International Journal of Sport Finance, 2011, **6**, 265-280, © 2011 West Virginia University

Match Rigging and the Favorite Long-Shot Bias in the Italian Football Betting Market

Marco Rossi

University La Sapienza, Rome

Daniel Larson is a lecturer in the Department of Health, Physical Education, and Sport Science in the Wellstar College of Health and Human Services. His research interests include industrial organization, sport economics, and participant-event management and marketing.

Abstract

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

Keywords: football wagering, match rigging, favorite long shot bias

Introduction

Wagering on sport events is a very old Italian tradition: In Roman imperial times, the Circus Maximus in Rome drew crowds of more than 250,000 people who were anxious about their bets on, for the most part, chariot races.[1] 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 amounts of almost €4 billion were spent in Italy on sport betting in 2008 (see Figure 1).[2] In this study, I examined the Italian soccer betting market, which is by far the most important sector of the Italian sport betting market: 93% of the sport bets placed in 2008 were on soccer events.

My data comprised the results and the odds of soccer games played in the highest Italian soccer league: the Serie A championship. The first dataset consisted of the results and odds posted by one bookkeeper on 6,369 games played from the 2002-2003 season until the 2007-2008 season. The second dataset consisted of the results and odds offered by three bookkeepers on 289 games played in the 2007-2008 season.

Football bets are simple financial assets: They have a short and Ill-defined end point, usually a week, when their value becomes certain. Also, there is no secondary market for bets. These factors avoid bubbles in the football betting market, simplifying the pricing problem. I started my analysis by assuming that the betting market is

Figure 1: Sport Wagering in Italy (1998-2008) Data source: Agipronews.

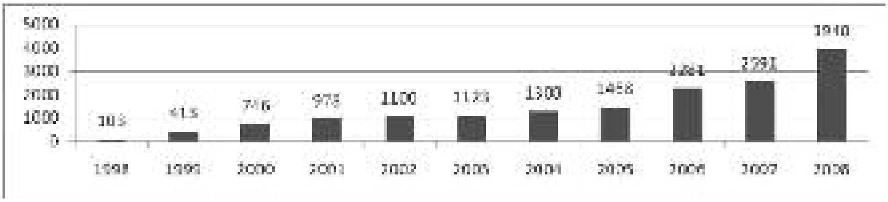
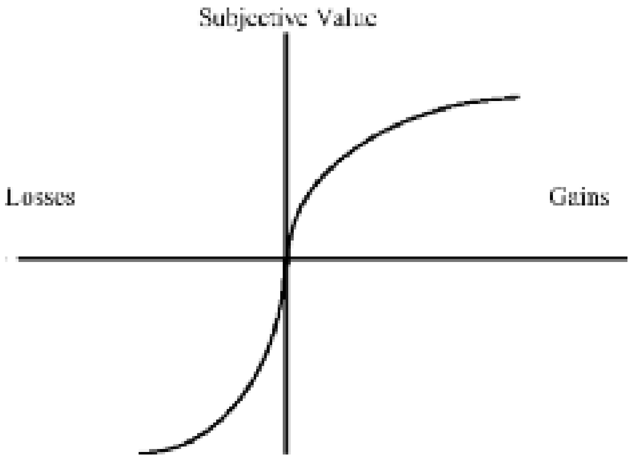


Figure 2: Prospect Theory



a fair game populated by rational representative agents. If this is the case, according to the constant expected returns (CER) model, 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 I found evidence of a very high predictive efficiency of the odds. But the representative agents, who populate the CER model, share the same information set; this is an assumption which seems implausible, particularly in the case of Italian football.

Italian football was plagued by several cases of insider trading. The most famous incidences are known as *Calcioscommesse* in 1980 and *Calcioscommesse 2* in 1986, while the last case of match-fixing, *Calciopoli* in 2006, had a larger scope and consequences bigger than any other before (Boeri & Severgnini, 2011). When there is a chance of match-fixing, the bookkeeper 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 (Bagehot, 1971; Copeland & Galai, 1983; Glosten & Milgrom, 1985; Kyle, 1985). There are several formal analyses of this problem; here I quote the Shin's model (1991). In its benchmark case—where there were no insiders—the odds were inversely proportional to the true probabilities, as it is in the CER model. But, if there is a chance that the bettor knows more about the outcome of the race than the bookkeepers, the optimal pricing response by the bookkeepers is a square root rule by which they trim the odds on long shots relative to favorites. Moreover, if bookkeepers 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* (F/L) bias implies that the bookkeeper returns on favorites should be less than those on long shots.

