DISCUSSION PAPER SERIES

IZA DP No. 3745

The Italian Job: Match Rigging, Career Concerns and Media Concentration in Serie A

Tito Boeri
Battista Severgnini

October 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Italian Job: Match Rigging, Career Concerns and Media Concentration in *Serie A*

Tito Boeri
Bocconi University-IGIER, CEPR and IZA

Battista Severgnini
Humboldt University Berlin and CEBR

Discussion Paper No. 3745
October 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3745
October 2008

ABSTRACT

The Italian Job: Match Rigging, Career Concerns and Media Concentration in *Serie A*[*]

This paper contributes to the literature on competition and corruption, by drawing on records from *Calciopoli*, a judicial inquiry carried out in 2006 on corruption in the Italian soccer league. Unlike previous studies, we can estimate the determinants of match rigging and use this information in identifying corruption episodes in years in which there are no pending judicial inquiries. We find evidence of corruption activity well before *Calciopoli*. Career concerns of referees seem to play a major role in match rigging. An implication of our study is that a more transparent selection of the referees and evaluation of their performance is essential in removing incentives to match rigging. Another implication is that in presence of significant "winners-take-all" effects, more competitive balance may increase corruption unless media concentration is also significantly reduced.

JEL Classification: D73, L82, L83

Keywords: concentration, corruption, career concerns, random effect ordered probit, Monte Carlo simulations, soccer

Corresponding author:

Tito Boeri
IGIER – Bocconi University
Via Salasco 3/5
20136 Milan
Italy
E-mail: tito.boeri@uni-bocconi.it

[*] We would like to thank Emanuela Audisio, Vincenzo Galasso, Aldo Grasso, Andrea Ichino, Jens A. Kovsted, Michele Pellizzari, Jan Rose Skaksen and Luigi Zingales as well as seminar participants at the Berlin Labour Market Research Network (BeNA), at IGIER-Bocconi, at Copenhagen Business School and at the conference Economics and Psychology of Football held at University of Innsbruck. Battista Severgnini acknowledges the research support of the InterVal project (01AK702A) which is funded by the German Ministry of Education and Research. All errors are our own.

The result is that Juve and Milan can often rig the system to assign themselves the most mediocre, provincially minded referees, who are (subconsciously) more deferential towards their prestige clubs. [...] Other referees who have issued critical penalties against Juve have found themselves working games in the lowly Serie B. [...] Only on a few occasions have some of the submerged sordid details come to surface. [...] Undeniably, the benefits of friendly refereeing accrue to Juventus and Milan more than any other clubs in Italy.
What's shocking [...] is how often Juventus have won the championship at the end of the season on a piece of dubious refereeing. It is worth seeing with one's own eyes the phantom penalties that have deprived Juve's opponents of vital goals. You'll see clips of the ball crossing Juve's goal line, yet inexplicably not counted against them. [...] Even though Juve committed more fouls than any club in the league, they received the least red cards, a statistical inconsistency that defies logical reckoning. [...] [G]ranted Juve a dubious penalty for a transparent piece of thespianism, where the cause of a player's flop to the ground could not be explained by any known law of physics.

Franklin Foer, *How Soccer explains the World*, Harper, 2004 (pp. 170 and 174).

1 Introduction

No question, Italian soccer has been for many years subject to repeated episodes of corruption aiming altering the outcome of crucial matches. Everybody operating in the field was aware of this fact, but no action was taken. The above quotation is drawn from a book of a careful observer of Italian soccer. It anticipates some of the main conclusions of *Calciopoli*, a judicial inquiry carried out in 2006 on corruption in the Italian soccer League. Actually, the degree of corruption turned out to be even worse than described in Foer's book: rather than simply being subconsciously deferential toward Juventus, referees were actually chosen in advance and blackmailed using the power that the corrupting managers had over the media. Moreover, the managers of the major clubs could not only control the assignments of referees to the match involving their own team, but also those of competing teams. Hence, they could plan ahead of matches with one of these teams, inducing the referees to give a red card to the strongest player of the opponent in the following match. Referees not cooperating were banned from the most important matches, reducing their career prospects. Moreover, major TV shows and complacent journalists (also concerned about their career prospects) were heavily criticizing the decisions precisely of those referees who did not cooperate. Reputation is essential in a career of a referee and being seriously criticized in a popular TV show, like the *Processo del Lunedì*,[1] means not being selected in the pool of referees for the international matches, those

[1]This is how Franklin Foer describes the *Processo del Lunedì*: "One of my favorite television shows around the world is a show called *Il Processo* in Italy, which means *The Trial*. And, literally, there's a panel of journalists, and inevitably one leggy woman with a dress with a slit up to her bellybutton who decide if the

best paid. Overall, corruption in Italian soccer was widespread and a longstanding equilibrium phenomenon, built on media power, carreer concerns and tacit conspiracy. Clearly not all games were rigged: the conspiracy was concentrated on a subset of crucial games. There was a very careful targeting of rigging efforts. These facts, together with the wealth of data available to researchers on various features of Italian soccer, makes *Calciopoli* a very interesting case study on corruption and media concentration.

Corruption has been widely investigated by economic theory. The pioneering works of Becker and Stigler (1974) and Rose-Ackerman (1975) provided a basic framework in terms of models of principal-agent and of production distortion. Shleifer and Vishny (1993) used a fishing game framework to analyse corruption in licensing arrangements while Bliss and Tella (1997) extended models of rent-seeking to analyse the relationship between competition and corruption. There has been much less empirical work to date on this issue. This is hardly surprising given the paucity of data on corruption. The key references are Gambetta (1993), Ades and Tella (1997), and, on the macroeconomic side, Mauro (1995), who analysed corruption in relation to economic growth. Duggan and Levitt (2002) and Wolters (2006) analysed corruption in Japanese sumo's wrestler and basketball, respectively. We develop on their methodology in one of the strategies to predict corruption in this paper. None of the studies we are aware of addressed empirically the links between corruption, career concerns in other domains. Also the implications on the relationship between corruption and market structure have some relevance for media concentration.

Three distinguishing features of this paper are: i) the possibility to compare theoretical and empirical predictions as to the rigged matches with information on the matches which are actually under inquiry because of documented pressures on the referee, ii) the fact of drawing from this analysis implications as to the relationship between career concerns and corruption and iii) the focus on the distribution of both, media and sport-competitive power, across clubs. Although the assumptions and the structure of the model and the empirical results in this paper are specific to the professional sport industry and the data refer to Italian soccer, our results may have some relevance also concerning the relationship between corruption and other industries characterised by "winner-takes-all" preferences.

We find that a simple model of optimal allocation of corruption fees across games of the Italian Championship can predict rather well the timing and the choice of matches to be rigged. This model also yields predictions as to the relationship between media power, competitive balance in sport events and corruption. In order to detect corruption before *Calciopoli* and test these implications of the model, we first estimate a probit model of corruption, using

referee has been judicious in handing out punishment or failing to hand out punishment. And you have these clips of players falling to the ground or getting tripped that are played in super-slow motion over and over again, and are debated. And there's always this whiff of conspiracy within this discussion that somebody is paying off somebody to get these charges, that the referees can't possibly be acting in the best interest of the game. It is the single most delightful spectacle."

the judicial records to identify rigged matches and we take into account of career concerns of referees. Next, we draw on a widely used empirical method predicting fair outcomes of soccer matches in the Italian league, based on information on the competitive strength of the various teams. Exploiting our unique dataset, we can carry out a careful specification search and choose the functional form that better fits the data and is orthogonal to the outcomes of documented corruption episodes in the 2004-5 Championship. This enables us to identify potential corruption episodes combining three key criteria: first, they must deviate from the predictions of the empirical model at different confidence intervals; second, these matches must have a high probability of corruption in light of the probit model of corruption; third there must be significant discretion in the selection of the referees assigned to these matches (according to the procedures detailed in Section 5, matches must be in grid A). We find that using only the first criterion, i.e., identifying corruption as outliers of regressions predicting fair outcomes, exposes to many type II errors (outlier games on which there is no evidence of corruption).

The plan is as follows. Section 2 offers a short history of the Italian scandal and presents and few descriptive statistics about the rigged matches. Section 3 provides a toy model of optimal allocation of "blackmail capital" in affecting outcomes of professional sport events. Section 4 applies our empirical strategies drawing on this model, first estimating the determinants of documented corruption episodes and then predicting outcomes of matches. Based on these results, Section 5 identifies corruption episodes before the 2004-5 Championship. Finally, Section 6 summarizes our results and provides suggestions for further research.

2 A Round History of *Calciopoli*

2.1 Italian Soccer before 2004

Corruption is not a new phenomenon in the Italian Soccer First Division, the Italian *Serie A*. Already in 1927 the Italian football federation revoked the championship just won by Torino Calcio since its managers bribed a Juventus soccer player before the Turin derby; in 1980 A.C. Milan and Lazio Rome were relegated to the second division after fixing a match and some of their players were found guilty of illegal gambling on soccer games.

Although the rigged matches before 2004 are unknown, there are a number of indications that widespread corruption was present at least since Luciano Moggi and his assistants were hired by Juventus in 1994.[2] According to Alessandro Gilioli's of the Italian weekly magazine *L'Espresso*,[3] and the book written by Garlando (2005) on Juventus penalties from 1929 to

[2]Before the season 1993/1994, Moggi was the general manager of Napoli and, then, of Torino Calcio. He had been involved in a scandal after offering prostitutes to referees in order to obtain preferential treatment during matches involving Torino Calcio.

[3]"L'Avvocato, la cupola e sette scudetti", May 25th 2005

2005, the following matches can be associated to corruption activity by Juventus managers.

Champ.	*M. Day*	*Match*	*Result*	*Rigged Episode*
1994/1995	18	**Juventus**-Brescia	2-1	Last minute irregular penalty
1996/1997	20	**Juventus**-Perugia	2-1	Perugia was denied penalty
1997/1998	3	**Juventus**-Brescia	4-0	Brescia was denied penalty
1997/1998	11	**Juventus**-Lazio	2-1	Penalty for Juventus
1997/1998	19	**Juventus**-Roma	3-1	Favors to Juventus
1997/1998	21	**Juventus**-Sampdoria	3-0	Inexistent goal for Juventus
1997/1998	25	**Juventus**-Napoli	2-2	Favors to Juventus
1997/1998	30	Empoli-**Juventus**	0-1	Empoli was denied goal
1997/1998	31	**Juventus**-Inter	1-0	Inter was denied penalty
1999/2000	33	**Juventus**-Parma	1-0	Parma was denied goal
2001/2002	3	**Juventus**-Chievo	3-2	Penalty for Juventus
2001/2002	14	A.C.Milan-**Juventus**	1-1	Penalty for Juventus
2001/2002	15	Inter-Chievo	1-2	Inter was denied penalty
2002/2003	17	Chievo-**Juventus**	1-4	2 penalties for Juventus
2002/2003	20	**Juventus**-Empoli	1-0	Penalty for Juventus
2002/2003	29	**Juventus**-Roma	2-1	Penalty for Juventus
2003/2004	10	Modena-**Juventus**	0-2	Favors to Juventus
2003/2004	16	Sampdoria-**Juventus**	1-2	Favors to Juventus
2003/2004	24	Brescia-**Juventus**	2-3	Favors to Juventus

Table 1: Matches likely to have been rigged by Juventus before the Championship 2004/5 according to Garlando and Gilioli

According to Garlando and Gilioli, the rigged episodes had, in most of the cases, a crucial role in affecting the outcome of the match. To give an example, in the 1994/1995 season, one minute before the end of the match Juventus-Brescia, the referee gave to Juventus a non existing penalty. Significantly no corruption episode was identified by these sources in the 1998/99 Championship, when the Italian Soccer Federation changed the systems for assigning the referees, moving from a discretionary (a small team of former referees was in charge) to a random assignment of referees to the different matches. In the following Championships the system for assigning referees was modified making it extremely complex and highly discretionary. Basically matches were ranked and classified in a number of levels (the so-called *griglie*), depending on their importance for the final outcome of the tournament. Then as many referees as matches in each *griglia* were selected on the basis of a (non-public) evaluation of their past performance. Finally, the referees were randomly selected from this pool and assigned to each match in the corresponding *griglia*. However, a number of exceptions (*preclusioni*) were introduced preventing some referees to be considered for a given match. For instance, a referee potentially eligible to *griglia A* matches could be excluded a priori from one match for unspecified reasons. Each griglia had on average 5 matches and 5 referees and the combination of *griglie* and *preclusioni* allowed in many cases to ultimately choose the referees.

