Title: Match rigging and the favorite long-shot bias in the Italian football betting market.

Author: Marco Rossi

Affiliation: Department of Law, University "La Sapienza", Roma.

Address: Piazzale Aldo Moro 5, 00185 Roma, Italy.

Phone: ++393201889856.

E-mail: rotorossi@yahoo.com

Abstract.

In this empirical study we compared the results of the matches played in the Italian football league "Serie A" with the odds offered by the bookmakers. We found that the market odds are good predictors of the actual game results but we also found that the distribution of returns for odds' sub-groups displayed the so-called "favorite long-shot" bias. Given the evidence of match-rigging in Italian football, we investigated if this bias was caused by a strategic behavior of bookmakers who were expecting to deal with insiders. Our results confirm that match rigging was associated with a larger F/L bias.

1. INTRODUCTION.

Wagering on sport events is a very old Italian tradition: in the Imperial times, the Circus Maximus in Rome drew crowds of over 250 thousands of people, anxious about their bets, mainly on chariot races[i]. Today, after its de facto liberalization which occurred in Italy from the year 2000, the volume of wagering on sport events is increasing at a very large average annual rate (about 60% each year from 1998 until 2008), and amount of almost 4 billion of euro was spent in Italy on sport betting in the year 2008 (graph 1)[ii]. Here we study the Italian soccer betting market, which is by far the most important slice of the Italian sport betting market (93% of the sport bets placed in the year 2008 were on soccer events). Our datasets consist of the results and the odds of the soccer games played in the highest Italian soccer league: the "Serie A" championship. The first dataset consists in the results and the odds posted by one bookmaker on 6369 games played from the 2002/03 season until the 2007/08 season. The second dataset consists in the results and the odds offered by three bookmakers on 289 games played in the 2007/08 season. Football bets are simple financial assets: they have a short and well-defined end-point (usually a week), when their value becomes certain, and there is no secondary market for bets. These factors avoid bubbles in the football betting market, simplifying the pricing problem. We start our analysis by assuming that the betting market is a fair game populated by rational representative agents. If this is the case, according to the "Constant Expected Returns Model" (CERM henceforth) the expected return from a unit bet on any event should be 1, and each odd should be the inverse of the frequency of the associated event. The latter hypothesis is empirically investigated in this paper, and we found evidence of a very high predictive efficiency of the odds. But the representative agents which populate the CERM share the same information set, an assumption which seems implausible, particularly in the case of Italian football. Actually, Italian football was plagued by several cases of insider trading, the most famous ones are known as "Calcioscommesse" (1980), "Calcioscommesse 2" (1986), while the last case of match-fixing, "Calciopoli" (2006), had scope and consequences bigger than any other one before. When there is a chance of match-fixing the bookmaker faces an adverse selection problem in which

a customer may be trading on the basis of superior information. In this case the bid-ask spread is determined in a trade-off between setting a large spread so as to minimize the profit of insider traders, and setting the optimal spread against the noise or liquidity traders[iii]. There are several formal analysis of this problem, here we quote the Shin's (1991) model. In its benchmark case, where there are no insiders, the odds are inversely proportional to the true probabilities, as it is in the CERM. But, if there is a chance that the bettor knows more about the outcome of the race than the bookmakers, the optimal pricing response by the bookmakers is a "square root rule" by which they trim the odds on long-shots relative to favorites. Moreover, if bookmakers expect insider trading to be more prevalent given that a long shot is tipped to win, the betting odds should understate the winning chances of a favorite relatively less than the winning chances of a long shot (Shin, 1992). This "favorite-long shot bias" implies that the bookmaker returns on favorites should be less than those on long-shots. The analysis of both our datasets found evidence of a "favorite-long shot bias" in the distribution of returns. In the literature there are several demand-side explanations for the existence of a "favorite-long shot bias" in betting markets (as bettors' local risk love and/or behavioral attitudes, etc.) [iv], but, given the well-known evidence of match-fixing in Italian football, we can use our data to check if the rumors of match rigging increased the F/L bias. Therefore we first investigated if the bias was larger for the suspected matches than for the other ones; then we investigated if the bias was larger for the matches of the club (Juventus) whose managers ran the rigging; and finally we checked if the bias was larger before the uncovering of the scandal than after the ban of the guilty managers. When we identified the rigged matches only with those investigated by the judicial authorities (about 80 cases in the 2004-05 season), the first hypothesis was empirically rejected, but this may be caused by under-sampling. The evidence of rigging in some suspicious cases was sufficient for the court to support the hypothesis that rigging was a systematic mis-behavior of the Juventus management, so that, even if the matches of the 2005-06 season were not under investigation, the "scudetto" won by Juventus was removed. The verdict confirmed the longstanding widespread rumor among Italian soccer fans about a systematic attitude of referees in

favor of Juventus[v]. Indeed, when we expanded the examination to the whole before "Calciopoli" dataset (1366 matches, from the 2002-03 season until the 2005-06 season), we found that the bias was much larger for the matches of the club (Juventus) whose managers ran the rigging than for the rest of the clubs. Finally, when we split the dataset into the two subsets: before and after "Calciopoli", we saw that the F/L bias was slightly larger before the uncovering of the scandal than after the ban of the guilty managers. Summing up, we took the Calciopoli case as a kind of natural experiment, where it was common knowledge that match rigging was on the run until 2006, and that rigging was focused in favor of one team (Juventus), and we found some empirical evidence in favor of the hypothesis that the F/L bias may also be caused by a supply side optimal pricing strategy by bookmakers who face the risk of dealing with insiders.

This paper consists of the current introduction, a summary of the literature on the "favorite long-shot" bias and its candidate explanations, an illustration of the "Calciopoli" case, an empirical analysis, and final conclusions. I am grateful to Agipronews and Sisal for their kindly supply of data, and to the anonymous referees.

2. "The Favorite long-shot bias".

The favorite-long shot bias (hereafter F/L bias) is a systematic tendency by subjects to under-bet or undervalue events characterized by high probability, and to over-bet or overvalue those with low probability. Evidence of a F/L bias was found in the laboratory experiments of Preston and Baratta (1948), Yaari (1965), Rossett (1971), Piron and Smith (1995), Hurley and McDonough (1995). Field evidence of the F/L bias was found in Us horseraces wagering markets by Griffith (1949), McGlothin (1956), Hoerl e Fallin (1974), Ali (1977), Snyder (1978), Asch, Malkiel e Quandt (1982,), Thaler and Ziemba (1988), and similar evidence was also found in UK racetrack wagering markets by Figgis (1951), Dowie (1976), the Royal Commission on Gambling (1978), Henery (1985), Vaughan Williams e Paton (1996, 1997). The hypothesis of a F/L bias was instead rejected by studies of the Japanese and Hong Kong horse-races wagering markets (Busche (1994) and

Busche e Hall (1988)), Us small racetrack (Swindler and Shaw (1995)), Us baseball and Hockey wagering (Woodland and Woodland (1994 and 2001)), and Australian football wagering (Schnytzer and Weinberg (2008)). For a review of empirical studies about the F/L bias see Snower and Wolfers (2007). Summing up, the F/L bias is a quite common (but not universal) feature of sport betting markets.