The analysis of both my datasets found evidence of an F/L bias in the distribution of returns. In the literature, there are several demand side explanations for the existence of a (F/L) bias in betting markets—for example, as bettors' local risk-love and/or behavioral attitudes—but, given the Ill-known evidence of match-fixing in Italian football, I used my data to check if the rumors of match-rigging increased the F/L bias.[3] Therefore, I first investigated if the bias was larger for the suspected matches than for others, then I investigated if the bias was larger for the matches of the club (Juventus) in which managers ran the rigging. Finally, I checked if the bias was larger before the uncovering of the scandal than after the ban of the guilty managers.

When I identified the rigged matches only with those investigated by the judicial authorities—approximately 80 cases in the 2004-2005 season—the first hypothesis was empirically rejected, but this may have been caused by undersampling. The evidence of rigging in some suspicious cases was sufficient for the court to support the hypothesis that rigging was a systematic misbehavior of the Juventus management. That is, even if the matches of the 2005-2006 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. Indeed, when I expanded the examination to include everything before the Calciopoli dataset—1,366 matches from the 2002-2003 season until the 2005-2006 season—I found that the bias was much larger for the matches of the club (Juventus) in which managers ran the rigging than for the rest of the clubs. Finally, when I split the dataset into the two subsets—before and after Calciopoli—I saw that the F/L bias was slightly larger before the uncovering of the scandal than after the ban of the guilty managers.

In summary, I took the Calciopoli case as a kind of natural experiment in which it was common knowledge that match rigging was on the run until 2006 and in which rigging was focused in favor of one team (Juventus). I 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 bookkeepers who face the risk of dealing with insiders.

The Favorite Long Shot Bias

The F/L bias is a systematic tendency of subjects to under-bet or undervalue events that arw characterized by high probability and to over-bet or overvalue those with low probability. Evidence of an F/L bias was found in various laboratory experiments (Preston & Baratta, 1948; Yaari, 1965; Rossett, 1971; Piron & Smith, 1995; Hurley & McDonough, 1995), and field evidence of the F/L bias was found in U.S. horserace wagering markets (Griffith, 1949; McGlothin, 1956; Hoerl & Fallin, 1974; Ali, 1977; Snyder, 1978; Asch, Malkiel, & Quandt, 1982; Thaler & Ziemba, 1988). Similar evidence was also found in

U.K. racetrack wagering markets (Figgis, 1951; Dowie, 1976; Royal Commission on Gambling, 1978; Henery, 1985; Vaughan, Williams, & Paton, 1996, 1997).

The hypothesis of an F/L bias was instead rejected by studies of the Japanese and Hong Kong horserace wagering markets (Busche, 1994; Busche & Hall, 1988), U.S. small racetrack (Swindler & Shaw, 1995), U.S. baseball and hockey wagering (Woodland & Woodland, 1994, 2001), and Australian football wagering (Schnytzer & Weinberg, 2008).[4] In summary, the F/L bias is a quite common, but not universal, feature of sport betting markets.

Several demand-side explanations have been suggested for the F/L bias. The evidence of an F/L 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 nonlinear probability weights.

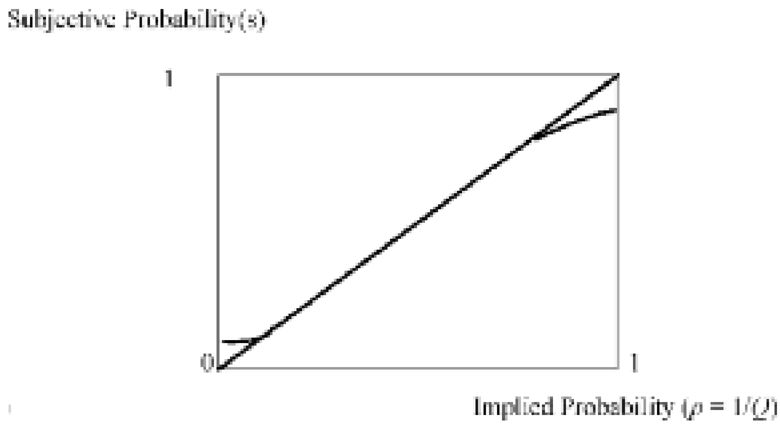
The models which assume a locally concave utility function may be justified by the observation that, in any lottery, the amount returned to the winners is less than the sum of all bets; the difference is the bookkeeper's profit, which is called *take out*. The take out comes from the difference between the odds posted by the bookkeeper and their fair value.[5] The bookkeepers' return π is positive if $\sum pi > 1 = \sum fj$, where pi is the probability implied by the odd Qi ($pi = 1/Qi$) and fj is the probability of the event i. So lotteries have a negative expected return for bettors, and, according to economic theory, only risk-loving agents could buy negative expected return assets.[6] 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-averse for high stakes. The 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 Ill 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 Kahneman and Tversky (1979, 1991), Machina (1982), and Camerer (1989). Finally, the prospect theory (Kahneman & Tversky, 1979) argues that the curvature of the value function is steeper in the loss domain than in the gain domain; this is another candidate explanation for the F/L asymmetry.[7]