The tapped phone conversations suggest that a favorable treatment of the referee was obtained by Moggi by "choosing" referees displaying a strong home bias in matches involving a big team at home. The Italian Championship displays a stronger degree of home bias than any other European Championship. Significantly the switch from the two points to the three points for a win rule, as well the reduction of travel costs, enabling more supporters to follow their team in away matches, do not seem to have reduced the home bias of Italian soccer. This is illustrated by Table 2, displaying the percentage of first division matches resulting in home wins, draws and away wins in Italy and other European countries – based on the data collected by Dobson and Goddard (2001) – before and after the introduction of the "three points for a win rule", as well as the percentage difference between the two periods.

Country	**Championship**	**Win**	**Draws**	**Loss**	**Goal Home**	**Goal Away**
	1982-1990	48.40	26.30	25.30	1.58	1.07
England	1991-1999	45.70	28.50	25.80	1.52	1.09
	Δ%	-5.58	8.37	1.98	-3.80	1.87
	1991-1994	51.70	31.70	16.60	1.45	0.76
France	1995-1999	49.20	31.00	19.80	1.47	0.87
	Δ%	-4.84	-2.21	19.28	1.38	14.47
	1991-1995	45.20	31.20	23.40	1.76	1.19
Germany	1996-1999	44.90	29.30	25.80	1.73	1.22
	Δ%	-0.66	-6.09	10.26	-1.70	2.52
	1992-1995	55.00	19.00	26.00	99.25	60.50
Italy	1996-1999	61.50	15.50	23.00	95.31	52.63
	Δ%	**11.82**	**-18.42**	**-11.54**	**-3.97**	**-13.02**
	1991-1994	50.40	28.40	21.20	1.51	0.94
Spain	1995-1999	47.90	27.10	25.00	1.60	1.09
	Δ%	-4.96	-4.58	17.92	5.96	15.96

Table 2: Home Bias in Italian Soccer

Significantly, Italy is the only country where an *increase* in the number of home wins was observed over time. It is also the only country were the number of goals scored abroad *declined* after the introduction of the new rules.

2.2 A Championship without Winners

In May 2006 a major scandal was uncovered by Italian prosecutors after tapping phone conversations in relation with an investigation on the use of doping at Juventus. They found that the general manager of the Turin based soccer team, Luciano Moggi, during the 2004-2005 season, won by Juventus, have had a large number of contacts with referees, football federation officials and journalists. These contacts were finalised to rig games by choosing referees favorable to Juventus and manipulating news on TVs and newspapers against referees damaging the Turin-based team. As mentioned above, with the exception of the 1998-99

Championship, referees were selected by a team of former referees, with whom Moggi had extensive phone conversations. Moggi was also, via his son Alessandro, controlling the Italian soccer player placement market and had a strong influence on the National team squad.

The final standings of the 2004-05 Italian Championship gave the victory to Juventus, while A.C. Milan, Inter and Udinese qualified for the Champions League, and Bologna, Brescia and Atalanta were relegated to the Second Division. Tables 16 and 17 in the Annex provide a list of all matches under investigation. The results of these matches are compared with a measure (IVS, *Indice di Valutazione di Squadra*, Team Preformance Index) measuring the relative strength of the two teams involved in each match, using an index which captures the fitness of the individual players (based on their performance during the match). This measure is based on the range of players actually available (no disciplinary measures, nor injuries) to each coach during the match, as well as the overall performance of the team in the previous matches. There is a total of 78 matches, that is about 2 per week of the tournament, likely to have been rigged. Importantly, they involve not only Juventus, but are mostly in favour of Juventus, as they condition the outcomes of the other matches in favor of Juventus. The other teams involved in the scandal were A.C. Milan, Fiorentina, Lazio, and Reggina. A.C. Milan was accused of having influenced the assignment of linesmen for its match against Chievo Verona (April 2005); while Diego Della Valle and Claudio Lotito, respectively Fiorentina owner and Lazio chairman, were accused of having used a method similar to Luciano Moggi's in rigging matches throughout referees' designation. The allegations against Reggina were also on the same vein.

The official judiciary documents, as reported by national newspapers,[4] suggest that a variety of methods had been used to affect the outcome of a match. Sometimes a strong player (e.g. Jankulowski in Udinese-Brescia) was given a red card (which means automatically missing the following match) for futile reasons in the match just before the one in which he should have played against Juventus. In other cases, it was the referee in the match under investigation to give a penalty kick or neglect an offside presumably in favor of one of the two teams. In all of these cases, tapped phone conversations certify direct contacts between the managers of the corrupting team, the official selecting the referees and sometimes the referees themselves. Tapped conversations involved also a number of journalists in popular TV shows. Corrupting managers were indeed threatening referees to destroy their reputation by using their media power in case they had not complied with their requests. A referee with a poor reputation would hardly qualify for refereeing international games, which are a major source of revenues for the referees. Thus referees with stronger career concerns were a natural target of match rigging.

Comparing the IVS rankings with the actual results, it would seem that in many cases the outcome would have not changed even in case of a fair match: the winner is always the

[4]We refer to the Italian daily newspapers *Corriere della Sera*, *La Gazzetta dello Sport* and *La Repubblica*.

strongest team. But this is not the point. As argued above, the tricky strategy used by Moggi & friends was to ask the referees to give a red card to the most important players of the rival team during the match *before* the rigged match in order to minimize the risk of a loss or a draw in the following match. This strategy affects the IVS index rather than altering the odds of a game: the match has ex-post, that is after the selection of players operated in the previous match, a ”fair” outcome. The fact that one or two key players are out of the game could significantly reduce the IV scores of a team. To give an example, A.C. Milan IVS index would have been 10% lower without Andrea Pirlo in 2004.

While the judicial inquiry is still pending, the sport justice has already given its verdict and enforced its (mild) sanctions. In particular, the *Federazione Italiana Giuoco Calcio*, the Italian soccer federation decided that Juventus should be relegated to the second division (the Italian Serie B) with a deduction of 9 points in the 2006-7 Championship; A.C.Milan was penalized by 8 points; Fiorentina was excluded by 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 in the first division. Very low pecuniary sanctions were given to the managers presumably involved in match rigging. For instance, Moggi was given a fine of about 30,000 Euros while his wage with Juventus in the year in which corruption was detected, amounted to some 2.7 million Euros. Most of these sanctions had small effects on the budgets of the teams involved, let alone the budget of managers presumably responsible of these episodes. The only ones to pay were *de facto* the supporters of the teams damaged by the corruption events, who found their favored team relegated to the second division.

With the exception of Reggina, all the teams involved in the corruption episodes had owners with some direct or indirect control over the media (see Table 3) and held a significant portion of TV rights. This media power was also exerted within the most popular TV shows on soccer, like the *Processo del Lunedì*, *Domenica Sportiva* and *Controcampo*, whose journalists were more or less explicitly representing the interests of specific teams, just like in a coalition government.

Team	*Owner*	*Main Activities*	*Media Control (direct)*	*Media Control (indirect)*	*Share of Tv Rights (%)*
Atalanta	Ivan Ruggeri	Real Estate	Local Newspaper (*Giornale di Bergamo*)	no	2
Bologna	Alfredo Cazzola	Exposition	no	no	7
Brescia	Luigi Corioni	Manifacturing	no	no	2
Cagliari	Massimo Cellino	Chemicals	no	no	1
Chievo	Luca Cempedelli	Food	no	no	2
Fiorentina	Diego Della Valle	Textile, Media	National Newspaper (*RCS*)	no	5
Inter	Massimo Moratti-Pirelli	Oil, TLC, Tyres, Media	National Newspaper (*RCS*) and TV (*la 7*)	yes	13
Juventus	Agnelli's Family	Automobile, Insurance, Media	National Newspaper (*La Stampa*)	yes	13
Lazio	Claudio Lotito	Real Estate	no	yes (Radios)	8
Lecce	Giovanni Semeraro	Bank	Local Newspapers (Lecce)	no	3
Livorno	Franco Spinelli	Transportation	no	no	2
Messina	Pietro Franza	Transportation	no	no	2
A.C. Milan	Silvio Berlusconi	Media	National TV (*Mediaset*) and Newspapers	yes	13
Palermo	Maurizio Zamparini	Real Estate	no	no	4
Parma	under judicial control	Food	no	no	7
Reggina	Pasquale Foti	n.a.	no	no	3
Roma	Francesco Sensi	Transportation	no	yes (Radios)	8
Sampdoria	Garrone	Oil	no	no	1
Siena	Paolo de Luca	Transportation	no	no	2
Udinese	Giampaolo Pozzo	Manifacturing	no	no	2

Table 3: Team owners and Media (Championship 2004/5)

3 A Toy Model of Match Rigging

The history of *Calciopoli* summarized above suggests that matches were rigged in Italy mainly by affecting the choice of the referees rather than by directly corrupting the referees. Pressure on the referees was made via a variety of methods acting on their career concerns. Referees not cooperating could be excluded from griglia A, reducing their chance to be involved in international competitions, have their reputation seriously reduced by a journalist in a popular show, highlighting all the mistakes of the referee. The important role played by media power in match rigging is also highlighted by the fact that several top sport journalists were involved in these episodes with the task of criticizing on TV the referees not cooperating.

In this section we develop a simple model of corruption and market concentration in media as well as competitive power. Although the two sources of power are closely interrelated in practice, it is preferable to treat them separately in order to evaluate their role in isolation.

In line with most of the literature on professional sports we shall assume that decisions on rigging are made by a team manager. This role of the manager is consistent with the crucial importance they have in Italian soccer. In most big teams, the manager is often more important than the coach, as it is responsible for the purchase of new players. This role of managers in Italian soccer is consistent with their potential involvement in match rigging. In other words, managers are hired also because they can do the "dirty jobs". There is no principal-agent problem as we posit that the objectives of the owner and the manager are the same.[5] This assumption draws also from the fact that monitoring in this industry is regular (it occurs once a week) while performance is easy to measure.

The manager is risk neutral and maximizes her profits, subject to a budget constraint. One may interpret this budget constraint as the amount of money that the manager is allowed to dispense to the official selecting the referees and the journalists or as a stock of "blackmail capital", that is, a maximum level of pressure that the manager can exert on the journalists, the official selecting the referees and the referees themselves (their carreer concerns imply that they cannot rig any single match they are involved in). This constraint is needed as a boundary condition. To keep things simple, we shall assume henceforth that the constraint holds as a maximum number of matches to be rigged and that each manager can at most rig one match per Championship so that the choice of whether or not to rig a match can be modelled as a simple binary choice.

