0	0	0
Biased Referees	3	4	6	13	11	5	4	20	7	1	5	13	12	4	1	17
Other Referees	6	11	8	25	15	3	0	18	5	9	11	25	11	6	4	21
All Referees	9	15	14	38	26	8	4	38	12	10	16	38	23	10	5	38

Table 5

Determinants of the Probability of Winning a Match: Biased Referees and Teams Involved in *Calciopoli*

The dependent variable is the probability of winning a match. Number of observations: 760. *, ** and *** stand for significance at 10%, 5% and 1% level, respectively.

Variables	Coefficient (S.E.)	Variables	Coefficient (S.E.)
Constant	-1.449*** (0.156)	Fiorentina × Ref 7	1.697 (1.664)
Strength	-0.075*** (0.013)	Fiorentina × Ref 6	-1.056 (1.696)
Home	1.050*** (0.176)	Juventus × Ref 1	-1.758 (1.511)
Fiorentina	-0.535 (0.495)	Juventus × Ref 2	-2.139* (1.277)
Juventus	2.647*** (0.699)	Juventus × Ref 3	0.221 (1.750)
Lazio	-0.639 (0.532)	Juventus × Ref 4	-2.679** (1.303)
Milan	1.144** (0.466)	Juventus × Ref 7	-0.868 (1.501)
Ref 1	-0.587 (0.463)	Juventus × Ref 6	-0.668 (1.520)
Ref 2	0.262 (0.441)	Lazio × Ref 3	1.535 (1.861)
Ref 3	0.584 (0.452)	Lazio × Ref 4	0.780 (1.279)
Ref 4	0.498 (0.414)	Lazio × Ref 7	0.147 (1.391)
Ref 5	-0.506 (0.994)	Lazio × Ref 5	2.797* (1.655)
Ref 6	-0.566 (0.454)	Lazio × Ref 8	1.656 (1.510)
Ref 7	-0.474 (0.408)	Milan × Ref 1	0.970 (1.191)
Ref 8	0.204 (0.666)	Milan × Ref 2	0.582 (1.468)
Fiorentina × Ref 1	-0.884 (2.001)	Milan × Ref 4	0.769 (1.698)
Fiorentina × Ref 2	0.425 (1.523)	Milan × Ref 7	2.833* (1.658)
Fiorentina × Ref 3	0.663 (1.457)	Milan × Ref 6	0.091 (1.263)

Table 6

Impact of Referees on Teams

Team	Positive Impact	Negative Impact
Fiorentina	-	-
Juventus	-	Ref 2; Ref 4
Lazio	Ref 5	-
Milan	Ref 7	-

Table 7

The Alternative Model: Probability of Winning a Match

The dependent variable is the probability of winning a match. In estimating models (A) and (B) we have used matches played by a team involved (not involved) in *Calciopoli*, respectively. In estimating models (C), (D), (E) and (F) we have used matches played by Fiorentina, Juventus, Lazio and AC Milan, respectively. *, ** and *** stand for significance at 10%, 5% and 1% level, respectively.

Variables	(A)	(B)	(C)	(D)	(E)	(F)
Constant	-0.764**	-1.440***	-2.646***	1.067	-1.907**	0.204
	(0.299)	(0.163)	(0.923)	(0.723)	(0.753)	(0.603)
Strength	-0.086***	-0.070***	-0.142	-0.061	-0.171**	-0.066
	(0.026)	(0.014)	(0.086)	(0.047)	(0.081)	(0.047)
Home	0.649*	1.070***	1.961*	1.304	0.125	-0.089
	(0.349)	(0.190)	(0.974)	(0.818)	(0.826)	(0.706)
Referee	0.410	-0.098	0.074	-1.545*	1.455*	0.583
	(0.352)	(0.198)	(1.010)	(0.845)	(0.852)	(0.723)

Table 8

The Alternative Model: Teams and Referees

The dependent variable is the probability of winning a match. We have used matches played by a team involved in *Calciopoli* (A), and by the remaining 16 teams (B). Models (C), (D), (E) and (F) were obtained selecting matches played by Fiorentina, Juventus, Lazio and AC Milan, respectively. *, ** and *** stand for significance at 10%, 5% and 1% level, respectively.