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

The F/L bias could also originate from bettors' loss-aversion, a behavioral 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-averse, they underbet the favorite more than usual and overbet horses at odds that would eliminate

Figure 3: Subjective Probabilities Misperceptions



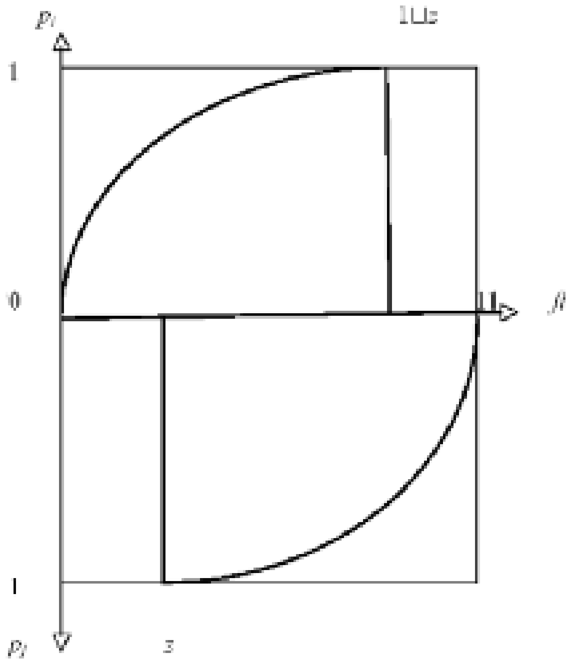
their losses (Kahneman & Tversy, 1979; Thaler & Ziemba (1988). As a result of this behavior, horses with shortened odds (bet down) are more likely to be favorites than those that lengthen long shots. The empirical studies of McGlothin (1956), Alì (1977), Asch, Malkiel, and Quandt (1982) support this loss aversion explanation, while Snowberg and Wolfers (2010)—examining exotic bets—did not find that the bias was more pronounced in the last race of the day in 5,610,580 horserace starts in the U.S. 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 & Wolfers, 2010, pp. 5-6).

Another behavioral explanation of the F/L bias stems from the observation that “… people adopt mental accounts and act as if the money in those accounts is not fungible” (Thaler & Ziemba, 1988, p. 171). Therefore, if bettors discount a fixed fraction of their losses, the F/L bias can arise as a consequence of bettors underweighing their losses compared with their gains. As Henery (1985) and Williams and Paton (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). Behavioral studies found that people are systematically poor at discerning between small and tiny probabilities (Slovic, Fishoff, & Lichtenstein, 1982), and hence price. In addition, people exhibit a strong preference for certainty over extremely likely outcomes (Kahneman & Tversky, 1979), leading highly probable gambles to be underpriced. A graphical illustration of the point may be seen in Figure 3 in which the odds’ implied probabilities ($0 \le p_i = 1/Q_i \le 1$) 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 underprice favorites and overprice long-shots.

While, in the experiments of Viscusi and O’Connor (1984) and Dwyer (1993), this behavioral bias disappears as subjects get experienced with the game, empirical support of this misperception explanation comes from the field study of Snowberg and

Figure 4: Optimal Prices (p) in Front of a z Percentage of Insiders



Wolfers (2010) on the U.S. racetrack wagering market. By using a very large dataset, they compared the neoclassical explanation of the F/L bias—bettors are rational and local risk lovers—with the alternative behavioral explanation—bettors are irrational and risk neutral—and they claim to have found evidence in favor of the view that misperceptions caused the F/L 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 bookkeepers 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 bookkeepers 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 the long shot. Shin (1992) formalized an extensive form game in which there is a constant probability (z) that the bookkeeper will face an informed bettor who knows exactly the outcome of the game. If this is the case, the problem for the price-setting bookkeeper is to set the odds ($Qi = 1/p_i$), such as to maximize them:

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

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:

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

If there are no insiders *($z = 0$)*, the result of this model is identical to that of the CER model: The prices coincide with the true probabilities (*pi = fi*), as the proportion of the total wealth wagered on each event (*Bi/ Bi*) is equal to its probability (*fi*). 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*/*fi* < z*j*/*fj* if *fi* > *fj*), the bookkeeper maximizes its profits by setting the prices so that *pi/pj* < *fi/fj*, whenever *fi*>*fj*.