In our simple setup, the choice set of the manager is then confined to allocating this blackmail capital to the game offering the highest net rewards from rigging. Importantly, the choice set is not confined to the match played by the own team of the manager, as there may be

[5]Differently from the UK, in Italy the team manager does not coincide with the coach. In our model we do not consider any direct role of the coach in match rigging as there is no evidence of any specific role played by them in the tapped phone conversations.

larger net benefits from rigging a match involving a rival team. Both the media power and the competitive strength of the team are considered exogenous for the manager. The media power is given at the beginning of each Championship while the competitive strength of the team is related to the "market" in the previous year as well as by other factors (injuries) that are outside the within-year planning of the team. Thus, also win, draw and loss probabilities in a fair match are exogenous for the manager. As it was discussed above, the manager can influence decisions on red and yellow cards to the players of the various teams, by blackmailing the referees. But the decision about match rigging is made taking into account the competitive strenght of the team in the scenario *without* red and yellow cards imposed by a friendly referee.

3.1 One Team with a Mogging Manager

In our baseline model there is just one team-manager endowed with media power, hence blackmail capital, sufficient to rig a match. We will call this manager the "mogging" manager and the persons being corrupted (those selecting the referees or the journalists) "mogged" agents. We will later on consider cases where there is competition among managers also in match rigging.

Denote win (loss) probabilities in fair games as π_{ij}^{W} and π_{ij}^{L}, where subscripts denote the teams involved in the match. The odds after the match has been rigged are denoted instead by $\overset{\frown}{\pi}_{ij}^{W}$ and $\overset{\frown}{\pi}_{ij}^{L}$. The payoffs from the corruption activity are given by the changes in the asset values of a team associated with the desired outcome of a match. These changes in asset values will clearly depend on the options offered by rigged matches to run for the victory in the Championship, for a slot in the Champions League or in the UEFA Cup, taking into account not only of the performance of the own team of the manager, but also of the performance of the other teams. To keep things simple, we initially take these values as given. Later on, we will endogenize them.

The expected match rigging costs are supposed to be increasing in the deviation of the odds of a match from its fair values. There are various reasons for this assumption. First, these costs will embody some expected fine in case the corruption is detected, and the detection probability is likely to be related to the "size" of the manipulation of the outcome: intuitively, it is easier to detect match rigging when it alters the outcome of a match, expected to have a completely different outcome. Second, they can include compensations to mogged agents, whose reputation is also likely to be negatively affected by the change in win probabilities that they implement with respect to a fair game. In particular, we specialize the expected costs from mogging, as follows:

$$C_{ij} = 2\int_{\pi_{ij}^{W}}^{\overset{\smile}{\pi}_{ij}^{W}} (\overset{\smile}{\pi}_{ij}^{W} - \pi_{ij}^{W})d\overset{\smile}{\pi}_{ij}^{W} = (\overset{\smile}{\pi}_{ij}^{W} - \pi_{ij}^{W})^2 \tag{1}$$

Thus expected match rigging costs are increasing and convex in the deviation of the outcome of the rigged match from the expected outcome of a fair match.

Denote by i the team whose outcome is improved by match rigging and by j the rival team in that match. Assume again, for simplicity, that corruption affects only the win-loss probabilities while leaving unaffected the probability of a draw.[6] The payoff from rigging a match is therefore given by

$$W^{W}\overset{\smile}{\pi}_{ij}^{W} + (1 - \overset{\smile}{\pi}_{ij}^{W})\ W^{L} - \Delta_{ij}^{2} \tag{2}$$

where W refers to the team of the mogging manager (which can be different from the team "i") and $\Delta_{ij} = (\overset{\smile}{\pi}_{ij}^{W} - \pi_{ij}^{W})$ denotes the variation in the win probability associated with match rigging. The expected payoff from a fair match is insted given by

$$W^{W}\pi_{ij}^{W} + (1 - \pi_{ij}^{W})\ W^{L} \tag{3}$$

Thus a necessary condition for match rigging is that

$$\Delta_{ij}(\Delta W - \Delta_{ij}) > 0 \tag{4}$$

where $\Delta W = W^{W} - W^{L}$.

The above makes it clear that rigged matches must have large effects on the asset value of the mogging team: crucial games are targeted by corruption. Holding the variation in the asset value of the team constant, the variations in the win probabilities associated with match rigging should not be too large; if the manipulation of the result is too large, then the (flow) costs may be as large as to exceed the gross benefits of corruption on the asset value of a team. Put it another way, when there is just one team with media power, match rigging should target crucial and rather balanced games, that is, games in which it is possible to improve the outcome in a way which is favorable to the team of the manager without altering too much (or too visibly) the odds. To give an example, rather than trying to affect the outcome of a match of a rival team (in the competition for the final victory in the League) with a weak team, it is better to allocate blackmail capital to a match involving the rival team and a team with competitive power comparable to that of the rival team.

[6] This is an easily altered assumption. It is justified by the fact that, after the introduction of the 3-points victory, a draw is not much different from a loss in terms of the position in the ranking. Quoting the experienced Italian coach Carlo Mazzone, "*Before this rule, the draw was a half victory, from now on it will be a half loss.*"

3.2 Two Teams with Mogging Managers

Consider now the case in which there are two mogging managers endowed with an equal stock of blackmail capital, sufficient to rig just one match per manager. In this case, we need to explicitly consider the preferences of the common (corrupted) agent. This common agent (team selecting the referees, journalists and referees themselves) is blackmailed by the two mogging managers and, when the latter compete against each other, may succeed in extracting some surplus, in terms of a compensating transfer from them. Put it another way, the mogging manager cannot any longer make a "take it or leave" offer to the mogged agent. Suppose that also these agents are risk-neutral so that the objective function of the common agent is given by

$$V = F_k - r(\Delta_{ij}^2)$$

where F denotes the "fee" paid by a mogging agent (indexed by k) and $r(.)$, $r' > 0$ is a reputation cost paid by the mogged agent whenever he alters the outcome of a match. The fee paid by the mogging team can be interpreted, as in the theory of lobbies (Grossman and Helpman (2001)), as a *compensating contribution function*. In the case of a single mogging manager, this compensating contribution allows the agent to remain on the same indifference curve as in the absence of match rigging, as the mogging managers extracts all surplus from match rigging.

In presence of two mogging managers, knowing each other, but allowed only to make conjectures as to the contributions offered by the other manager, there are two relevant scenarios to be discussed. The first is when the two managers do not compete for the same target, e.g., only one of two teams is running for the victory in the national League or for a slot in the Champions League whilst the other is trying to avoid the relegation. The other scenario is one where the two teams with mogging managers are competing against each other for the same target, e.g., the victory in the Championship.

To simplify our discussion, we will simply take the case where $r(\Delta_{ij}^2) = \Delta_{ij}^2$ and the detection probability is zero so that mogging costs coincide with the compensating contribution to the mogged agent.

3.2.1 The Two Teams are not Competing

In the case where the two teams with mogging managers do not compete against each other, then three types of matches have to be considered.

First there are matches affecting the asset value of only one of the two teams with media power. Define these games as "type-M" games. As the absence of a competing contribution, assigns a monopoly power to the mogging manager, we go back to eq.(1) and the mogging manager extracts all surplus from the mogged agents, i.e.:

$$C_{ij}^{M} = \Delta_{ij}^{2}$$

Secondly, there are matches in which both teams with media power have a vested interest in affecting the outcome, but in opposite directions. Define these games as "type-O" games. For instance, the first manager would like team "i" to win, whilst the second manager would like team "j" to win. As shown by Bernheim and Whinston (1998), appropriate restrictions to the contribution function must be made to rule out multiple equilibria in this case. In particular, one needs to assume that the contributions are chosen from the class of continuous (this restriction is consistent with each mogging manager making conjectures as to the contributions offered by the other manager), and differentiable functions. This unique equilibrium contribution must not only compensate the common agent for the reputation loss (1), but also should "beat" the offer of the competing manager. Thus, in this case

$$C_{ij}^{O} > \Delta_{ij}^{2}$$

Finally both teams with blackmail capital may be interested in affecting the outcome of a match in the same direction, say both want team "i" to win. Define these game as "type-L" games. In this case, the equilibrium contribution cannot be larger than in the monopoly case. It will be lower if both teams receive a relatively high payoff from rigging that match, so that the other manager expect them to make a non-zero contribution to match rigging. In other words

$$C_{ij}^{L} \leq \Delta_{ij}^{2}$$

3.2.2 Competing Teams with Media Power

Relax now the one-match constraint in the endowment of blackmail capital. Mogging managers are allowed to rig as many games as they wish up to exhausting this endowment, denoted by B, that is

$$\sum_{ij} C_{ij} \leq B \tag{5}$$

Mogging managers will allocate this capital by choosing the games with the highest net payoffs from rigging. Per any given revenue from rigging, this choice will be a cost minimizing one.

When only matches type-O are present, as discussed above, corruption is more costly. Hence, there will be less rigged matches in this case per any given endowment of B.

Importantly, competition only in media power does not necessarily make corruption more costly, as there can be some crucial type-L games, allowing mogging managers to save some blackmail capital with respect to the scenario where there is a monopoly of media power.

Put it another way, the costs of match rigging increase together with competition on both media power and competitive power. At the same time, competition only in competitive-sport power does not necessarily increase match rigging costs insofar as a monopoly in match rigging allows mogging managers to extract all surplus from the corrupted agents.

This result can be extended to the case of 3 or more teams with media power. The key implication is once again that less concentration in media power does not, by itself, reduce corruption. By the same token, more competitive balance without less concentration in media power does not imply less match rigging as actually more competitive balance reduces the costs of match rigging (odds only need to be slightly adjusted to obtain some desired outcomes). An additional reason for a competitive balance not to discourage match rigging comes from the revenue side, as discussed below.

3.3 Endogenizing the (Gross) Payoffs

So far we have taken the asset values of teams, W as given. Let us now derive them.

A main source of revenues of teams in professional soccer after the introduction of the pay-TV is represented by TV rights.[7] In Italy TV rights account for roughly 2/3 of total revenues of soccer teams. We can model the demand for TV events, as an "adjusted" winner-takes-all demand function, as sport fans highly value talents, who are in short supply. Only teams with some soccer star, type-S teams, attract a significant audience, going beyond the close supporters of the club. At the same time, consumers of TV events love competitive balance, hence the demand of sport events is negatively affected by a competitive deficit.

Define the inverse demand for each single match i, j as $p(\pi_{ij})$. We specialize this match-specific demand as follows

$$p(\pi_{ij}^W) = \begin{cases} \pi_{ij}^W \left(1 - \pi_{ij}^W\right) & \text{if } i \text{ or } j \text{ are type-s teams} \\ 0 & \text{otherwise} \end{cases}$$

which captures the idea that most preferred matches are those balanced and with talented players.

Only teams in the first division have stars. TV rights are assigned to the teams with stars, a necessary condition to yield some revenue, at the beginning of the championship. Part of these revenues is then split across teams without stars according to a given sharing rule assigning a fraction $0 < \phi < 1$ of the revenues from the "virtual stadium" also to the teams without stars (e.g., in Italy this share has been for several years about 18%). Thus, for a team i without stars, the payoff from remaining in the first division (ΔW^A) rather than being

[7] See, for example, Szymanski and Kuypers (1999) and Deloitte and Touche (2003).

relegated (whose asset value is normalized to zero) is given by

$$\Delta W^A = 2\phi \sum_{i\in S} \sum_{j\neq i} (\pi_{ij}^w * (1-\pi_{ij}^w)) \tag{6}$$

Soccer stars play only in teams involved in the Champions League as this makes them eligible to international rankings increasing their market value, hence bargaining power with respect to the owner of the team (after the Bosman ruling). Thus, the payoff, in terms of national TV rights,[8] of a team qualifying for the Champions League (ΔW^C) rather than simply remaining in the first division is given by

$$\Delta W^C = 2\left[(1-\phi)\sum_{j\neq i}(\pi_{ij}^w * (1-\pi_{ij}^w)) - \phi \sum_{i\in S}\sum_{j\neq i}(\pi_{ij}^w * (1-\pi_{ij}^w))\right]$$

which highlights the capital loss associated with not qualifying for the Champions League and suggests that the greater is the sharing rule, the lower is the payoff.