Several demand-side explanation have been suggested for the F/L bias. The evidence of a favorite long shot bias is not compatible with a model in which representative bettors maximize a function that is linear in probabilities and linear in payoffs. A demand side explanation of the bias can be based on a representative bettor with either a (locally) concave utility function, or a subjective utility function employing non linear probability weights. The models which assume a (locally) concave utility function may be justified by the observation that in any lottery the amount returned to the winners is less than the sum of all bets, the difference is the bookmaker's profit, and it is called "take out". The take out comes from the difference between the odds posted by the bookmaker and their "fair value"[vi]. So lotteries have a negative expected return for bettors, and, according to economic theory, only risk-loving agents could buy negative expected return assets[vii]. But the people who buy lottery tickets are the same people who buy home insurances, so they act as risk-lovers for small stakes and as risk-averter for high stakes. Friedman and Savage (1948) explanation is based on the assumption of a convex (increasing marginal utility) segment in the middle of an otherwise concave utility function. Hence, individuals in the first concave segment are predicted to purchase low probability, high payoff gambles that reach well into the convex segment, while simultaneously insuring against wealth-decreasing risks. Markowitz (1952) refined this approach by placing the convex segment of utility at current wealth, allowing all segment of the income distribution to make rational gambles. Evidence of this "reference dependence" was found in the experimental analyses of Kanheman and Tversky (1979, 1991), Machina (1982), and Camerer (1989). Finally, the "Prospect Theory" (Kanheman and Tversky, 1979) argue that the

curvature the value function is steeper in the loss domain than in the gain domain, a candidate explanation for the F/L asymmetry[viii].

Racetrack wagering models based on local risk loving attitudes were tested and not rejected by Weitzmann (1965), Alì (1997), and Quandt (1986). Finally, Quandt (1986) proved that, if we assume that bettors are local risk lovers, a corollary necessary condition for a pari-mutuel market equilibrium is that favorites should yield higher expected returns than long shots.

The F/L bias could also originate from bettors' loss-aversion, a behavioural attitude that can make bettors act as risk lovers in order to close their betting day without losses. In racetrack wagering, for the bettors who are losing at the end of the day, the last race provides them with a chance to recoup losses. If bettors are loss-averter, they underbet the favourite more than usual, and overbet horses at odds that would eliminate their losses (Kahneman and Tversy (1979) and Thaler and Ziemba (1988)). As a result of these behaviour, horses whose odds shorten ("bet down") are more likely to be favourites, and those who lengthen long shots. The empirical studies of McGlothin (1956), Alì (1977), Asch, Malkiel and Quandt (1982) support this loss aversion explanation while Snowberg and Wolfers (2010) did not find that the bias was more pronounced in the last race of the day in 5.610.580 horse race starts in the United States from 1992 to 2001: "If there were evidence of loss aversion in earlier data, it is no longer evident in recent data, even as the favourite-long shot bias has persisted" (Snowberg and Wolfers, 2010, p.5-6).

Another behavioral explanation of the F/L bias steps from the observation that: "… people adopt mental accounts and act as if the money in those accounts is not fungible" (Thaler and Ziemba, 1988, p. 171). Therefore, if bettors discount a fixed fraction of their losses, the favorite long-shot bias can arise as a consequence of bettors underweight their losses compared to their gains. As Henery (1985) and Williams and Patton (1997) argue, it may be that bettors discount a constant proportion of the gambles in which they bet on a loser. Because long shots lose more often, this discounting yields perceptions in which betting on a long shot seems more attractive.

The F/L bias could instead originate from systematic biases in bettors' subjective probabilities (s_i). Behavioural studies found that people are systematically poor at discerning between small and tiny probabilities (Slovic, Fishoff and Lichtenstein, 1982), and hence price both similarly, and that people exhibit a strong preference for certainty over extremely likely outcomes (Kahneman e Tversky (1979), leading highly probable gambles to be underpriced. A graphical illustration of the point may be fig.2, where the odds' implied probabilities ($0 \leq p_i = 1/Q_i \leq 1$) are compared with the bettors' subjective probabilities ($0 < s_i < 1$). While the 45° line shows the linear probability case ($s_i = p_i$), the misperception bias makes bettors under-price favourites and overprice long-shots.

While in the experiments of Viscusi and O'Connor (1984) and Dwyer (1993) this behavioural bias disappears as subjects get experienced with the game, an empirical support to this misperception explanation comes from the field study of Snowberg and Wolfers (2010) about the Us racetrack wagering market. By using a very large dataset, they compared the neoclassical explanation of the F/L bias (bettors are rational and locally risk lovers), with the alternative behavioural explanation (bettors are irrational and risk neutral) and they claim to have found evidence in favour of the view that misperceptions caused the favourite-long shot bias.

An alternative supply side explanation for the F/L bias is based on the hypothesis of asymmetric information distribution among the traders. If there is asymmetric information, the bookmakers should act strategically in order to minimize the losses they face when they deal with informed bettors. Shin (1991) showed that it is then optimal for the bookmakers to employ a "square root rule" in which the ratio of the posted price is set equal to the square root of the ratio of winning probabilities: $p_i/p_j = \sqrt{(f_i/f_j)}$. One consequence of this rule is that the betting odds tend to understate the winning chances of favorites, and to overstate the winning chances of long-shot. Shin (1992) formalized an extensive-form game in which there is a constant probability (z) that the bookmaker will face an informed bettor who knows exactly the outcome of the game. If this is the case, the problem for the price-setting bookmaker is to set the odds ($Qi = 1/pi$) such as to maximize:

$1 - [\Sigma z_i f_i + (1-z_i)f^2_i/p_i]$,

where p_i is the price of the basic security which pays one monetary unit if the event i is realized, and $0 \leq p_i \leq 1$ for all i. The solution of the problem is:

$$p_i = \sqrt{[z_i f_i + (1-z)]f^2_i}\{\Sigma\sqrt{[z_s f_s + (1-z)f^2_s]}\}.$$

If there are no insiders *(z=0)*, the result of this model is identical to that of the CERM: the prices coincide with the true probabilities ($p_i=f_i$), as the proportion of the total wealth wagered on each event ($B_i/\Sigma B_i$) is equal to its probability (f_i). But if there is a positive chance of insider trading ($z\geq 0$), and if this chance relatively decreases with the probability of the event itself ($z_i/f_i<z_j/f_j$ *if* $f_i>f_j$), the bookmaker maximizes its profits by setting the prices so that: $p_i/p_j < f_i/f_j$, whenever $f_i>f_j$. Figure 3 (from Shin, 1991, fig.1), shows the quadratic rule used by bookmakers to set the prices (odds) in front of an expected percentage z of informed traders. If bookmakers expect insider trading to be more prevalent given that a long shot is tipped to win, the betting odds should understate the winning chances of a favorite relatively less than the winning chances of a long shot. In terms of the bookmaker's profit, this implies that the ex-post take out rate should be higher for long shots than for favorites. The F/L bias may therefore be a supply side (bookmakers) optimal pricing response to an adverse selection problem: unknown insiders among the bettors. Shin (1993) and Williams and Paton (1997) claim to have found empirical evidence of insider trading in horse racing supporting Shin's model. In particular, the percentage of unknown insiders (Sin's z) in the bettors' pool has been estimated to be small but significant: about 2% by Shin (1993) and Vaughan Williams and Paton (1997). Other evidence of insider trading was found by Schnytzer and Shilony (1995) in the Australian horse betting market.