Figure 4 (from Shin, 1991; see Figure 2) shows the quadratic rule used by bookkeepers to set the prices (odds) in front of an expected percentage (z) of informed traders. If bookkeepers 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 bookkeeper's profit, this implies that the ex-post takeout rate should be higher for long shots than for favorites. The F/L bias may therefore be a supply side (bookkeepers) 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.

Evidence of Match Rigging in Italian Football: Calciopoli

Italian football is used to match fixing. As early as 1927, the Italian Football Federation (FIGC) revoked a championship won by Torino since its managers bribed a Juventus player before the Turin derby. In the 1980s, two famous cases of insider trading were uncovered: Calcioscommesse in 1980 and Calcioscommesse 2 in 1986. The Calcioscommesse cases mainly involved players who gained their monetary profit from betting in the black market. The last case of match-fixing, Calciopoli in 2006, was bigger than any other before.

In the Calciopoli case, referees, federation officials, owners, and team managers cooperated in collusion to alter the final ranking of the Italian First League (Serie A) tournaments in favor of certain teams—mainly in favor of the Juventus team, in which 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 on the Juventus team. The scandal, commonly referred to as Calciopoli, resolved after around 20 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 rigged games by choosing referees favorable to Juventus and manipulating news on television and newspapers against the referees who did not display a favorable attitude toward the team of Moggi.[7]

Table 1. Odds: Descriptive Statistics

	Home			Draw			Visitor		
Bookkeeper	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
Standard Deviation	1.23	1.29	1.25	0.69	0.75	0.69	2.94	2.89	3.02

Note. Italian Serie A, 2007-2008 season; $n = 289$.

Table 2. Odds: Paired *t*-tests

t-test	Home			Draw			Visitor		
$n = 289$	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

Note. Italian Serie A, 2007-2008 season; $n = 289$. Home: odds posted on the victory of the home team; Draw: odds posted on the draw event; Visitor: odds 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

Note. Italian Serie A, 2007-2008 season; $n = 289$. Home: odds posted on the victory of the home team; Draw: odds posted on the draw event; Visitor: odds posted on the victory of the visiting team

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 the time—at least since Luciano Moggi became general manager of Juventus in 1994.[8] Actually, well before the uncovering of Calciopoli evidence, rumors of rigging were already widespread among Italian football fans (Garlando, 2005).

The verdict from the sport justice panel on Calciopoli was determined as follows: Even if the matches of the 2005-2006 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 2006-2007 season. Milan was kept in Serie A but penalized by 8 points. Fiorentina was banned from the Champions League and penalized with a deduction of 15 points. Lazio was sanctioned with a reduction of 3 points and the exclusion from the Union of European Football Associations (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 FIGC and of the Italian Association of Referees (AIA).[9]

Table 4. Odds: OLS Analysis

Odd	Dependent Variable	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

Note. Italian Serie A, 2007-2008 season; $n = 289 \times 3$. Home: odds posted on the victory of the home team; Draw: odds posted on the draw event; Visitor: odds posted on the victory of the visiting team.

Empirical Analysis

Data and Descriptive Statistics

In a football game, a bettor can 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 bookkeeper will return for a unit bet if the event i (i=H,V,D) is realized. I started my analysis of the Italian football wagering market by analyzing a panel dataset which consisted of the odds posted by three bookkeepers—MatchPoint (MP), MisterToto (MT), and SportingBet (SB)—on 289 games played in the 2007-2008 Serie A season.[10] By comparing the odds offered by the three bookkeepers, I saw that their distributions were quite similar in their first two moments (see Table 1). Actually, the t-tests cannot reject the hypotheses: $Q_{D,LM} > Q_{D,MT} > Q_{D,SB}$ and $Q_{H,LM} > Q_{H,MT}$ (see Table 2). The distributions are homoskedastic, and the paired linear correlation coefficients are very close to unity (see Table 3). I looked for system-

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)*

Note. Italian Serie A, 2007-2008 season; 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