Finally, there may be an extra payoff from winning the national League even when victory in the League is not required for access to the Champions League (in the Italian League the first four teams qualify for the Champions League). As the final victory in the national championship has more a reputational effect, than an effect on national TV rights (which are strictly related to the audience of the team), we shall model this event as yielding a fixed increase in the asset value of the team, ΔW^L.

We have not considered the case of a team without stars qualifying for the Champion League. The latter obtains the option to buy soccer stars, but has still to make the required investment. Hence, we may consider that the primary target of a team without stars is simply to remain in the first division or, if in a position to do so, to compete for the final victory in the Championship in order to get the reputational premium (which is independent of future investiments).

Summarizing, the target in the national League of the teams with soccer stars is a slot in the following year Champions League whilst the target of the other teams is primarily remaining in the national League. These payoffs are increasing in the competitive balance in the national League. More competition involves potentially higher net payoffs from attaining the targets, but more concentration in sport power makes it easier for some teams to attain their targets. The TV rights sharing rule is also important in affecting the payoffs from fair and rigged matches. All teams, independently of the presence of soccer stars, may target the victory in the national League, which is independent of the competitive balances as it carries with it mainly a reputational effect.

[8]We ignore TV rights related to Champions League events as this would require modelling also the competitive balance at the supra-national level, something with is beyond the scope of this paper.

TV rights, allocation rules and reputational premia associated with the victory in the National League are all defined at the beginning of the Championship, based on an assessment of the competitive balance, hence of the demand for pay-TV. What evolves over time is the relative strength of the various teams and the probability associated to attaining any of these targets.

3.4 The Option Value of Waiting

Unlike TV rights, decisions about match rigging need not to be made once and far all at the beginning of the Championship, based on a given distribution of the competitive strength. Mogging managers may decide not only *whether*, but also *when* to rig a match, knowing that the victory in some games may not be essential to attain their target. Only *marginal teams*, that is, close to the threshold position in the ranking allowing to compete for any of three targets characterised above are in a position to potentially benefit from match rigging. The uncertainty associated with sport events and the fact that match rigging is more convenient for marginal teams create an option value of waiting before rigging a match. A mogging manager intervening too early in the Championship may well waste blackmail capital as it turns out that the target can be attained (or cannot be attained) even without (with) match rigging.

The option value of waiting before rigging a match can be easily characterized. To keep things simple, we ignore discounting and assume that blackmail capital allows only to rig one match and that there are only two days to go before the end of the tournament. At this late stage, there will be a well defined threshold score in the ranking ensuring victory in the League (yielding ΔW^L), access to the Champions League (ΔW^C) or survival in the first division (ΔW^A). Consider now the problem of a *marginal* team (one lacking 3 points in the ranking to attain its target) having to meet a team without media power (and not competing with a team with media power) and with the same competitive strenght in the last two games of the League. The reward from rigging the match at the last but one game are given by

$$\tilde{\pi}_{ij}^W \Delta W_i - C_{ij} + (1 - \tilde{\pi}_{ij}^W)\pi_{ij}^W \Delta W_i \tag{7}$$

whilst the rewards from playing a fair match and, if needed, rigging the last game are given by

$$\pi_{ij}^W \Delta W_i + (1 - \pi_{ij}^W)(\tilde{\pi}_{ij}^W \Delta W_i - C_{ij}) \tag{8}$$

It is easy to check that (8) yields a higher expected payoff than (7) because it involves potential savings ($\pi_{ij}^W C_{ij}$) in match rigging costs.

Waiting may not be an optimal strategy if the final match involves a stronger team or a match for which there is a conflict of interests with another team with media power. Moreover, match rigging should target rather balanced matches, due to the convexity of match rigging costs, and towards the end of the Championship a number of factors (including the motivation of teams having already attained or missed their targets) may militate against this requirement.

Thus, even if there is in principle an obvious case for waiting before rigging a match towards the end of the Championship, mogging managers may be forced to intervene before getting too close to the final days of the tournament. It is likely that match rigging is concentrated in the second half of the Championship, but not in the final games of the tournament.

3.5 Empirical Implications of the Model

We can now summarize the main predictions of our model in a way as to guide the following empirical analysis.

Prediction 1: Concentration and corruption. Match rigging is larger the more concentrated is media power. Less concentration in competitive power actually increases the net payoffs from rigging matches when there is only one team-manager rigging matches. When there is some competition in match rigging, less concentration in competitive power has ambiguous effects on corruption, as it increases both the costs and the revenues from match rigging.

Prediction 2: Marginal teams. Conditioning on a given distribution of media power, match rigging should preferably involve marginal teams, that is teams close to attaining their target.

Prediction 3: The timing of corruption. Match rigging should be more frequent in the second half of the Championship although not necessarily in the last few days of the tournament.

Prediction 4: The nature of rigged matches. Matches targeted for corruption should be rather balanced, so that a change in the expected outcome does not require a large variation in the win probability with respect to a fair match.

4 A Closer Look at *Calciopoli*

4.1 The Determinants of Match Rigging

In order to identify potential corruption episodes in the Italian First Division and test the above implications of the model, we proceed in three steps. First, we use information on *documented* (via tapped phone conversations) corruption episodes in the 2004-5 Championship to analyse the criteria being followed by mogging managers in that season in allocating their "blackmail capital". This procedure is useful to gain insights as to match rigging and on the characteristics of corruption episodes. Next, in our second step, we predict fair outcomes in the 2004-5 Championship (where we can also check the orthogonality of our predictions with outcomes of matches under judicial inquiry) as well as in previous Championships. Third, based on the results of step one and two, we identify potential corruption episodes also in Championships in which there is no ongoing judicial inquiry.

We estimate the probability of match rigging, the dependent variable being the matches under investigation in the Championship 2004-5, based on the following probit model:

$$CORRUPTION_{ijt} = \beta_1 CR_{t-1} + \beta_2 MATCH\_DAY_t + \beta_3 MEDIA\_POWER_{ijt} + \beta_4 PRECLUSIONS\_GRID\_ijt + \tag{9}$$

$$\beta_5 EXP\_REFEREE_t + \beta_6 ELITE\_REFEREE_t + \epsilon_{ijt}$$

where the subscripts i, j and t denote, respectively, the home team, the away team and the day of the Championship, the dependent variable $CORRUPTION_{it}$ is a binary variable taking value 1 if the match is under investigation and 0 otherwise; CR is the competitive power measured, as in Koning (2000), by the sum of the number of points obtained by the top K teams, relative to the potential number of points they could have gained, i.e.

$$CR_K = \frac{\sum_{k=1}^{K}}{P_{(k)}} KW\,(2J - K - 1) \tag{10}$$

where $W = 3$, i.e. the points in case of victory, and J is the number of competing teams, $MEDIA\_POWER$ is a vector representing the different types of Media Power of the weak teams, defined in terms of a twofold taxonomy (Strong[9] or Weak media power), based on information on TV rights and media ownership, against Juventus, A.C. Milan and Inter, $MATCH\_DAY_t$ represents the number of days before the beginning of the Championship, entering in some specifications with the quadratic form.

$PRECLUSIONS\_GRID$ is a dummy variable controlling the grid for the most important matches with more than one preclusion, $EXP\_REFEREE_t$ and $ELITE\_REFEREE_t$ are the spline of the number of matches refereed during the previous championship by experienced referees (i.e. between 10 and 18 matches, i.e. between the median and the 90th percentile) and elite referees (i.e. with more 18 matches refereed, i.e. corresponding to the 90th percentile) respectively. These variables aim at capturing the blackmail capital, i.e., the influence that mogging managers could exert on referees because of their career concerns. Elite referees can be less subject to this type of pressures as they already have reached an international standing, while experienced referee are about to get a promotion to international competitions. Preclusion grid measures the importance of the match to be refereed (grid A collects the most important games) and the discretion of the officials in choosing the referees.

[9]The following teams are considered endowed with strong media power: Fiorentina, Inter, Juventus, Lazio Rome, A.C. Milan and Roma.

Our model predicts that corruption will rarely target matches involving the team of the mogging manager and another team with media power, as clearly in this case there is a conflict of interests and match rigging can be very costly. We include dummies denoting matches involving, on the one hand, a given team with media power, and, on the other hand, teams with weak (e.g., only local) media power. These dummies, however, are not significant.

The panel structure of the data allow us to estimate a stacked probit with separate intercepts for each match day and team, and interpret the results as discrete-time proportional hazards. In these cases, as shown by Kloek (1981), random effect estimators are generally more efficient than OLS estimators. Thus, we implement a cluster-specific random effect model, where clusters are defined as observations concerning the same team, assuming independence across clusters. Table 4 displays the results of this robust cluster probit estimation, where corruption is defined based on all the matches under investigation in the Championship 2004-5.

We provide two different specifications. In the first specification (whose results are displayed in the first four columns), competitive power is contemporaneous to the match. In the second specification, competitive power is measured before the match, embodying the idea that match rigging makes the outcome more predictable for the mogging manager.

The regression results suggest that the competitive power of Juventus is significant and always *negative*, the competitive power of A.C. Milan is *positive* when it is statistically significant, while the competitive power of Inter is never significant. In other words, at times in which Juventus is stronger, there is less corruption activity and vice versa. While the probability of match rigging, if anything, increases with the competitive power of A.C.Milan. This provides some support to the view that match rigging was targeted to support Juventus against A.C. Milan. The competitive strength of Inter is, instead, irrelevant. This is not surprising as Inter in that Championship was soon out of the game as far as the final victory in the League was concerned. Another important result is that the probability of rigging a match is increasing at a decreasing rate with time. This is consistent with the model prediction that match rigging is concentrated in the second half of the Championship, but occur before the final days of the tournament. The probability of match rigging is higher when Juventus plays with a team not having media power ($WEAK\_JUVENTUS$), according to the dummy variable characterised above. This is also consistent with the theoretical predictions. We tried also with dummies interacting A.C. Milan and Inter with the media power of their challengers, but these were never significant (and are not reported herein for brevity).

	(1)	(2)	(3)	(4)	(5) ($CR\_$ att -1)	(6) ($CR\_$ att -1)
CR_JUVENTUS	-0.064***	-0.028**	-0.024*	-0.024*	-2.246*	-2.541**
	(0.015)	(0.009)	(0.011)	(0.011)	(0.905)	(0.517)
CR_MILAN	7.664***	0.704	0.696	0.541	-1.098	
	(1.958)	(1.542)	(1.524)	(1.577)	(1.371)	
CR_INTER			-0.904	-0.623	1.280	
			(1.657)	(1.682)	(1.850)	
WEAK_JUVENTUS				0.738*	0.771*	
				(0.362)	(0.351)	
MATCH_DAY	-0.056***	0.157**	0.167*	0.169*	0.180***	0.152
	(0.011)	(0.057)	(0.066)	(0.068)	(0.053)	(0.050)
$MATCH\_DAY^2$		-0.005***	-0.005***	-0.005**	-0.005***	-0.005**
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EXP_REFEREE						0.062*
						(0.030)
ELITE_REFEREE						-0.558*
						(0.240)
PRECLUSIONS_GRID						0.523**
						(0.145)
N. Obs.	380	380	380	380	370	370
Log pseudolikelihood	-179.46	-171.41	-171.32	-169.90	-170.14	-159.28
χ^2	141.444	105.677	116.479	112.176	98.85	195.789
$*p < 0.05$, $**p < 0.01$, $***p < 0.001$						

Table 4: Estimation results : Robust Cluster Probit. Corruption as dependent variable.