3. Evidence of match rigging in Italian football: "Calciopoli".

Italian football is used to match fixing. As early as 1927 the Italian Football Federation revoked the championship won by Torino since its managers bribed a Juventus player before the Turin derby. In the eighties, two famous cases of insider trading were uncovered: the so-called "Calcioscommesse" (1980), and "Calcioscommesse 2" (1986) cases. The "Calcioscommesse" cases involved mainly players, who gained their monetary profit from betting in the black market[ix], but the last case of

match-fixing, "Calciopoli" (2006), was bigger than any other one before. In the "Calciopoli" case referees, federation officials, owners and team managers cooperate in a kind of criminal organization to alter the final ranking of the Italian First League ("Serie A") tournaments in favor of some teams (mainly in favor of the Juventus team, whose managers ran the rigging system). In may 2006 the scandal was uncovered by Italian prosecutors after tapping phone conversations in relation to an investigation on the use of doping at Juventus team. The scandal, commonly referred to as "Calciopoli", resolve around twenty months of wiretapped conversations involving key figures of Italian Football. The prosecutors found that the general manager of Juventus, Luciano Moggi, had exerted pressure on referees, officials of the football federation and journalists, ahead of crucial matches involving Juventus or rival teams. These contacts were finalized to rig games by choosing referees favorable to Juventus and manipulating news on television and newspapers against the referees not displaying a favorable attitude toward the team of Moggi[x]. The matches that were likely to be rigged did not only involve Juventus, but were mostly in favor of Juventus, as they condition the outcomes of other matches in favor of Juventus. The other teams involved in the scandal were Milan, Fiorentina, Lazio and Reggina. Although there is no pending judicial inquiry on match rigging before 2004, there are indications that match fixing based on corruption of referees was present at least since Luciano Moggi became general manager of Juventus in 1994[xi], and actually, well before the uncovering of the Calciopoli evidence, rumors of rigging were already widespread among Italian football fans (Garlando, 2005). The verdict of the sport justice on Calciopoli was the following: even if the matches of the 2005-06 season were not under investigation, the" scudetto" won by Juventus was removed and awarded instead to Inter, Juventus was also relegated to play in the Second Division ("serie B") with a deduction of 9 points in the 2006/07 season; Milan was kept in "serie A", but penalized by 8 points; Fiorentina was banned from the Champions League and was penalized with a deduction of 15 points; Lazio was sanctioned with a reduction of 3 points and the exclusion from the UEFA cup; finally Reggina was sanctioned with a deduction of 15 points. Moreover, the investigation on the "Calciopoli" case sanctioned the

top management of the Italian Federation of Football (FIGC) and of the Italian Association of Referees (AIA).

4. EMPIRICAL ANALYSIS.

4.1 Data and descriptive statistics.

In a football game a bettor could bet on three alternative results: victory of the home team (H), victory of the visiting team (V), and a draw (D). The odd ($Qi :1$) is the amount of money that a bookmaker will return for a unit bet if the event i ($i=H,V,D$) is realized. We start our analysis of the Italian football wagering market analyzing a panel dataset which consists of the odds posted by three bookmakers: MatchPoint, MisterToto and SportingBet (henceforth: MP, MT and SB), on 289 games played in the 2007/08 "Serie A" season[xii]. By comparing the odds offered by the three bookmakers, we can see that their distributions are quite similar in their first two moments (tab. 1). Actually, the t-tests cannot reject the hypotheses: $Q_{D,LM} > Q_{D,MT} > Q_{D,SB}$ and $Q_{H,LM} > Q_{H,MT}$. (tab. 2), but the distributions are homoskedastic, and the paired linear correlation coefficients are very close to unity (tab. 3). We looked for systematic differences in the odds by regressing the odds of one firm on the odds of another (tab. 4). Under the null hypothesis of no systematic differences between the firms, the intercept should be indistinguishable from zero and the slope coefficient should equal unity. The results of our OLS regressions cannot reject the hypothesis $H_0: \beta \approx 1$ for any regression, but they rejects the hypothesis $H_0: \alpha \approx 0$ in six cases on nine (mostly because MT's odds are lower than the others), and all the determination coefficients of the nine regressions are very high, close to unity. Summing up, our results suggest that the odds offered by the three firms are highly correlated, although some firm systematically offer odds which are (marginally, but often significantly) lower/higher than the others. Given these similarities, we summarize the market odd distribution by the distribution of a synthetic odd (named: "Delphi"), which is the un-weighted average of the odds offered by the three bookmakers. The frequency distributions of the "Delphi" odds are quite normal for the victory of the home/visiting team event (but the last one is

asymmetric), but the frequency distribution for the "draw" odds is strongly leptokurtic and asymmetric[xiii].

4.2 The correlation between the game results and the associated market odds.

The bookmaker ex-post profit is: $\pi=\sum B_i-\sum(f_i \cdot B_i \cdot Q_i)$, where B_i is the amount bet on the event i, and f_i its frequency. Therefore, given f_i and $\sum B_i$, the iso-profit curves of the bookmaker map a family of hyperboles in the Cartesian plane described by the odd of the event (Q_i), and the share of bets placed on it (B_i). In this way the odds on a sport event are like asset prices: they are both market equilibrium clearing prices and market forecasts of actual game outcomes. The simplest model to study a betting market is the "Constant Expected Return Model" reviewed by Sauer (1998). According to this model, where betting is a fair game played by rational risk-neutral representative agents, the expected return on any unit bet should be 1, and, as a corollary, the share of the pool which is bet on the event i ($B_i/\sum B_i$) should be the same as the event i probability (f_i). As an example, in a fair football game where the events were equally likely ($f_H=f_D=f_V=1/3$), the bets should be $B_H=B_D=B_V=1/3\sum B_i$, and the odds should be $Q_H=Q_D=Q_V=3:1$. If the CERM is true[xiv], each odd should be the inverse of the frequency of the associated event, so the hypothesis $H_0: Qi \approx 1/fi$ can be empirically tested. In order to measure this kind of predictive efficiency of the Italian soccer wagering market we used a Multinomial Logit analysis. We ran three Multinomial Logit regressions, one for each bookmaker, where we regressed the vector of games results (y_i) on the matrix of the associated odds ($Q_{i,j}$) offered by each bookmaker: $y_i = f(Q_{Hi}, Q_{Di}, Q_{Vi})$ where y= H,D,V and i=1, …, 289. We found that, for each bookmaker, the odds have a significant and very strong predictive power (tab. 5). All the preudo-R^2 are very high and all the F^s are highly significant, the matching between the predicted and actual results is very high: for each of the three bookmakers, the multinomial logit models correctly predicted about 80-90% of the results (tab. 6). Moreover, if we regress the game results on the un-weighted average of the associated odds offered by the three