	Percent 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
Percent Total	84.4	83.7	86.2	99.0

Note. Italian Serie A, 2007-2008 season; n = 289 x 3. Home: victory of the home team; Draw: draw result; Visitor: victory of the visiting team.

atic differences in the odds by regressing the odds of one firm with the odds of another (see Table 4). Under the null hypothesis of no systematic differences between the firms, the intercept should be indistinguishable from 0 and the slope coefficient should equal unity. The results of my ordinary least squares (OLS) regressions cannot reject the hypothesis $H_0: \; \approx 1$ for any regression, but they rejected the hypothesis $H_0: \; \approx 0$ in 6 of 9 cases, mainly because MT's odds are loIr than the others. All the determination coefficients of the 9 regressions are very high, close to unity.

In summary, my results suggested that the odds offered by the three firms were highly correlated, although some firms systematically offered odds which were marginally, but often significantly, lower/higher than the others. Given these similarities, I summarized the market odds distribution by the distribution of a synthetic odd, named Delphi, which is the unweighted average of the odds offered by the three bookkeepers. The frequency distributions of the Delphi average odds Were close to normal for the victory of the home/visiting team event—but the last one was asymmetric; the frequency distribution for the draw odds was strongly leptokurtic and asymmetric.[11]

The Correlation Between the Game Results and the Associated Market Odds

The bookkeeper ex-post profit is $\pi = \sum B_i - \sum(f_i \, B_i \, Q_i)$, where B_i is the amount bet on the event i, and f_i is its frequency. Therefore, given f_i and $\sum B_i$, the iso-profit curves of

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)*			

Note. Standard deviations of estimated coefficients are shown in parentheses.

Table 8. Bookkeeper Take-Out: Descriptive Statistics

	Home	Draw	Visitor	Total
Frequency	137	78	74	289
Frequency (%)	0.47	0.27	0.26	1.00
Implied Probability	0.48	0.32	0.30	1.10
Take Out Rate	0.01	0.05	0.04	0.10

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

the bookkeeper map a family of hyperboles in the Cartesian plane, described by the odds 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 in which to study a betting market is the Constant Expected Return Model (CERM), 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 in which 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 CER model is true, 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.[12] In order to measure this kind of predictive efficiency of the Italian soccer wagering market, I used a multinomial logit analysis.

I ran three multinomial logit regressions, one for each bookkeeper, in which I regressed the vector of games results (y_i) on the matrix of the associated odds ($Q_{i,j}$) offered by each bookkeeper, $y_i = f(Q_{Hi}, Q_{Di}, Q_{Vi})$, where y = H,D,V and i = 1, …, 289. I found that, for each bookkeeper, the odds had a significant and very strong predictive power (see Table 5). All the pseudo-R^2 Were very high, and all the F^s were highly significant; the matching between the predicted and actual results was very high. For each of the three bookkeepers, the multinomial logit models correctly predicted about

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

Note. Italian Serie A, 2007-2008 season; n = 289 x 3. 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 Datasets	0.77% *(538)*	5.97% *(650)*	0.56% (*332*)	5.91% (*413*)
Juventus Only	4.28 *(107)*	9.67% (*152*)	—	—

Note. Before Calciopoli dataset: 2002-2006 seasons, odds posted by one bookkeeper; After Calciopoli dataset: 2007-2008 season, odds posted by three bookkeepers; number of observations in parentheses.

80-90% of the results (see Table 6). Moreover, if I regressed the game results on the unweighted average of the associated odds offered by the three bookkeepers (the Delphi odds), I got an almost perfect prediction of the actual game results. The predictive efficiency of the market was also confirmed by an OLS analysis. As in Kuypers (2000), for each result—home win (H), visitors win (V), and draw (D)—I 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: $y = a + bq +$. According to the results showed in Table 7, the hypothesis that the implied probabilities Were good predictors of the actual game result ($H_0 : b = 1$) cannot be rejected at the 95% confidence level.[13]

The Returns Distribution and the F/L Bias

By comparing the odds with the games results (see Table 8), I measured an ex-post take out rate of approximately 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 F/L bias indicated that the bookkeeper returns on favorites Were less than those on long shots, H_0

$: \pi(f_j) > \pi(f_i)$, $iff_i > f_j$. In order to find evidence of this bias, I identified the F/L events in my 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, I 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 classified the F/L events in the $\text{High}_{ij}/\text{Low}_{ij}$ subsets, which consisted of the events in which expected probability was higher/lower. By construction, the subset $\text{Low}_{ij}/\text{High}_{ij}$ was the subset consisting of the events associated with the higher/lower odds, it was the lower/higher tail of the q_{ij} distribution, and it consisted of about 1/6 of the whole distribution. The subset Inside_{ij} was the central body of the distribution and it consisted of about 2/3 of the distribution itself.