We work with 380 observations (370 when using competitive power lagged one period). In order to increase degrees of freedom, we perform 1,000 Monte Carlo simulations.[10] Figures 1 and 2 display the mean (as well as the upper, 95%, and lower, 5%, confidence intervals) of the estimated match rigging probability as a function of the competitive power of Juventus and of the days of the Championship, respectively. This shows that the probability of match rigging is steeply decreasing in the competitive power of Juventus and corruption activity is likely to be concentrated between the 10th and the 30th days of the Championship.

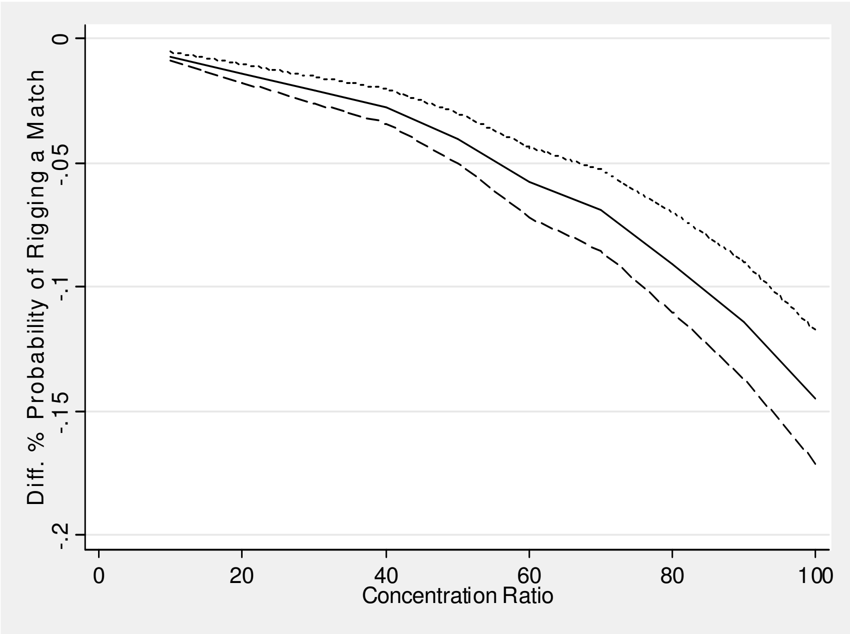

Figure 1: Δ% Probability of Rigging a Match and Juventus Competitive Power (Concentration Ratio) according to the Model Estimates

4.2 Predicting Fair Outcomes in the 2004-5 Championship

The above results are consistent with a negative relationship between concentration of competitive power and match rigging and a positive one between media concentration and corruption. In order to analyze the time variation of media power, we need to extend our analysis to other Championships, capturing some variation in media power, changing at relatively low frequencies. Our next step is therefore to identify potential corruption episodes in the 2002-2004 period. We combine two criteria to identify the matches likely to have been rigged: i. they must have a high probability of match rigging in light of the above regression results, and ii. they must be outliers in a regression predicting fair outcomes. This section is devoted to

[10] The Monte Carlo simulation has been performed using the Clarify program and the procedure explained in Tomz, Wittenberg, and King (2001).

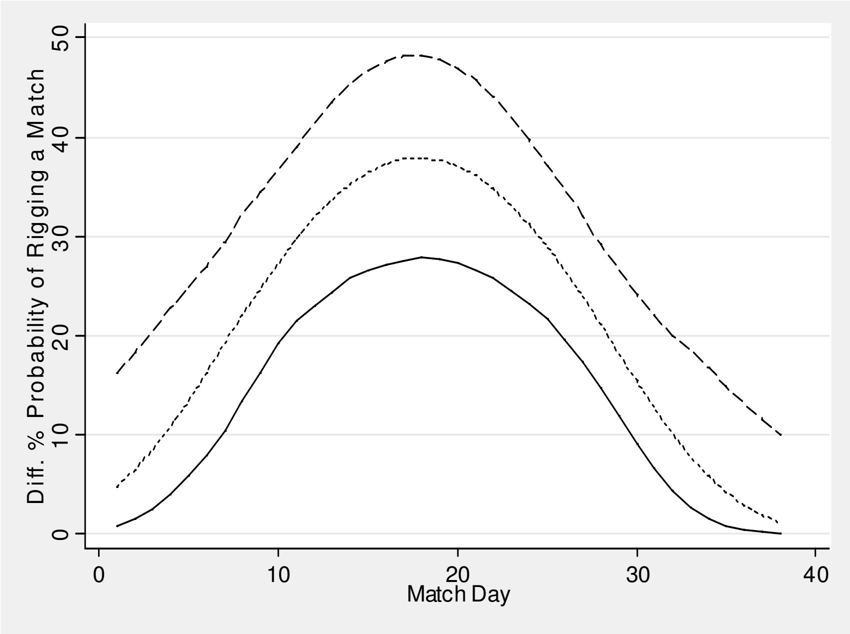

Figure 2: Δ% Probability of Rigging a Match and Match Day according to the Model Estimates

finding an appropriate specification of the latter equation. The literature on soccer economics typically uses and ordered probit or a bivariate Poisson model to predict the outcomes of a match. We draw on the model introduced by Koning (2000) and developed by Dobson and Goddard (2001) to analyze the competitive balance and uncertainty of outcomes. Due to the presence of sample selection generated by the presence of promotions and relegations, their model aims at disentangling the effect of team heterogeneity from the conditional win probabilities. The dependent variable is represented by the points P_i obtained by home team i, normalized in terms of victory (represented by value 1), draws (equalized to 0.5) or losses (identified by 0). Moreover, we define an unobserved index function for the points P^* as

$$P^* = X\beta + \epsilon \tag{11}$$

where X are the set of explanatory variables and assume

$$P_i = \begin{cases} 1 & \text{if } \mu_2 < \sigma_i - \sigma_j + \epsilon_{i,j} \\ 0.5 & \text{if } \mu_1 < \sigma_i - \sigma_j + \epsilon_{i,j} < \mu_2 \\ 0 & \text{if } \sigma_i - \sigma_j + \epsilon_{i,j} < \mu_1 \end{cases}$$

where μ_1 and μ_2, the cut points, are parameters controlling the proportions of victory, draws and losses. The values of $\epsilon_{i,j}$ are simulated from a normal distribution. Thus, the conditional probabilities can be defined as

$$\begin{cases} \text{Win probability} & Pr\left(\epsilon_{i,j} > \hat{\mu_2} - \hat{P}^*_{i,j}\right) = 1 - \Phi\left(\hat{\mu}_2 - \hat{P}^*_{i,j}\right) \\ \text{Draw probability} & Pr\left(\hat{\mu}_1 - \hat{P}^*_{i,j} < \epsilon_{i,j} < \hat{\mu}_2 - \hat{P}^*_{i,j}\right) = \Phi\left(\hat{\mu}_2 - \hat{P}^*_{i,j}\right) - \Phi\left(\hat{\mu}_1 - \hat{P}^*_{i,j}\right) \\ \text{Loss probability} & Pr\left(\epsilon_{i,j} < \hat{\mu}_1 - \hat{P}^*_{i,j}\right) = \Phi\left(\hat{\mu}_1 - \hat{P}^*_{i,j}\right) \end{cases}$$

Our specification for $P^*_{i,j}$ is similar to the regression framework developed by Koning (2000), but differs from that in two main respects. First, in his model, Koning (2000) considers whether the match has championships, promotion or relegation significance for one team and not for the other.[11] We believe that this assumption is too subjective, and prefer to use instead the variables proposed by Kuypers (2000) in his investigation of the fixed odds betting market, notably the cumulative points for the season, the goals scored and the teams' position in the overall ranking. Moreover, we consider only first division data as we believe that measurement errors are more serious in the second division.
The econometric equation defining $P^*_{i,j}$ is therefore specified as follows

$$\begin{aligned} P_{i,t} &= \textstyle\sum_{y=2002}^{2004} \alpha_y AVERAGEP\_HOME_{y,t-1} + \sum_{y=2002}^{2004} \beta_y AVERAGEP\_VISITOR_{y,t-1} + \\ &+ \textstyle\sum_{y=1}^{2} \gamma_y LAST\_HOME_{2004,y} + \sum_{y=1}^{2} \delta_y LAST\_VISITOR_{2004,y} + \lambda_1 MATCH\_DAY + \\ &\lambda_2 MATCH\_DAY^2 + \lambda_3 DIFFERENCE\_RANKING_{2004,t-1} + \lambda_4 NEW\_PROMOTED_{2004,t} \\ &\lambda_5 C.LEAGUE_{2004,t} + \lambda_6 GOAL\_HOME_{2004,t-1} + \lambda_7 GOAL\_VISITOR_{2004,t-1} + \\ &+ \lambda_8 NEW\_COACH\_HOME_{2004,t} + \lambda_9 IVS\_HOME_{2004,t} + \lambda_{10} CORRUPTED\_HOME_{2004,t} + \\ &+ \lambda_{11} CORRUPTED\_VISITOR_{2004,t} + \epsilon_{i,t} \end{aligned}$$

where the suffixes $HOME$ and $VISITOR$ denote the home and the visitor team respectively, the subscripts the Championship the measure is refereed to, $AVERAGEP$ stands for the fraction of potential points gone to any of the two teams (i.e. it is a sort of *win ratio*) in previous matches opposing the same teams (either in the same Championship or in previous ones), $LAST$ denotes the last performance in the match, $NEW\_PROMOTED$ is a dummy which is equal to 1 whether the team has just been promoted to the first division, $DIFFERENCE\_RANKING$ denotes the difference in the position in the ranking of the home team vis-a-vis the visitor, $GOAL$ are the average goals scored in the last game before the match considered by the dependent variable and $CLEAGUE\_HOME$ is a dummy variable taking the value one when the home team is competing for qualifying to the Champions League and zero otherwise, and NEW_COACH_HOME is another dummy taking value one

[11] They also choose a "crude" algorithm to identify the matches with significances to one or both teams during the last few weeks of the season.

when the home team has a new coach.[12] Finally IVS_HOME is the measure of competitive strength discussed in section 2.2. We initially concentrate our analysis on the Championship 2004-2005 as it allows us to evaluate the appropriateness of our strategy to identify rigged matches as outliers. Table 5 provides our estimates.

The results are fairly encouraging. As pointed out by the last but one row of table 5, we succeed in predicting up to 70% of the results, while ordered probit models used by the best specifications offered by the literature on soccer economics do not successfully predict more than 4 matches out of 10. The sign of the coefficients is broadly in line with a priori expectations (except $AVERAGEP\_VISITOR_{2002}$), although many variables are not statistically significant (possibly because of multicollinearity). More importantly, when we include in the regressions dummies capturing matches that are under judicial inquiry as they favored the home team ($CORRUPTED\_HOME$) or the visitor ($CORRUPTED\_VISITOR$), these variables turn out to be statistically significant. Furthermore, their inclusion in the model increases the predictive capacity of our model (from 55 to 57%, from 55 to 59% or from 68 to 70% depending on the specification) while it declines the percentage of outliers of our regression corresponding to matches under judicial inquiry. Outliers are identified, in this context, according to a very restrictive criterion. Matches with most likely outcomes according to our regression are considered as outliers. We interpret these results as broadly supportive of our empirical strategy.