bookmakers (the "Delphi" odds), we get an almost perfect prediction of the actual game results. The predictive efficiency of the market is also confirmed by an OLS analysis. As in Kuypers (2000), for each result (Home win, Visitors win and Draw) we ordered the odds (Q_{Hi}, Q_{Di}, Q_{Vi}) and grouped them into 23/25 categories, then the average implied probabilities of each category (q) were used as explanatory variables in the following OLS regression: $y = a + bq + \varepsilon$. According to the results showed in table 7, the hypothesis that the implied probabilities are good predictors of the actual game result (H_0: b=1) cannot be rejected at the 95% confidence level[xv].

4.3 The return distribution and the favorite long shot bias.

By comparing the odds with the games results (tab. 8), we measured an ex-post take-out rate of about 10%, it was mainly coming from the bets placed on the draw event (π_D=5%) and on the victory of the visiting team (π_V=4%). The "favorite-long shot bias" means that the bookmaker returns on favorites are less than those on long-shots H_0: $\pi(f_j) > \pi(f_i)$, if $f_i > f_j$. In order to found evidence of this bias, we identify the favorite/long shot events in our dataset and compared their associated ex-post take out rate. The identification procedure started by ordering the events on the basis of their (odd-implied) expected probability $q_{ij} = 1/Q_{ij}$, where: i =H,D,V, and j=MT,MP,SB. Then, we split the q_{ij} frequency distribution in the subsets: Low_{ij}, $Inside_{ij}$, $High_{ij}$. The low/high cut-off value was the average of q_{ij} distribution less/more the distribution standard deviation. This partition classifies the "favorite/long shot" events in the $High_{ij}/Low_{ij}$ subsets, which consist of the events whose expected probability is higher/lower. By construction, the subset $Low_{ij}/High_{ij}$ is the subset consisting of the events associated with the higher/lower odds, it is the lower/higher tail of the q_{ij} distribution and it consists of about 1/6 of the whole distribution; the subset $Inside_{ij}$ is the central body of the distribution and it consists of about 2/3 of the distribution itself. By comparing the odds with the associated results we have an ex-post take out distribution for each subset (tab. 9). The hypothesis of favorite long shot bias is H_0: $\pi(Low_{ij}) > \pi(High_{ij})$, and our results support this hypothesis insofar $\pi(Low_{ij})$ was actually bigger than $\pi(High_{ij})$ for any i and j. On average, the take

out rate was 6% for long shots and 1% for favorites. Specifically, $\pi(High_{ij})$ was negative for i=H,D, while $3\%<\pi(Low_i)<8\%$. This result is confirmed by the analysis of another dataset, consisting in the results and the odds offered by one bookmaker (MP) on 6369 games played from the 2002/03 season until the 2007/08 season. For this dataset too: $\pi(Low_i)>\pi(High_i)$ for any i, On average, the take out rate was 6% for long shots and -1% for favorites[xvi]. Specifically $\pi(High_i)<0$ for i=H,D, while $4\%<\pi(Low_i)<8\%$. Summing up, the analyses of both datasets support the hypothesis that the F/L bias is an empirical feature of the Italian football wagering market.

4.4 Match rigging and the F/L bias.

The Calciopoli case is a kind of natural experiment where we can see if the rumors of rigging (which proved to be true) influenced the wagering market: a) there is a judicial evidence that some matches were rigged; b) the court found this evidence sufficient to prove that some managers manipulated the Italian "Serie A" tournaments until 2006; c) the verdict banned those manager from the Italian football henceforth, so that rigging was not on the run after 2006 (or, at least, it was not ran by the same people). If we assume that the bookmakers were able to detect the rigged matches, so that they can adjust their odds according to the expected of percentage of insiders, the first test may consist in comparing the bias displayed by the odds posted on suspected matches with the odds posted on the rest of the dataset. If the expectation of insider trading was higher about these games than the rest then, according to Shin's model, the bias should be larger for the 80 suspected games than for the other 340 ones[xvii]. Our analysis of the 2004/05 Serie A season dataset found the opposite result: the F/L bias was larger for the rest of the dataset ($\pi(Low) = 5\% > \pi(High) =0\%$) than for the suspected games ($\pi(Low) = 4\% > \pi(High) =2\%$). There may be many explanations for this result. First, the verdict states that rigging was finalized to alter the matches results in order to promote some teams (mainly the Juventus one) in the "Serie A" ranking, it does not state that rigging was finalized to yield any monetary profit from betting on the rigged matches. Therefore, it could be that rigging did not influence the wagering market simply because the informed people

involved in Calciopoli did not bet. Another explanation for our result may come from the inferential procedure which was followed by the court: the investigators did not look for every rigging episode, but the evidence of rigging in some cases was sufficient to prove that rigging was a systematic mis-behavior of the Juventus management. The verdict of the Calciopoli case confirmed that rigging was a systematic malpractice of Juventus' managers since Moggi became its general manager in 1994, and because of this the court banned him and the other main figures of the scandal off the Italian football from 2006 up to 5 years. Therefore, we split the whole dataset into two subsets, the first one consists of the results and the odds offered by the MP bookmaker from the 2002/03 season until the 2005/06 season, that is when rigging was on the run, while the second dataset consists of the already reviewed (see above par. 4.1) results and odds posted by three bookmakers for the 2007/08 seasons, that is after the Calciopoli verdict. For each subset we applied the same procedure than before to select the favorite/long shot subsets. Comparing the F/L bias realized before and after Calciopoli (showed in tab.10), we can see the F/L bias was slightly higher (1,4%) before Calciopoli than afterward. Moreover, if we analyze the subset consisting only in the matches played by the most suspected team (Juventus) before Calciopoli, we can see that the F/L bias on these games (14%) was the double than that realized on the whole dataset (7%)[xviii]. Summing up, if we expand our analysis beyond the judicial cases, we can see that the F/L bias was slightly larger before the Calciopoli case than later on, and also that, when rigging was on the run, the F/L bias was much larger than the average for those games involving the most suspected team (Juventus).