Variables	(A)	(B)	(C)	(D)	(E)	(F)
Constant	-0.770**	-1.448***	-4.142**	1.001	-2.181**	-0.072
	(0.303)	(0.163)	(1.532)	(0.715)	(0.889)	(0.677)
Strength	-0.087***	-0.071***	-0.255*	-0.047	-0.233**	-0.071
	(0.026)	(0.014)	(0.138)	(0.050)	(0.102)	(0.049)
Home	0.692*	1.083***	3.424**	1.968	-0.065	0.489
	(0.357)	(0.192)	(1.557)	(1.024)	(1.004)	(0.838)
Ref 1	0.055	-0.611	-12.690	-2.919*		0.365
	(0.673)	(0.460)	(14.95)	(1.604)		(1.166)
Ref 2	0.407	0.224	0.576	-2.235*		1.243
	(0.699)	(0.437)	(1.678)	(1.262)		(1.539)
Ref 3	1.057	0.520	2.884	2.655	4.175	
	(0.872)	(0.451)	(2.212)	(4.564)	(3.654)	
Ref 4	0.354	0.408		-2.727*	1.958	17.105
	(0.730)	(0.418)		(1.448)	(1.427)	(11.09)
Ref 5	0.326	-0.463	0.824	-1.127	-3.278	9.590
	(0.666)	(0.411)	(2.833)	(1.531)	(4.795)	(12.14)
Ref 6	2.475	-0.867			28.733	
	(2.763)	(1.149)			(17.090)	
Ref 7	-0.055	-0.526	-15.171	-1.579		-0.255
	(0.783)	(0.453)	(15.63)	(1.583)		(1.201)
Ref 8	0.827	0.247			1.685	
	(1.262)	(0.668)			(1.595)	

Table 9

Robustness Check I: Orderd Probit Analysis

The dependent variable is the probability of winning a match. Number of observations: 760. Standard Errors are in parentheses. *, ** and *** stand for significance at 10%, 5% and 1% respectively.

Variables	(A)	(B)	(C)	(D)	(E)	(F)
Constant	0.116	0.169**	0.123*	0.191***	0.137*	0.125*
	(0.073)	(0.070)	(0.070)	(0.071)	(0.070)	(0.074)
Strength	-0.043***	-0.044***	-0.043***	-0.043***	-0.044***	-0.043***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Home	0.581***	0.564***	0.581***	0.572***	0.574***	0.594***
	(0.084)	(0.084)	(0.084)	(0.084)	(0.084)	(0.085)
Referee	-0.030	0.060	0.043	-0.042	0.022	-0.031
	(0.099)	(0.090)	(0.090)	(0.090)	(0.090)	(0.099)
Team	0.246*					
	(0.133)					
Referee × Team	0.282					
	(0.218)					
Fiorentina		-0.105				-0.060
		(0.240)				(0.241)
Fiorentina × Referee		-0.252				-0.154
		(0.397)				(0.399)
Juventus			1.445***			1.456***
			(0.355)			(0.356)
Juventus × Referee			-0.831*			-0.761*
			(0.451)			(0.453)
Lazio				-0.517**		-0.450*
				(0.231)		(0.232)
Lazio × Referee				1.112**		1.151**
				(0.449)		(0.451)
Milan					0.676**	0.695***
					(0.263)	(0.264)
Milan × Referee					0.205	0.265
					(0.410)	(0.412)

Table 10
Robustness Check II: Panel Data Analysis

The dependent variable is the probability of winning a match. Number of observations: 760. Standard Errors are in parentheses. *, ** and *** stand for significance at 10%, 5% and 1% level, respectively.

Variables	(A)	(B)	(C)	(D)	(E)	(F)
Strength	-0.045***	-0.046***	-0.046***	-0.045***	-0.045***	-0.045***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Home	0.606***	0.598***	0.602***	0.603***	0.604***	0.608***
	(0.102)	(0.102)	(0.102)	(0.102)	(0.102)	(0.102)
Referee	-0.156	-0.101	-0.093	-0.143	-0.135	-0.157
	(0.119)	(0.111)	(0.109)	(0.116)	(0.110)	(0.119)
Referee × Team	0.271					
	(0.280)					
Fiorentina × Referee		-0.118				-0.059
		(0.464)				(0.466)
Juventus × Referee			-0.471**			-0.408**
			(0.235)			(0.181)
Lazio × Referee				0.731*		0.745**
				(0.372)		(0.349)
Milan × Referee					0.743	0.766
					(0.598)	(0.599)

Figure 1

Behaviour of the Referees and Fair Outcome of the Championship: Alternative Cases