By comparing the odds with the associated results, I had an ex-post take out distribution for each subset (see Table 9). The hypothesis of the F/L bias is $H_0 : \pi(\text{Low}_{ij}) > \pi(\text{High}_{ij})$, and my results supported this hypothesis since $\pi(\text{Low}_{ij})$ was actually larger than $\pi(\text{High}_{ij})$ for any i and j. On average, the take out rate was 6% for long shots and 1% for favorites. Specifically, $\pi(\text{High}_{ij})$ was negative for I = H,D, while $3\% < \pi(\text{Low}_i) < 8\%$. This result was confirmed by the analysis of another dataset, consisting in the results and the odds offered by one bookkeeper (MP) on 6,369 games played from the 2002-2003 season until the 2007-2008 season. For this dataset, $\pi(\text{Low}_i) > \pi(\text{High}_i)$ for any i. On average, the take out rate was 6% for long shots and 1% for favorites.[14] Specifically, $\pi(\text{High}_i) < 0$ for i = H,D, while $4\% < \pi(\text{Low}_i) < 8\%$. In summary, the analyses of both datasets support the hypothesis that the F/L bias is an empirical feature of the Italian football wagering market.

Match Rigging and the F/L Bias

The Calciopoli case presented a natural experiment in which I could see if the rumors of rigging—which were shown to be true—influenced the wagering market. That is, (a) there was 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 before 2006; c) the verdict banned those manager from the Italian football so that rigging was not run after 2006—or, at least, it was not run by the same people.

If I assume that the bookkeepers were able to detect the rigged matches—so that they could adjust their odds according to the expected percentage of insiders—the first test may have consisted 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 in these games than the rest, then according to Shin's model, the bias should have been larger for the 80 suspected games than for the other 340.[15] My analysis of the 2004-2005 Serie A season dataset found the opposite result: The F/L bias was larger for the rest of the dataset, $\pi(\text{Low}) = 5\% > \pi(\text{High}) = 0\%$, than for the suspected games, $\pi(\text{Low}) = 4\% > \pi(\text{High}) = 2\%$. There may be many explanations for this result. First, the verdict states that rigging was finalized to alter the match results in order to promote some teams (mainly, Juventus) 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 my results may come from the inferential procedure that 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 show that rigging was a systematic misbehavior 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 from Italian football up to 5 years starting from 2006.

Therefore, I split the whole dataset into two subsets. The first consisted of the results and the odds offered by the MP bookkeeper from the 2002-2003 season until the 2005-2006 season, when rigging was on the run. The second dataset consisted of the already reviewed results and odds posted by three bookkeepers for the 2007-2008 seasons, after the Calciopoli verdict. For each subset I applied the same procedure as before to select the F/L subsets. Comparing the F/L bias shown before and after Calciopoli (see Table 10), I saw the F/L bias was slightly higher (1.4%) before Calciopoli than after. Moreover, when I analyzed the subset consisting only of the matches played by the most suspected team (Juventus) before Calciopoli, I saw that the F/L bias in these games (14%) was double what was presented in the whole dataset (7%).[16] In summary, by expanding my analysis beyond the judicial cases, I saw that the F/L bias was slightly larger before the Calciopoli case than later on and 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).

Conclusions

In this paper, I investigated 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. This was a case of systematic match rigging run by top federation and team managers until 2006. Preliminarily, I tested, using a multinomial logit analysis and ordinary least square procedure, whether or not the market odds predicted the football game results by. My results confirmed that the odds did have a very significant predictive power. I then looked for evidence of the F/L bias, an empirical feature in which the bookkeepers' return on favorites Were less than those on long-shots. I found that the F/L bias was an empirical feature of the Italian football wagering market. There may be many demand-side reasons, which could add together to explain the F/L bias, but the F/L bias may also be the equilibrium result of a game in which bookkeepers expect to deal with unknown insiders (Shin, 1991). Therefore, I took Calciopoli as a kind of natural experiment that could be useful to determine whether or not 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 bookkeepers 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. I 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.