5. CONCLUSIONS.

In this paper we investigated on the correlation between the results and the odds of the football games played in the highest Italian league ("serie A"), which was plagued by the notorious "Calciopoli" scandal, a case of systematic match rigging, which was run by top federation and team managers until 2006. Preliminarily, we tested if the market odds did predict the football game results by using a Multinomial Logit analysis and Ordinary Least Square procedure. Our results

confirm that the odds did have a very significant predictive power. We then look for evidence of the "favorite-long shot bias", an empirical feature where the bookmakers' returns on favorites are less than those on long-shots. We found that the favorite long shot bias is an empirical feature of the Italian football wagering market. There may be many demand-side reasons which could add up together to explain the F/L bias, but the F/L bias may also be the equilibrium result of a game where bookmakers expect to deal with unknown insiders (Shin, 1991). Therefore, we took "Calciopoli" as a kind of natural experiment, useful to see if the rumors of match rigging were associated with a larger F/L bias. Actually, well before the uncovering of the Calciopoli evidence, the rumors about a systematic referees' bias in favor of some team (Juventus in particular) were already widespread. It is plausible that these rumors were known to bookmakers too, who could have replied to these rumors by altering the odds in order to minimize the losses they would incur if they had the chance to deal with insiders. We found evidence that, while rigging was on the run, the F/L bias was much larger on the matches of the most involved team (Juventus), and that the F/L bias was slightly larger before the eruption of the scandal than after the verdict. Summing up, we found that the Italian football betting market was weakly efficient, but also that the odds displayed a persistent F/L bias. This bias was larger during "Calciopoli", but it did not vanish after that. Therefore, we guess that the F/L bias could be the sum of several demand-side factors (bettors' local risk love or bettors' behavioral attitudes), but that this bias may also be originated by the strategic behavior of bookmakers who are expecting to deal with unknown insiders.

References.

AgiProNews, Agenzia di stampa giochi a pronostico e scommesse, Roma (http://www.agipronews.it/).

M.M. Alì, *Probability and Utility estimates for Racetrack Bettors*, Journal of Political Economy, 1977, 85:4, 803-15.

P. Asch, B.G. Malkiel, and R.E. Quandt, *Racetrack Betting and Informed Behavior*, Journal of Financial Economics, 1982, 10, 2, 187-94.

P. Asch, B.G. Malkiel, and R.E. Quandt, 1984, *Market Efficiency in Racetrack Betting*, Journal of Business, 57, 165-75.

W. Bagehot, *The Only Game in Town*, Financial Analysts' Journal, 1971, 22, 12-14.

Boeri and Severgnini, 2010, *Match Rigging and the career concerns of referees*, Labour Economics, 2010, 18, 3, 233-54.

K. Busche, and C. Hall, *An Exception to the Risk Preference Anomaly*, Journal of Business, 1988, 61, 3, 337-46.

K. Busche, *Efficient Market Results in an Asian Setting*, in D.B. Hausch, V.S.Y. Lo and W.T. Ziemba (eds.) *Efficiency of Racetrack Betting Markets*, Academic Press, San Diego, 1994 .

C. Camerer, 1989, Does the Baseball Market Believe in the Hot Hand?, American Economic Review, 79, 5 , 1257-61.

T. Copeland, and D. Galai, *Information Effects on the Bid Ask Spread*, Journal of Finance, 1983, 38, 1457-69.

J. Conlinsk, 1993, *The Utility of Gambling*, Journal of Risk and Uncertainty, 6,3 , 255-75.

N.F.R. Crafts, *Some Evidence of Insider Trading Knowledge in Horse Race Betting in Britain*,Economica, 1985, 52, 295-304.

J. Dowie, *On the Efficiency and Equity of Betting Markets*, Economica, 1976, 3, 139-50.

G. Dwyer et al., 1993, *Tests of Rational Expectations in a Stark Setting*, Economic Journal, 103, 586-601.

E.L. Figgis, *Focus on Gambling*, London, Barker 1951.

S. Figlewsky, *Subjective Information and Market Efficiency in a Betting Market*, 1979, Journal of Political Economy, 87:1, 75-89.

M. Friedman and L. Savage, *The Utility Analysis of Choices Involving Risk*, 1948, The Journal of Political Economy, 56, 4, 279-304.

P. Gabriel and J. Marsden, *An Examination of Market Efficiency in British Racetrack Betting*, 1990, The Journal of Political Economy, 98, 4, 874-85.

P. Gabriel and J. Marsden, *An Examination of Market Efficiency in British Racetrack Betting: Errata and Corrections,* 1991, Journal of Political Economy, 99, 3, 654-59.

L. Garlando, *Nostra Signora del Dischetto*, 2005, Kowalsy, Milano.

T. Gilovich, *Biased Evaluations and Persistence in Gambling*, 1983, Journal of Personality and Psychology, 44, 6, 1110-26.

T. Gilovich, and C. Douglas, *Biased evaluations of randomly-determined gambling outcomes*, 1986, Journal of Experimental Social Psychology, 22, 228-41.

L. Glosten and P. Milgrom , *Bid Ask and transaction Prices in a Specialist Market with Heterogeneous Informed Traders*, 1985, Journal of Financial Economics, 14, 71-100.

R. Griffith, *Odds Adjustments by American Horse Race Bettors*, American Journal of Psychology, 1949, 62, 290-94.

R.J. Henery, *On the average probability of losing bets on horses with given starting price odds*, Journal of the Royal Statistic Society,1985, 148:4, 342-9.

A.E. Hoerl, and H.K. Fallin, *Reliability of Subjective Evaluations in a High Incentive Situation,* Journal of the Royal Statistic Association, 1974, 137, 227-30.

W. Hurley, and L. McDonough, *A Note on the Hayek Hypothesis and the Favorite Long Shot Bias in Parimutuel Betting,* American Economic Review, 1995. 85:4, 949-55.

D. Kanheman and A. Tversky, *Prospect Theory: An Analysis of Decision under Risk*, 1979, Econometrica, 47, 2, 263-92.

D. Kanheman and A. Tversky, 1991, *Loss Aversion in Riskless Choice: A Reference –Dependent Model*, Quarterly Journal of Economics, 106, 1039-61.

Kyle, *Continuous Auctions and Insider Trading*, Econometrica, 1985, 63, 1315-35.

R.L. Losey, and J.C. Talbott Jr., *Back on the Track with the Efficient Market Hypothesis*, Journal of Finance, 1980, 35:4, 1038-43.

M. Machina, *Expected Utility Analysis Without the Independence Axiom*, Econometrica, 1982, 50, 277-323.

M. Machina, 1987, Choices Under Uncertainty, Problems Solved and unsolved, Journal of Economic Perspectives, 1, 1, 121-54.

H. Markowitz, The Utility of Wealth, 1952, *Journal of Political Economy*, 56, 4, 151-8.

W.H. McGlothin, 1956, Stability of Choices Among Uncertain Alternatives", *American Journal of Psychiatry*, 69, 604-15.