Our second specification is based on a bivariate Poisson regression of the goals scored and conceded. In particular, we assume that, $GOAL^H_{i,o}(= GOAL^V_{j,o})$ and $GOAL^H_{i,o}(= GOAL^V_{j,o})$ follow the joint probability function P:

$$P\left(GOAL^H_{i,o} = s, GOAL^V_{i,o} = c\right) = exp\left(-\lambda_{1,i,j} - \lambda_{2,i,j} + \lambda_{3,i,j}\right) \sum_{k=0}^{min(s,c)} \frac{\lambda^{s-k}_{1,i,j}\lambda^{c-k}_{2,i,j}\lambda^{k}_{3,i,j}}{(s-k)!\,(c-k)!k!} \quad (12)$$

where $\lambda_{1,i,j}$ and $\lambda_{2,i,j}$ denote the expected number of goals scored by the home and the away teams respectively. The covariates are the same as in Table 5 except for the dummies ($FRIENDS$) and ($ENEMIES$) capturing teams with whom there are traditional links among supporters or boards (see Appendix 2). The results are displayed in Table 6, and are presented separately for the home team and the visitor.

[12] All these data have been collected from the Italian State Television *Rai* website, http://www2.raisport.rai.it/mcalcio/, and from the Italian sport newspaper *Gazzetta dello Sport*, http://www.gazzetta.it/speciali/statistiche/2006/squadre/squadre.shtml.

Table 5: Estimation results : oprobit

Dependent Variable: *POINTS*	(1a) β/se	(1b) β/se	(2a) β/se	(2b) β/se	(3a) β/se	(3b) β/se
$AVERAGEP\_HOME_{2004,t-1}$	1.806	1.737	1.751	1.887	0.956	0.884
	(1.027)	(1.056)	(1.062)	(1.031)	(0.918)	(1.024)
$AVERAGEP\_VISITOR_{2004,t-1}$	-1.207***	-1.000**	-1.155**	-0.918*	1.317**	1.509**
	(0.352)	(0.358)	(0.376)	(0.376)	(0.451)	(0.475)
$AVERAGEP\_HOME_{2003,t-1}$	0.977	2.530	1.451	2.769	-0.734	1.348
	(3.256)	(3.125)	(3.491)	(3.178)	(4.278)	(4.469)
$AVERAGEP\_VISITOR_{2003,t-1}$	-3.610	-2.585	-3.636	-2.761	-5.134	-3.215
	(3.422)	(3.352)	(3.485)	(3.230)	(4.581)	(4.633)
$AVERAGEP\_HOME_{2002,t-1}$	4.798	4.633	5.598	4.659	0.636	0.825
	(3.827)	(3.815)	(3.828)	(3.792)	(3.493)	(3.530)
$AVERAGEP\_VISITOR_{2002,t-1}$	9.463**	9.964*	10.853**	10.145**	6.504*	7.566*
	(3.641)	(3.893)	(3.435)	(3.849)	(3.062)	(3.239)
$LAST\_HOME_{2004,t-1}$	0.102	0.018	0.126	0.020	0.168	0.089
	(0.153)	(0.175)	(0.148)	(0.172)	(0.145)	(0.148)
$LAST\_VISITOR_{2004,t-1}$	-0.166	-0.234	-0.135	-0.211	-0.060	-0.114
	(0.202)	(0.200)	(0.206)	(0.206)	(0.241)	(0.241)
$LAST\_HOME_{2004,t-2}$	-0.124	-0.045	-0.130	-0.054	-0.150	-0.052
	(0.091)	(0.094)	(0.084)	(0.089)	(0.167)	(0.150)
$LAST\_VISITOR_{2004,t-2}$	-0.064	-0.059	-0.072	-0.073	-0.212	-0.202
	(0.139)	(0.129)	(0.146)	(0.134)	(0.162)	(0.171)
$MATCH\_DAY$	0.049	0.060	0.042	0.064	0.025	0.035
	(0.028)	(0.033)	(0.033)	(0.034)	(0.035)	(0.039)
$MATCH\_DAY^2$	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$DIFFERENCE\_RANKING_{2004,t-1}$	-0.219***	-0.215***	-0.240***	-0.236***	-0.310***	-0.324***
	(0.054)	(0.064)	(0.049)	(0.063)	(0.062)	(0.090)
$NEW\_PROMOTED_{2004,t}$	-0.074	0.006	-0.072	0.005	-0.078	0.062
	(0.199)	(0.200)	(0.203)	(0.204)	(0.211)	(0.219)
$C.LEAGUE\_HOME_{2004,t}$	1.237***	1.234***	1.246***	1.239***	1.321***	1.450***
	(0.185)	(0.241)	(0.182)	(0.236)	(0.250)	(0.265)
$GOAL\_HOME_{2004,t-1}$			0.163		0.118	
			(0.254)		(0.214)	
$GOAL\_VISITOR_{2004,t-1}$			-0.029		0.025	
			(0.046)		(0.035)	
$NEW\_COACH\_HOME_{2004,t}$			-0.232	-0.226	-0.385**	-0.382**
			(0.141)	(0.135)	(0.121)	(0.120)
$IVS\_HOME_{2004,t}$					0.119***	0.126***
					(0.008)	(0.009)
$CORRUPTED\_HOME_{2004,t}$		0.830***		0.846***		1.126**
		(0.206)		(0.202)		(0.349)
$CORRUPTED\_VISITOR_{2004,t}$		-1.142***		-1.127***		-1.221***
		(0.190)		(0.190)		(0.301)
‘						
N.obs	330	330	330	330	320	320
Pseudo R^2	0.086	0.146	0.090	0.150	0.400	0.463
Log-likelihood	-319.711	-298.599	-318.225	-297.421	-203.759	-182.465
% Matches Forecasted	55 %	57 %	55%	59%	68 %	70%
% Outliers which are Rigged	46 %	33 %	46%	32%	38 %	19%

Table 6: Estimation results : poisson

Dependent Variable:	(1) $GOAL\_HOME$ β/se	(2) $GOAL\_VISITOR$ β/se	(3) $GOAL\_HOME$ β/se	(4) $GOAL\_VISITOR$ β/se
$MATCH\_DAY$	-0.014*	-0.010**	-0.013*	-0.008
	(0.006)	(0.004)	(0.006)	(0.004)
$FRIENDS$	-0.387	-0.111	-0.380	-0.083
	(0.227)	(0.148)	(0.230)	(0.144)
$ENEMIES$	-0.167	-0.123	-0.154	-0.079
	(0.157)	(0.069)	(0.160)	(0.067)
$VISITOR$	-0.024**		-0.021**	
	(0.008)		(0.007)	
$AVERAGEP\_HOME_{2004,t-1}$	-0.935*	-0.121	-0.896	0.160
	(0.473)	(0.381)	(0.472)	(0.398)
$AVERAGEP\_VISITOR_{2004,t-1}$	-0.361		-0.402	
	(0.677)		(0.686)	
$GOAL\_HOME_{2004,t-2}$	0.552***	-0.387***	0.626**	-0.118
	(0.080)	(0.055)	(0.205)	(0.147)
$GOAL\_VISITOR_{2004,t-2}$	0.229***	0.658***	0.235***	0.671***
	(0.031)	(0.036)	(0.035)	(0.038)
$NEW\_COACH\_HOME_{2004,t}$	0.156**	0.114*	0.157**	0.120*
	(0.057)	(0.057)	(0.057)	(0.056)
$NEW\_COACH\_VISITOR_{2004,t}$	0.065	0.069	0.060	0.060
	(0.064)	(0.070)	(0.061)	(0.071)
$HOME$		-0.018***		-0.015***
		(0.004)		(0.004)
$AVERAGEP\_VISITOR_{2004,t-1}$		-0.048		0.381
		(0.530)		(0.537)
IVS_HOME			-0.003	-0.009*
			(0.007)	(0.005)
N.obs	320.000	320.000	320.000	320.000
Log-likelihood	-436.540	-296.957	-436.396	-296.443
% Non Rigged Matches Forecasted	63 %		62 %	
% Outliers which are Rigged	37 %		37 %	

Lagged dependent variables seem to explain rather well the time variation in the number of goals scored. This suggests that persistence across matches, e.g. of the fisical fitness of key players, induces an autoregressive structure to the data. By combining the predicted goals of the home team and the visitor, we succeed in correctly predicting about 62-63% of the results. The percentage of outliers (defined as in the previous regression) which are under judicial inquiry is lower than in the ordered probit. Consistently with the soccer economics literature (Karlis and Ntzoufras (2003)), the Poisson model seems to perform worse than the ordered logit in predicting matches in Italian *Serie A*, as it underperforms for low-scoring draws.

Although the ordered probit turns out to provide more satisfactory estimates than the Poisson model, we can use the latter within a two-stage estimation procedure involving a larger number of observations. The literature on soccer economics suggests that the normal transformations (and even less so linear probability models) are not satisfactory when applied to more than one Championship. For this reason we use a *random effect unbalanced panel ordered logit* when dealing with the larger sample, in order to exploit also the time effects. Being our dependent variable $P^*_{i,j}$ ordinal and the rule given by (30), the conditional probability of observing $P_{i,j}$ is given by

$$\begin{aligned} Prob\left(P_{i,j}=k|X_{i,j}\right)=Prob\left(\mu_{k-1}\leq P^*_{i,j}\leq\mu_k\right)=\\ =Prob\left(\mu_{k-1}\leq\beta X+\epsilon_{it}\leq\mu_k\right)=\\ =Prob\left(\mu_{k-1}-\beta X\leq\epsilon_{it}\leq\mu_k-\beta X\right)=\\ =Prob\left(\epsilon_{it}\leq\mu_k-\beta X\right)-Prob\left(\epsilon_{it}\leq\mu_{k-1}-\beta X\right)\end{aligned}$$

We assume that ϵ_{it} follows a logistic distribution. This choice is motivated by different papers in the literature.[13] However, we have also checked the consistency of our model considering also a standard normal distribution and an ordered probit model and the results are still consistent. We assume that the individual specific effects are not correlated with the explanatory variables (as there is randomness in soccer events), so we can rewrite the previous equation as

$$P^*_{i,t}=X_{i,t}\beta+u_i+\epsilon_{i,t} \tag{13}$$

with

$$i=1,\ldots,N \tag{14}$$

and

$$t=1,\ldots,T \tag{15}$$

where u_i catches for the team i at week t the individual's time constant specific effect, assumed to be normally distributed with zero mean and variance σ_u, and ϵ_{it} is the time varying error term distributed as a logistic function. Defining $v_{it}=\epsilon_{it}+u_{it}$, we obtain that the variance

[13]See Brillinger (2006).

$var\left(v_{it}\right)$ and the correlations $\rho_v$are respectively equal to

$$var\left(v_{it}\right) = \sigma_\epsilon^2 + \sigma_u^2 = 1 + \sigma_u^2 \tag{16}$$

and

$$\rho_v = \frac{\sigma_u^2}{1+\sigma_u^2} \tag{17}$$

The log-likelihood function correspondent to our model can be therefore written as

$$logL = \sum_i^N logProb\left(P_{i1}, \ldots, P_{iT}\right) \tag{18}$$

with

$$Prob\left(P_{i1}, \ldots, P_{iT}\right) = \int_{\mu_0 - X_{it}\beta}^{\mu_1 - X_{it}\beta} \cdots \int_{\mu_{T-1} - X_{it}\beta}^{\mu_T - X_{it}\beta} \phi\left(v_{i1}, \ldots, v_{it}\right) dv_{i1} \ldots, dv_{iT} = \tag{19}$$

$$= \int_{\mu_0 - X_{it}\beta}^{\mu_1 - X_{it}\beta} \cdots \int_{\mu_{T-1} - X_{it}\beta}^{\mu_T - X_{it}\beta} \int_{-\infty}^{+\infty} \phi\left(\epsilon_{it}|u_i\right) \phi\left(u_i\right) du_i d\epsilon_{iT} \ldots d\epsilon_{i1} = \tag{20}$$

$$= \int_{-\infty}^{+\infty} \phi\left(u_i\right) \prod_{t=1}^{T} \left[\Phi\left(\mu_k - X_{it}\beta\right) - \Phi\left(\mu_{k-1} - X_{it}\beta\right)\right] du_i \tag{21}$$

where ϕ and Φ are the density and the cumulative distribution of the logistic distribution respectively. This log-likelihood can be generalized exploiting the technique in Butler and Moffitt (1982), proposing a new computational procedure.[14]

In our case, the computational feasibility is due to the number of points at which the integrand must be evaluated. Having many observations and regressors, the calculation is still extremely burdensome[15], hence we use the procedure proposed by Pregibon (1980), i.e., we employ the standard econometric packages in STATA in analyzing ordinal data exploiting conditional probabilities.