In summary, I found that the Italian football betting market was weakly efficient but that the odds displayed a persistent F/L bias as Ill. This bias was larger during Calciopoli, but it did not vanish following it. Therefore, I suspect that the F/L bias could be the sum of several demand-side factors—bettors' local risk love or bettors' behavioral attitudes)—but I also anticipate that this bias may have originated from the strategic behavior of bookkeepers who were expecting to deal with unknown insiders.

References

Alì, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy, 85,* 803-815.

Asch, P., Malkiel, B. G., & Quandt, R. E. (1982). Racetrack betting and informed behavior. *Journal of Financial Economics, 10,* 187-194.

Asch, P., Malkiel, B. G., & Quandt, R. E. (1984). Market efficiency in racetrack betting. *Journal of Business, 57,* 165-175.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal, 22,* 12-14.

Boeri, T., & Severgnini, B. (2010). Match rigging and the career concerns of referees. *Labour Economics, 18,* 233-254.

Busche, K. (1994). Efficient market results in an Asian setting. In D. B. Hausch, V. S. Y. Lo, & W. T. Ziemba (Eds.), *Efficiency of racetrack betting markets* (pp. 615–616). New York: Academic Press.

Busche, K., & Hall. C. (1988). An exception to the risk preference anomaly. *Journal of Business, 61,* 337-346.

Camerer, C. (1989). Does the baseball market believe in the hot hand? *American Economic Review, 79,* 1257-1261.

Copeland, T. & Galai, D. (1983). Information effects on the bid ask spread. *Journal of Finance, 38,* 1457-1469.

Conlinsk, J. (1993). The utility of gambling. *Journal of Risk and Uncertainty, 6,* 255-275.

Crafts, N. F. R. (1985). Some evidence of insider trading knowledge in horse race betting in Britain. *Economica, 52,* 295-304.

Dowie, J. (1976). On the efficiency and equity of betting markets. *Economica, 3,* 139-150.

Dwyer, G., Williams, A. W., Battalio, R. C., & Mason, T. I. (1993). Tests of rational expectations in a stark setting. *Economic Jounal, 103,* 586-601.

Figgis, E. L. (1951). *Focus on gambling.* London: Barker.

Figlewsky, S. (1979). Subjective information and market efficiency in a betting market. *Journal of Political Economy, 87,* 75-89.

Friedman, M., & Savage, L. (1948). The utility analysis of choices involving risk. *Journal of Political Economy, 56,* 279-304.

Gabriel, P., & Marsden, J. (1990). An examination of market efficiency in British racetrack betting. *Journal of Political Economy, 98,* 874-885.

Gabriel, P., & Marsden, J. (1991). An examination of market efficiency in British racetrack betting: Errata and corrections. *Journal of Political Economy, 99,* 654-659.

Garlando, L. (2005). *Nostra signora del dischetto.* Milan, Italy: Kowalski.

Gilovich, T. (1983). Biased evaluations and persistence in gambling. *Journal of Personality and Psychology, 44,* 1110-1126.

Gilovich, T., & Douglas, C. (1986). Biased evaluations of randomly-determined gambling outcomes. *Journal of Experimental Social Psychology, 22,* 228-241.

Glosten, L., & Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneous informed traders. *Journal of Financial Economics, 14,* 71-100.

Griffith, R. (1949). Odds adjustments by American horse race bettors. *American Journal of Psychology, 62,* 290-294.

Henery, R. J. (1985). On the average probability of losing bets on horses with given starting price odds. *Journal of the Royal Statistic Society, 148,* 342-349.

Hoerl, A. E., & Fallin, H. K. (1974). Reliability of subjective evaluations in a high incentive situation. *Journal of the Royal Statistic Association, 137*, 227-30.

Hurley, W., & McDonough, L. (1995). A note on the Hayek hypothesis and the favorite long shot bias in parimutuel betting. *American Economic Review, 85*, 949-55.

Kanheman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-92.

Kanheman, D., & Tversky, A. (1991) Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106*, 1039-61.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica, 63*, 1315-35.

Losey, R. L., & Talbott, J. C., Jr. (1980). Back on the track with the efficient market hypothesis. *Journal of Finance, 35*, 1038-43.

Machina, M. (1982). Expected utility analysis without the independence axiom. *Econometrica, 50*, 277-323.

Machina, M. (1987). Choices under uncertainty: Problems solved and unsolved. *Journal of Economic Perspectives, 1*, 121-54.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy, 56*, 151-8.

McGlothin, W. H. (1956). Stability of choices among uncertain alternatives. *American Journal of Psychiatry, 69*, 604-15.