R. Piron, and L.R. Smith, *Testing risk love in an experimental racetrack*, Journal of Economic Behavior and Organization, 1995, 27, 465-74.

P. Pope, and D.A. Peel, *Information, Prices, and Efficiency in a Fixed-Odds Betting Market*, *Economica*, 1989, 56, 322-41.

M.G. Preston, and P. Baratta, *An experimental study of the auction-value of an uncertain outcome*, American Journal of Psychology, 1948, 61, 183-93.

R. Quandt, *Betting and Equilibrium*, 1986, The Quarterly Journal of Economics, 101, 1, 201-08.

R.N. Rossett, *Weak experimental verification of the expected utility hypothesis*", Review of Economic Studies, 1971, 38, 481-92.

Royal Commission on Gambling, *Final report*, *1*, London, HMSO, 1978.

P. Samuelson, *Probability, Utility, and the Independence Axiom*, 1952, Econometrica,20, 4, 670-8.

R.D. Sauer, *The Economics of Wagering Markets*, Journal of Economic Literature, 1998, 36, 2021-2064.

Schnytzer, and Y. Shilony, *Inside Information in a Betting Market*, 1995, Economic Journal, 105, 963-71.

Schnytzer, and G. Weinberg, *Testing for Home Team and Favorite Biases in the Australian Rules Football Fixed-Odds and Point Spread Betting Markets*, Journal of Sports Economics, 2008, 9:2, 173-90.

H.S. Shin, 1991, *Optimal Odds Against Insider Traders*, The Economic Journal, 101, 1179-85.

H.S. Shin, *Prices of State Contingent Claims with Insider Traders and the Favorite-Longshot Bias*, 1992, The Economic Journal, 102, 426-35.

H.S. Shin, 1993, *Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims*, The Economic Journal, 103, 1141-53.

W. Snyder, *Horse Racing: Testing the Efficient Market Models*, Journal of Finance, 1978, 33, 1109-18.

P. Slovic, B. Fishoff and S. Lichtenstein, 1982, *Facts versus Fears: Understanding Perceived Risk*, in *Judgment Under Uncertainty: Heuristic and Biases*, , D. Kahneman, P. Slovic and A. Tversky eds., Cambridge: Cambridge University Press.

E. Snowberg and J. Wolfers, 2007, *The Favorite-Longshot Bias: Understanding a Market Anomaly*, in *Efficiency of Sports and Lottery markets*, D. Haush and W. Ziemba (eds.), Elsevier, Handbook in Finance seires.

E. Snowberg and J. Wolfers, *Explaining the Favorite-Longshot Bias: Is It Risk-Love or Misperceptions?*, 2010, NBER WP # 15923.

S. Swindler and R. Shaw, *Racetrack wagering and the uninformed bettor: a study of market efficiency*, Quarterly Review of Economics and Finance, 1995, 35:3, 305-14.

Tertullianus (160-220 d.c.) *De Spectaculis* (191-202 d.c.), Cambridge, Massachusettes, Harvard University Press, 1997.

R.H. Thaler, and W.T. Ziemba, *Anomalies-Parimutuel Betting Markets: Racetrack and Lotteries*, Journal of Economic Perspectives, 1988, 2, 161-74.

W.L. Vaughan Williams, and D. Paton, *Risk, return, and adverse selection: a study of optimal behavior under asymmetric information*, Rivista di Politica Economica, 1996, 11:12, 63-81.

W.L. Vaughan Williams, and D. Paton, *Why is There a Favorite Long Shot Bias in British Racetrack Betting Markets*, Economic Journal, 1997, 107, 150-58.

K. Viscusi and C. O'Connor, 1984, *Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?*, American Economic Review, 942-56.

M. Weitzmann, *Utility Analysis and Group Behavior: An Empirical Study*, 1965, The Journal of Political Economy, 73, 1, 18-26.

L.M. Woodland, and B.M. Woodland, *Market Efficiency and the Favorite Long Shot Bias: The Baseball Betting Market*, Journal of Finance, 1994, 49:1, 269-80.

L.M. Woodland, and B.M. Woodland, *Market efficiency and profitable wagering in the national hockey league: Can bettors score on long shots?*, Southern Economic Journal, 2001, 67, 983-995.

E. Yaari, *Convexity in the Theory of Choice Under Risk*, Quarterly Journal of Economics, 1965, 79, 278-90.

Tables.

Table 1. Odds: descriptive statistics.

Bookmaker	Home			Draw			Visitor		
	MT	MP	SB	MT	MP	SB	MT	MP	SB
Average	2.44	2.48	2.45	3.27	3.29	3.20	4.56	4.59	4.53
st.dev.	*1.23*	*1.29*	*1.25*	*0.69*	*0.75*	*0.69*	*2.94*	*2.89*	*3.02*

Italian "serie A", season 2007-2008, n = 289.

Table 2. Odds: paired t-tests.

T-test n =289	Home			Draw			Visitor		
	MP-MT	SB-MT	MP-SB	MP -MT	MT-SB	MP -SB	MP -MT	MT-SB	MP -SB
Average	0.05	0.02	0.03	0.03	0.06	0.09	0.03	0.03	0.06
t value	*3.27**	*1.14*	*3.04**	*2.49**	*3.99**	*5.97**	*0.65*	*0.44*	*1.20*

Italian "serie A", season 2007-2008, n = 289. Home: odd posted on the victory of the home team; Draw: odd posted on the draw event; Visitor: odd posted on the victory of the visiting team.

Table 3. Odds: linear correlations.

Correlation	MT/ MP	MT/SB	MP /SB
Home	0.98	0.98	0.99
Draw	0.97	0.93	0.94
Visitor	0.96	0.92	0.96

Italian "serie A", season 2007-2008, n = 289. Home: odd posted on the victory of the home team; Draw: odd posted on the draw event; Visitor: odd posted on the victory of the visiting team.

Table 4. Odds: OLS analysis.

Odd	Dep. Var.	Regressor	Alpha	Beta	R squared	F
Home	MT	MP	0.11 (3.99)*	0.94 (91.47)	0.97	8367
Home	MT	SB	0.07 (2.60)*	0.96 (93.16)	0.97	8678
Home	M	SB	-0.02 (-0.88)	1.02 (127.5)	0.98	16260
Draw	MT	MP	0.34 (7.53)*	0.89 (66.67)	0.94	4445
Draw	MT	SB	0.28 (3.90)*	0.93 (42.94)	0.87	1844
Draw	MP	SB	-0.01 (-0.16)	1.03 (48.03)	0.89	2308
Visitor	MT	SB	0.48 (4.02)*	0.90 (41.15)	0.86	1693
Visitor	MT	MP	0.07 (0.77)	0.98 (61.11)	0.96	3735
Visitor	MP	SB	0.42 (4.86)*	0.92 (57.83)	0.92	3344

Italian "serie A", season 2007-2008, n = 289x3. Home: odd posted on the victory of the home team; Draw: odd posted on the draw event; Visitor: odd posted on the victory of the visiting team.