In particular, we consider the probability $Prob\left\{P_{i,j} = 1\right\}$ with respect to $Prob\left\{P_{i,j} \neq 1\right\}$ and the probability $Prob\left\{P_{i,j} = 0.5|P_{i,j} \neq 1\right\}$ with respect to $Prob\left\{P_{i,j} = 0|P_{i,j} \neq 1\right\}$. We estimate our model using the fact that the multinomial probability mass can be represented as a product of binomials. Tables 6 and 8 display the two steps followed in generating the final predictions: first, estimates of the number of goals in the Poisson model (Columns (1) and (3) of Table 6), and then the estimates of the match outcome conditioning on the goals estimated in the first stage (Column (1a) of Table 8). We applied for the single estimation a random effect ordered logit with robust error estimation, while the forecasts calculate the probability of a positive outcome assuming that the random effect is zero.

[14]More in detail, they substitute the trapezoidal integration or its variants as Romberg integration with the Gaussian quadrature.

[15]See, for example, Frechette (2001).

Estimation Method DEPENDENT VARIABLE	(1a) Cl. Logit WIN vs. (DRAW & LOSS) β/se	(2a) Cl. Logit LOSS vs. (WIN & DRAW) β/se	(1b) Cl. Logit WIN vs. (DRAW & LOSS) β/se	(2b) Cl. Logit LOSS vs. (WIN & DRAW) β/se
$AVERAGEP\_HOME_{2004,t-1}$	-3.683	-3.683	-2.913	-5.142
	(4.158)	(4.158)	(4.152)	(4.216)
$AVERAGEP\_HOME_{2004,t-1}$	-5.626	-5.626	-4.324	-6.423
	(4.543)	(4.543)	(4.461)	(4.605)
$AVERAGEP\_HOME_{2003,t-1}$	-3.379	-3.379	-3.229	-6.041
	(3.718)	(3.718)	(3.871)	(3.919)
$AVERAGEP\_HOME_{2003,t-1}$	-5.942	-5.942	-6.662	-4.894
	(3.782)	(3.782)	(3.965)	(4.001)
$AVERAGEP\_HOME_{2002,t-1}$	10.824	10.824	12.537	16.183*
	(6.275)	(6.275)	(6.755)	(7.272)
$AVERAGEP\_HOME_{2002,t-1}$	-1.769	-1.769	-4.984	-5.723
	(6.942)	(6.942)	(7.531)	(7.429)
$POINTS\_HOME_{2004,t-1}$	0.206	0.206	0.142	0.161
	(0.304)	(0.304)	(0.314)	(0.315)
$POINTS\_VISITOR_{2004,t-1}$	-0.241	-0.241	-0.261	-0.300
	(0.268)	(0.268)	(0.277)	(0.285)
$G\hat{OAL}\_HOME_{2004,t-1}$	-0.657*	-0.657*	-0.675*	-0.529
	(0.316)	(0.316)	(0.324)	(0.323)
$G\hat{OAL}\_VISITOR_{2004,t-1}$	-0.226	-0.226	-0.178	-0.071
	(0.351)	(0.351)	(0.368)	(0.368)
$G\hat{OAL}\_HOME_{2004,t}$	0.807***	0.807***	0.881***	0.897***
	(0.217)	(0.217)	(0.225)	(0.235)
IVS_HOME				
$CORRUPTION\_HOME$			1.615***	
			(0.424)	
$CORRUPTION\_VISITOR$				-3.882***
				(1.066)
$CONSTANT$	0.081	0.081	-0.207	-0.077
	(0.841)	(0.841)	(0.868)	(0.878)
N	310	310	310	310
Log-likelihood	-191.328	-191.328	-183.029	-173.320
% Non Rigged Matches Forecasted	54 %		53 %	
% Outliers which are Rigged	56 %		9 %	

Table 7:

Estimation Method DEPENDENT VARIABLE	(1a) Ordered Probit POINTS β/se	(1a) Ordered Probit POINTS β/se	(2) Ordered Probit POINTS β/se	(2a) Ordered Probit POINTS β/se
e $AVERAGEP\_HOME_{2004,t-1}$	-1.240	-1.657	-4.629*	-4.994*
	(1.807)	(2.100)	(2.071)	(2.134)
$AVERAGEP\_HOME_{2004,t-1}$	-4.936**	-4.759*	0.743	0.753
	(1.844)	(2.015)	(2.489)	(2.628)
$AVERAGEP\_HOME_{2003,t-1}$	-0.955	-2.076	-3.314	-3.856
	(2.253)	(2.058)	(2.629)	(2.160)
$AVERAGEP\_HOME_{2003,t-1}$	-2.060	-1.340	2.249	2.391
	(2.341)	(2.042)	(2.091)	(2.512)
$AVERAGEP\_HOME_{2002,t-1}$	6.450*	9.041**	5.088	6.652
	(3.010)	(3.272)	(3.620)	(3.726)
$AVERAGEP\_HOME_{2002,t-1}$	-2.316	-6.163	-2.591	-6.238
	(3.412)	(3.411)	(2.655)	(3.951)
$POINTS\_HOME_{2004,t-1}$	0.085	-0.038	-0.029	-0.180
	(0.195)	(0.167)	(0.139)	(0.201)
$POINTS\_VISITOR_{2004,t-1}$	-0.155	-0.184	-0.093	-0.105
	(0.155)	(0.147)	(0.150)	(0.175)
$G\hat{O}AL\_HOME_{2004,t-1}$	-0.284	-0.198	-0.458*	-0.409*
	(0.192)	(0.177)	(0.192)	(0.181)
$G\hat{O}AL\_VISITOR_{2004,t-1}$	-0.011	0.148	-0.324	-0.209
	(0.220)	(0.179)	(0.188)	(0.220)
$G\hat{O}AL\_HOME_{2004,t}$	0.422**	0.505***	-0.677***	-0.699***
	(0.156)	(0.130)	(0.173)	(0.129)
IVS_HOME			0.139***	0.142***
			(0.017)	(0.012)
$CORRUPTION\_HOME_{2004,t}$		0.951***		1.012***
		(0.214)		(0.289)
$CORRUPTION\_VISITOR_{2004,t}$		-1.238***		-1.076***
		(0.217)		(0.260)
N. obs	310	310	310	310
Pseudo R^2	0.068	0.147	0.395	0.448
Log-likelihood	-307.222	-281.329	-199.616	-181.908
% Non Rigged Matches Forecasted	52 %	53 %	73 %	71%
% Outliers which are Rigged	59 %	36 %	41 %	20 %

Table 8:

Figure 3 shows the distribution of win probabilities obtained via Monte Carlo simulations. The two dotted lines are obtained from the entire sample and from the sample of matches on which there is not an impeding judicial inquiry (apparently non-rigged games). Significantly, the two distributions almost overlap. The continuous line denotes the distribution of rigged matches. It has visibly a different distribution than non-rigged matches, notably a larger mean (a higher average win probability). There are indeed 78 apparently rigged games compared with 310 observations in the full sample. To evaluate this sample size effect, we performed Monte Carlo simulations over a sample of 78 matches not under judicial inquiry. This distribution is denoted by the bold line in the left-hand-side of Figure 3. While it also departs from the distribution of non-rigged matches, it is skewed to the left, rather than to the right. These results sound in support of our strategy to identify rigged matches on the basis of their probability of being outliers. At the same time, they also suggest that the distribution of non-rigged matches is not distinguishable from the distribution of all matches. This means that our method may involve a relatively large number of type II errors (outliers not corresponding to matches under investigation).

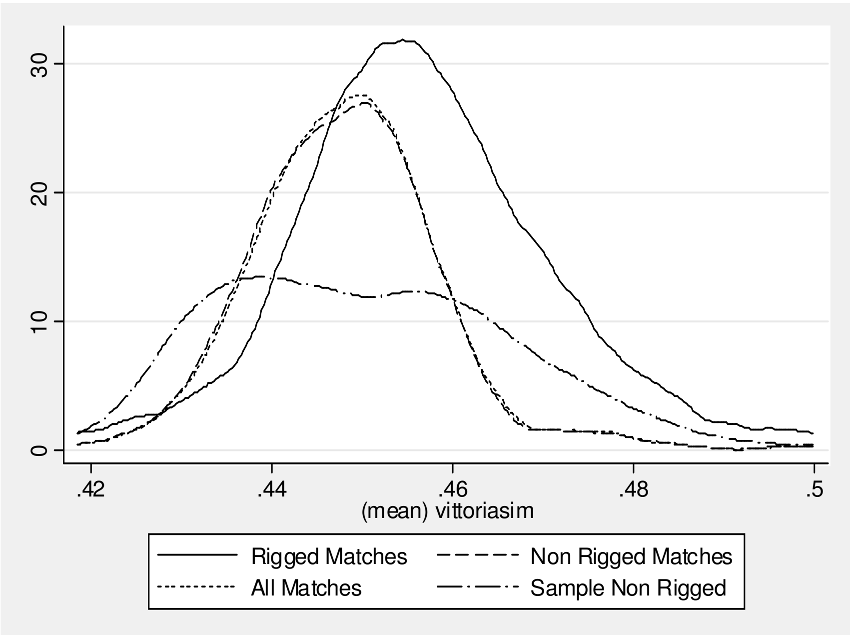

Figure 3: Monte Carlo Simulations: Densities for Different Types of Match

Tables 9, 10, 11 and 12 in the appendix provide information as to the probability of being outliers, based on the discriminant analysis technique[16], of the different matches under investigation. It is noteworthy that 70 out of 78 matches under judicial inquiry (the exceptions being A.C. Milan-Atalanta, Atalanta-Fiorentina, Inter-Atalanta, Cagliari-Lecce, Lazio-Atalanta, Lazio-Parma, Reggina-Chievo and Atalanta-Parma) display a probability of being

[16]See, for example, Greene (2004)

outliers which is higher than 50 per cent. Tables 9 and 10 also in the appendix list type I and type II errors. There are only two matches under investigation which are not outliers (Juventus-Lazio and A.C. Milan-Lazio) but there are 12 type II errors. The identification of corruption episodes based on outliers may lead to overstating the actual number of rigged matches. When making inferences as to rigged matches in other Championships we will therefore combine the outlier criterion with additional information on the determinants of match rigging in order to minimize type II errors.