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

Pope, P., & Peel, D. A. (1989). Information, prices, and efficiency in a fixed-odds betting market. *Economica, 56*, 322-41.

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

Quandt, R. (1971). Betting and equilibrium. *The Quarterly Journal of Economics, 101*, 201-08.

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

Royal Commission on Gambling (1978). *Final report* (Vol. 1). London: H. M. Stationary Office.

Samuelson, P. (1952). Probability, utility, and the independence axiom. *Econometrica, 20*, 670-8.

Sauer, R. D. (1998). The economics of wagering markets. *Journal of Economic Literature, 36,* 2021-2064.

Schnytzer, A., & Shilony, Y. (1995). Inside information in a betting market. *Economic Journal, 105*, 963-71.

Schnytzer, A., & Weinberg, G. (2008). Testing for home team and favorite biases in the Australian rules football fixed-odds and point spread betting markets. *Journal of Sports Economics, 9,* 173-90.

Shin, H. S. (1991). Optimal odds against insider traders. *The Economic Journal, 101*, 1179-85.

Shin, H. S. (1992). Prices of state contingent claims with insider traders and the favorite-longshot bias. *The Economic Journal, 102*, 426-35.

Shin, H. S. (1993). Measuring the incidence of insider trading in a market for state-contingent claims. *The Economic Journal, 103*, 1141-53.

Snyder, W. (1978). Horse racing: Testing the efficient market models. *Journal of Finance, 33*, 1109-18.

Slovic, P., Fishoff, B., & Lichtenstein, S. (1982). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristic and Biases.* Cambridge, MA: Cambridge University Press.

Snowberg, E., & Wolfers, J. (2007). The favorite-longshot bias: Understanding a market anomaly. In D. Haush and W. Ziemba (Eds.), *Efficiency of Sports and Lottery markets.* Amersterdam: Elsevier.

Snowberg, E., & Wolfers, J. (2010). Explaining the favorite-longshot bias: Is it risk-love or misperceptions? *Journal of Political Economy, 118*, 723-746.

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

Tertullianus (1997). *De Spectaculis*, Cambridge, MA, Harvard University Press. (Original work published c. 191 202 AD)

Thaler, R. H., & Ziemba, W. T. (1988). Anomalies-parimutuel betting markets: Racetrack and lotteries. *Journal of Economic Perspectives, 2*, 161-74.

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

Vaughan Williams, W.L. & Paton, D. (1997). Why is there a favorite long shot bias in British racetrack betting markets. *Economic Journal, 107*, 150-58.

Viscusi, K., & O'Connor, C. (1984). Adaptive responses to chemical labeling: Are workers Bayesian decision makers? *American Economic Review, 74*, 942-56.

Weitzmann, M. (1965). Utility analysis and group behavior: An empirical study. *The Journal of Political Economy, 73*, 18-26.

Woodland, L. M., & Woodland, B. M. (1994). Market efficiency and the favorite long shot bias: The baseball betting market. *Journal of Finance, 49*, 269-80.

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

Yaari, E. (1965). Convexity in the theory of choice under risk. *Quarterly Journal of Economics, 79*, 278-90.

Endnotes

1 "Aspice populum ad id spectaculum iam venientem, iam tumultuosum, iam caecum, iam de sponsionibus concitatum," Tertullianus (c. 155-230 AD; pp. 271-272).

2 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 on the black market.

3 For a review, see Sauer (1998).

4 For a review of empirical studies about the F/L bias, see SnoIr and Wolfers (2007).

5 Another explanation, alternative to the assumption of a risk-loving attitude of bettors, may stem 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: $E(G,p,W) = pU(W+G) + (1-p)U(W-L) + V(G,p)$, 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 [$V(G,p)$] 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 $V(G,p)$, where 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 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.

[6]"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 & Tversky, 1979, p. 269). Figure 3 gives a graphical illustration of the *prospect theory*.

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

[8]Boeri et al. (2010) listed the most dubious referee decisions favorable to Juventus in the years between 1994 and 2004.

[9]See Boeri and Servignini (2010) for further analysis of the Calciopoli case, with particular emphasis on the behavior of referees approached by the sanctioned clubs.

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

[11]"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 & Peel, 1989, p.328).

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

[13]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. See the Pope and Peel quote in comment cited in endnote.

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

[15]I draw the list of suspected matches from Boeri and Severgnigni (2010, Appendix A, Tables 9 & 10).

[16]My 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.

Author's Note

I am grateful to Agipronews and Sisal for their kind supply of data and to the anonymous referees.