Table 5. Odds and game results: Multinomial Logit analysis.

Provider	Pseudo-R^2	-2LL	(-2LL) Home	(-2LL) Draw	(-2LL) Visitor
MP	0,87 (N)	415,61	246,81	182,72	278,41
	0,68 (MF)	(0,01)	(0,20)	(0,69)	(0,09)
SB	0,91 (N)	473,54	111,43	42,79	158,67
	0,75 (MF)	(0,00)	(0,24)	(0,52)	(0,09)
MT	0,92 (N)	460,37	134,74	34,55	183,33
	0,78 (MF)	(0,00)	(0,06)	(0,95)	(0,02)

Italian "serie A", season 2007-2008, n = 289. Home: victory of the home team; Draw: draw result; Visitor: victory of the visiting team.

Table 6. Odds and game results: classification table.

	% correct			
Game results	MP	SB	MT	Delphi
Home	87,6	89,8	89,8	100,0
Draw	80,8	78,2	84,6	98,7
Visitor	82,4	78,4	81,1	97,3
% tot.	84,4	83,7	86,2	99,0

Italian "serie A", season 2007-2008, n = 289x3. Home: victory of the home team; Draw: draw result; Visitor: victory of the visiting team.

Table 7: OLS regression of actual results on odds' implied probabilities.

Game results	constant	b	R^2	F	n
Home	-0.02	1.03	0.72	58.22	25
	(0.07)	(0.14)			
Draw	-0.04	1.00	0.67	44.28	24
	(0.05)	(0.15)			
Visitor	-0.09	1.16	0.32	9.73	23
	(0.12)	(0.37)			

Standard deviations of estimated coefficients are showed under parentheses.

Table 8. Bookmaker take-out: descriptive statistics.

	Home	Draw	Visitor	Total
Frequency	137	78	74	289
Frequency (%)	0.47	0.27	0.26	1.00
Implied prob.	0.48	0.32	0.30	1.10
Take out rate	0.01	0.05	0.04	0.10

Italian "serie A", season 2007-2008, n = 289. Home: victory of the home team; Draw: draw result; Visitor: victory of the visiting team.
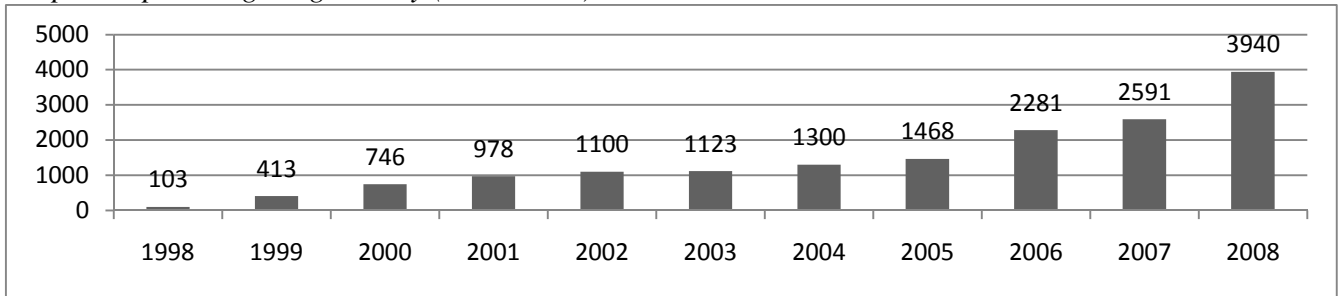
Table 9. Takeout distribution for sub-groups.

		Take	out	rate
Event	cluster	MT	MP	SB
Home	Low probability	0,04	0,03	0,06
Home	Inside probability	0,01	0,01	0,00
Home	High probability	0,00	-0,01	-0,01
Draw	Low probability	0,08	0,06	0,07
Draw	Inside probability	0,04	0,05	0,06
Draw	High probability	-0,02	-0,08	-0,03
Visitor	Low probability	0,06	0,06	0,08
Visitor	Inside probability	0,04	0,03	0,04
Visitor	High probability	0,04	0,06	0,01

Italian "serie A", season 2007-2008, n = 289x3. Home: victory of the home team; Draw: draw result; Visitor: victory of the visiting team.

Table 10: Take out rates before and after Calciopoli.

	Before Calciopoli		After Calciopoli	
	Favorites	Long Shots	Favorites	Long Shots
All dataset	-0,77% *(538)*	5,97% *(650)*	0,56% *(332)*	5,91% *(413)*
Juventus only	-4,28 *(107)*	9,67% *(152)*	n.a.	n.a.

BC :Before Calciopoli dataset (2002-06 seasons, odds posted by one bookmaker); AC: After Calciopoli dataset (season 2007-08, odds posted by three bookmakers); number of observations on parentheses.

Figures.

Graph 1: Sport wagering in Italy (1998-2008).

Figure 1: Prospect Theory.
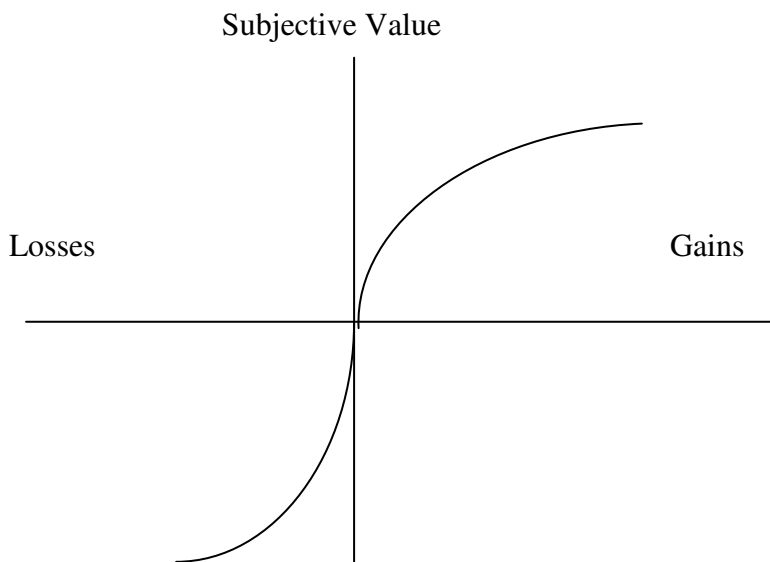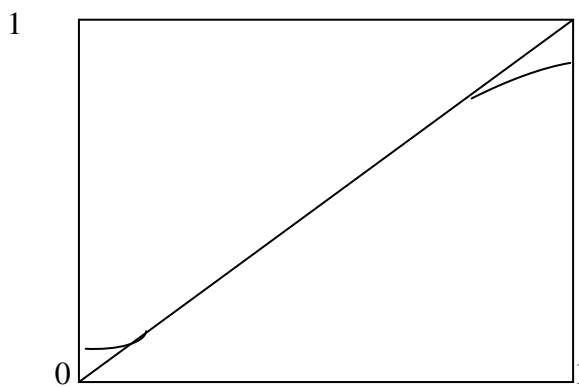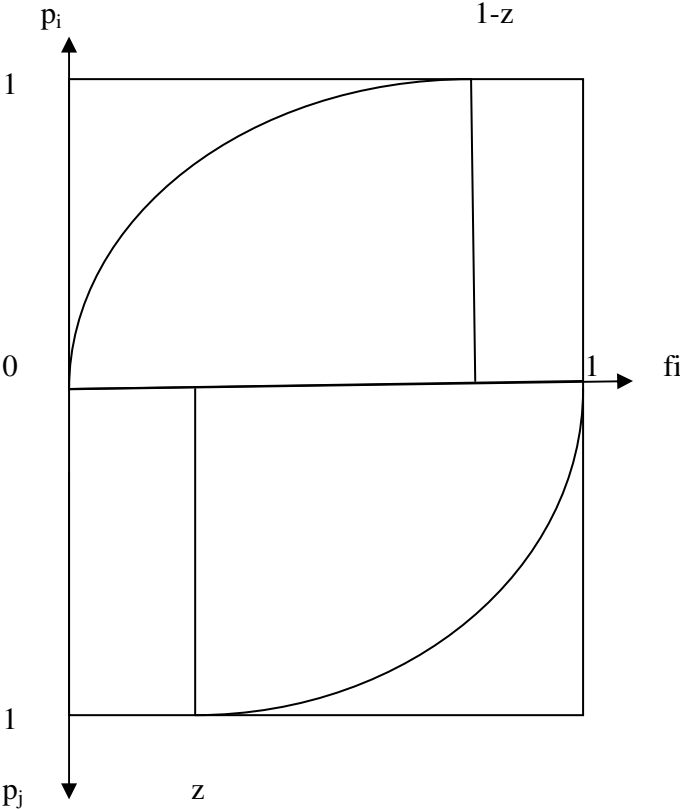

Subjective Value

Losses

Gains

Figure 2: Subjective probabilities misperceptions.

Subjective probability (s)

Implied probability (p=1/Q)

Figure 3: Optimal prices (p) in front of a z percentage of insiders.

Endnotes.

[i] "*Aspice populum ad id spectaculum iam venientem, iam tumultuosum, iam caecum, iam de sponsionibus concitatum*", Tertullianus (AD 155-230), p.271-2.

[ii] Before that time, authorized betting on soccer events was limited to only one lottery ("Totocalcio"), and illicit bets were posted on the black market. The subsequent increase in the volume of wagering was also due to the increase in the varieties of available lotteries and to the diversion of trading from the black market.

[iii] Bagehot (1971), Copeland and Galai (1983), Glosten and Milgrom (1985) and Kyle (1985).

[iv] For a review you can see Sauer (1998).

[v] You can still find on the web (http://www.youtube.com/watch?v=PQDFux6q-BQ) a milestone in the history of Italian soccer rumors: a denied penalty kick in favor of Inter in a match against Juventus which was decisive to assign the 1998 "Scudetto".

[vi] The bookmakers return π is positive if $\sum p_i > 1 = \sum f_j$, where p_i is the probability implied by the odd Q_i ($p_i=1/Q_i$) and f_j is the probability of the event i.

[vii] Another explanation, alternative to the assumption of a risk-loving attitude of bettors, may step from Samuelson (1952), who argued that gambling is not only wealth-oriented: "*Warning: what constitutes a prize is a tricky concept. When I go to a casino, I go not alone for the dollar prizes but also for the pleasure of gaming*" (p. 671). Conlisk (1993) included this "pleasure for gambling" in the following preference function: $E(G,p,W) = pU(W+G) + (1-p)U(W-L) + \varepsilon V(G,p)$; where the amount G is won with probability p, the amount L is lost with probability $1-p$, and W is the initial wealth. In this model an additional utility of gambling [$\varepsilon V(G,p)$] is added to a utility of wealth function [$U(W)$]. The function $U(W)$ is the standard utility of wealth, which is bounded and exhibits decreasing absolute risk aversion: $U(0) = 0; U' > 0; U'' < 0$. The additional utility of gambling is: $\varepsilon V(G,p)$, where ε is a non negative scale parameter, and $V(G,p)$ has the properties: $V(0,p)=V(G,0)=0; V_1(G,p)>0; V_2(G,p)> 0; V_{11}(G,p)< 0$, for $G>0$. Assuming that ε is sufficiently small, this model predicts the acceptance of small gambles and the purchase of insurance when risks are large. Intuitively, the basis for these implications is that for small gambles, the utility-of-gambling is first-order small, whereas the risk aversion effect is second order small.

[viii] "Analysis suggests that a person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwise" (Kanheman and Tversky, 1979, p. 269). Figure 2 gives a a graphical illustration of the "Prospect Theory".

[ix] In 1982 Milan and Lazio were relegated to the Second Division ("Serie B") after fixing a match and some of their players were found guilty of illegal gambling on soccer games.

[x] Boeri e Severgnini (2010) found that past involvement in match rigging increased the likelihood that referees were assigned to the most important matches in the tournament. Significantly, this increase occurred precisely for those referees who were candidates for promotion to an international standing, a crucial step in their career: "Our results indicate that career concerns may be a substitute for financial bribes, reducing substantially the monetary outlays involved in match fixing" (pag.10 in press, 2010).

[xi] Boeri et al (2010, tab.1) listed the most dubious referee decisions favorable to Juventus in the year 1994-2004. We also remember that when Moggi was the manager of Torino (before 1994), he was found offering prostitutes to referees in exchange for preferential treatment in favor of its team.

[xii] Because of missing data, 71 games were dropped from the pool.

[xiii] "This behavior could simply reflect a general inability to predict draw outcomes with any degree of reliability, in which case the unconditional (constant) probability might be the most appropriate basis for setting the odds", Pope and. Peel (1989,p.328).

[xiv] Griffith (1949), Hoerl and Fallin (1974) found a congruence between the win pool share and the frequency of event in Us pari-mutuel horse races.

[xv] Actually the predictive efficiency of the odds about the draw event is not as good as it is about the victory of the home/visiting team events. About this result we again quote the Pope and Peel's comment cited in the endnote xiii.

[xvi] Dowie (1976) found the same pattern of returns in the British horse races bookmaking market: his figures indicate that bookmakers lost money when taking bets on extreme favorites.

[xvii] We draw the list of suspected matches from Boeri and Severgnigni (2010, Appendix A, Annex, Tables 9 and 10).

[xviii] Our available dataset is too short to make any significant inference about the F/L bias displayed by the odds associated with the matches played by the Juventus team in Serie A after Calciopoli.