Match Day	Match	Final Result	Probability to be outliers (%)
5	Udinese-Juventus	0-1	50
7	Siena-Juventus	0-3	50
8	Chievo-Sampdoria	0-2	83
8	Juventus-Roma	2-0	51
8	**A.C. Milan-Atalanta**	3-0	46
9	Juventus-Chievo	3-0	51
9	Messina-Reggina	2-1	62
9	Siena-Bologna	1-1	78
10	Fiorentina-Inter	1-1	72
10	A.C. Milan-Roma	1-1	72
10	Reggina-Juventus	2-1	94
11	Brescia-A.C. Milan	0-0	78
11	Siena-Lecce	1-1	78
12	Cagliari-Inter	3-3	94
12	A.C. Milan-Siena	2-1	82
13	Atalanta-Reggina	0-1	72
13	Messina-Fiorentina	1-1	89
13	Siena-Roma	0-4	50
14	Parma-A.C. Milan	1-2	56
14	Reggina-Brescia	1-3	72
14	Roma-Sampdoria	1-1	72
15	Bologna-Juventus	0-1	56
15	Chievo-Palermo	2-1	94
15	Lazio-Lecce	3-3	78
16	Juventus-A.C. Milan	0-0	94
16	Palermo-Cagliari	3-0	72
16	Roma-Parma	5-1	51
16	Siena-Livorno	1-1	72
17	**Atalanta-Fiorentina**	1-0	46
17	Brescia-Bologna	1-1	83
17	Reggina-Palermo	1-0	78
18	Inter-Sampdoria	3-2	51
18	Lecce-Reggina	1-1	83
18	Messina-Brescia	2-0	60
19	Atalanta-Siena	1-1	78
19	Cagliari-Juventus	1-1	72
19	Livorno-Messina	3-1	57
19	Sampdoria-Bologna	0-0	89

Table 9: Outliers & Matches under Investigation: First Part

Match Day	Match	Final Result	Probability to be outliers (%)
20	Inter-Chievo	1-1	72
20	Lecce-Atalanta	1-0	51
20	Messina-Parma	1-0	63
21	A.C. Milan-Bologna	0-1	100
21	Parma-Udinese	1-0	51
21	Roma-Messina	3-2	57
21	Sampdoria-Siena	1-1	72
22	**Inter-Atalanta**	1-0	46
22	Juventus-Sampdoria	0-1	94
22	Lazio-Brescia	0-0	72
22	Messina-A.C. Milan	1-4	56
23	**Cagliari-Lecce**	3-1	46
23	Chievo-Messina	1-0	62
23	Palermo-Juventus	1-0	51
23	Roma-Bologna	1-1	89
24	Fiorentina-Parma	2-1	83
24	**Lazio-Atalanta**	2-1	46
24	Siena-Messina	2-2	78
25	Palermo-Lecce	3-2	57
25	Siena-Fiorentina	1-0	51
25	Udinese-Inter	1-1	72
26	Inter-A.C. Milan	0-1	83
26	**Lazio-Parma**	2-0	46
26	Lecce-Messina	1-0	71
26	**Reggina-Chievo**	1-0	46
27	Atalanta-A.C. Milan	1-2	94
27	Messina-Lazio	1-0	64
27	Roma-Juventus	1-2	67
27	Siena-Brescia	2-3	100
28	**Atalanta-Parma**	1-0	46
28	Reggina-Messina	0-2	67
29	Inter-Fiorentina	3-2	51
30	Fiorentina-Juventus	3-3	72
30	Udinese-Roma	3-3	94
31	Livorno-Fiorentina	2-0	89
33	Bologna-Fiorentina	0-0	72
33	Lazio-Juventus	0-1	50

Table 10: Outliers & Matches under Investigation: Second Part

Match Day	Match	Final Result	Probability to be outliers (%)
14	Juventus-Lazio	2-1	0
23	A.C. Milan-Lazio	1-0	0

Table 11: Type I Error: Matches under Investigation & Non Outliers

Match Day	Match	Final Result	Probability to be outliers (%)
9	Inter-Lazio	1-1	100
10	Atalanta-Sampdoria	0-0	100
10	Lecce-Udinese	3-4	100
11	Inter-Bologna	2-2	100
12	Fiorentina-Livorno	1-1	100
14	Juventus-Lazio	2-1	100
17	Chievo-Siena	1-3	100
23	A.C. Milan-Lazio	2-1	100
25	Messina-Juventus	0-0	100
27	Udinese-Bologna	0-1	100
29	Messina-Bologna	0-0	100
32	Juventus-Inter	0-1	100

Table 12: Type II Error: Matches not under Investigation & Outliers

5 Detecting Corruption in the Previous Championships

Tables 13 and 14 reproduce the two steps of our preferred specification in predicting fair matches applied to data covering the 1997-2003 period. We work now with a sample of roughly 2,000 observations. The signs of the coefficients are once more in line with theoretical (and a priori) expectations. Compared with the results obtained when focusing on the 2004-5 Championship (Table 6), the lagged dependent variable is no longer statistically significant, indicating that at least part of the autoregressive structure of the data is lost when we operate over this larger sample. Importantly, the coefficients of the variables capturing the past performance of the home and away teams are larger when considering matches played in the 1999-2000 and 2000-2001 Championships. According to the reconstruction offered by the press, these two Championship are those where there were less dubious episodes, that could be possibly associated to match rigging.

Figures 4 and 5 compare the normal distribution with the distribution of win and loss probabilities obtained via Monte Carlo simulations. In the case of the win probabilities, the Monte Carlo simulations involve higher win probabilities than those implied by the normal distribution in proximity of the mean. In the case of loss probabilities, it is just the opposite: there is an underprediction of loss probabilities with respect to the normal distribution. Due to the strong home bias of Italian football, our empirical model assigns a high probability of victory to the home team.

We therefore use the following criteria in identifying potential corruption episodes in the previous Championships:

1. The win/loss probabilities should be in the region where the Monte Carlo simulation deviates from the normal distribution: the difference between the estimated density and the normal one should more than the 90% percentile for the win probabilities of the

Dependent Variable	(1) $GOAL\_HOME$	(2) $GOAL\_VISITOR$
$MATCH$	0.005**	0.002
	(0.002)	(0.002)
$FRIEND$	-0.022	-0.029
	(0.136)	(0.121)
$ENEMIES$	-0.076	0.080
	(0.052)	(0.105)
$VISITOR$	0.007***	
	(0.002)	
$AVERAGEP\_HOME_{2003,t-1}$	1.045***	-0.366
	(0.127)	(0.192)
$AVERAGEP\_VISITOR_{2003,t-1}$	-0.698***	0.633***
	(0.190)	(0.148)
$AVERAGEP\_HOME_{2002,t-1}$	0.812***	-0.528**
	(0.097)	(0.161)
$AVERAGEP\_VISITOR_{2002,t-1}$	-0.432**	0.613***
	(0.142)	(0.129)
$AVERAGEP\_HOME_{2001,t-1}$	1.027***	-0.355*
	(0.148)	(0.171)
$AVERAGEP\_VISITOR_{2001,t-1}$	-0.534**	0.573***
	(0.177)	(0.171)
$AVERAGEP\_HOME_{2000,t-1}$	0.751***	-0.644**
	(0.130)	(0.200)
$AVERAGEP\_VISITOR_{2000,t-1}$	-0.247	0.920***
	(0.198)	(0.152)
$AVERAGEP\_HOME_{1999,t-1}$	1.262***	-0.327
	(0.250)	(0.235)
$AVERAGEP\_VISITOR_{1999,t-1}$	-0.689**	0.285
	(0.250)	(0.161)
$AVERAGEP\_HOME_{1998,t-1}$	1.037***	-0.423
	(0.187)	(0.370)
$AVERAGEP\_VISITOR_{1998,t-1}$	-0.291	0.489
	(0.214)	(0.268)
$AVERAGEP\_HOME_{1997,t-1}$	1.459***	-0.423
	(0.212)	(0.335)
$AVERAGEP\_VISITOR_{1997,t-1}$	-0.902***	0.692*
	(0.258)	(0.289)
$NEW\_COACH\_HOME_t$	0.019	0.084
	(0.050)	(0.053)
$NEW\_COACH\_VISITOR_t$	0.075*	-0.065
	(0.036)	(0.043)
$GOAL\_HOME_{t-1}$	0.022	-0.017
	(0.013)	(0.014)
$GOAL\_VISITOR_{t-1}$	-0.038	0.027
	(0.042)	(0.058)
$HOME$		-0.005
		(0.003)
χ^2	.	.
N	1,998	1,998
$*p < 0.05$, $**p < 0.01$, $***p < 0.001$		

Table 13: First Step

Dependent Variable: $POINTS$	$\beta/s.e.$
$AVERAGEP\_HOME_{2003,t-1}$	1.314***
	(0.258)
$AVERAGEP\_VISITOR_{2003,t-1}$	-1.467***
	(0.209)
$AVERAGEP\_HOME_{2002,t-1}$	1.067***
	(0.188)
$AVERAGEP\_VISITOR_{2002,t-1}$	-0.954***
	(0.169)
$AVERAGEP\_HOME_{2001,t-1}$	1.994***
	(0.431)
$AVERAGEP\_VISITOR_{2001,t-1}$	-2.006***
	(0.363)
$AVERAGEP\_HOME_{2000,t-1}$	1.558***
	(0.411)
$AVERAGEP\_VISITOR_{2000,t-1}$	-1.628***
	(0.393)
$AVERAGEP\_HOME_{1999,t-1}$	2.087***
	(0.487)
$AVERAGEP\_VISITOR_{1999,t-1}$	-1.747***
	(0.455)
$AVERAGEP\_HOME_{1998,t-1}$	1.467***
	(0.294)
$AVERAGEP\_VISITOR_{1999,t-1}$	-0.860**
	(0.286)
$AVERAGEP\_HOME_{1997,t-1}$	1.380***
	(0.284)
$AVERAGEP\_VISITOR_{1997,t-1}$	-1.384***
	(0.333)
$LAST\_HOME$	-0.100
	(0.068)
$LAST\_VISITOR$	0.000
	(0.069)
$G\hat{O}AL\_HOME_{t-1}$	0.053
	(0.044)
χ^2	1436.321
N	1,951
$*p < 0.05$, $**p < 0.01$, $***p < 0.001$	

Table 14: Second Step

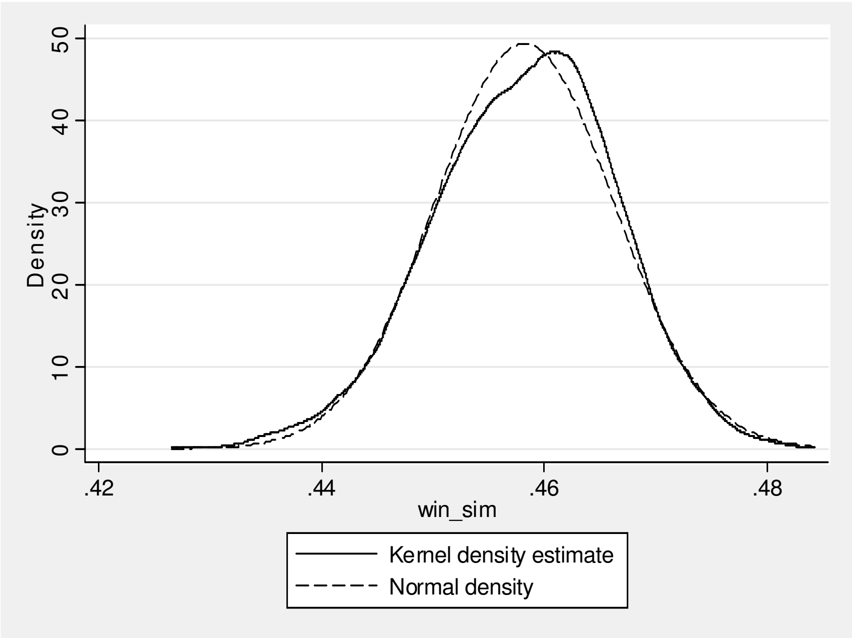

Figure 4: Monte Carlo Simulations: Densities for the Win Probabilities of the Home Team and Normal Distribution

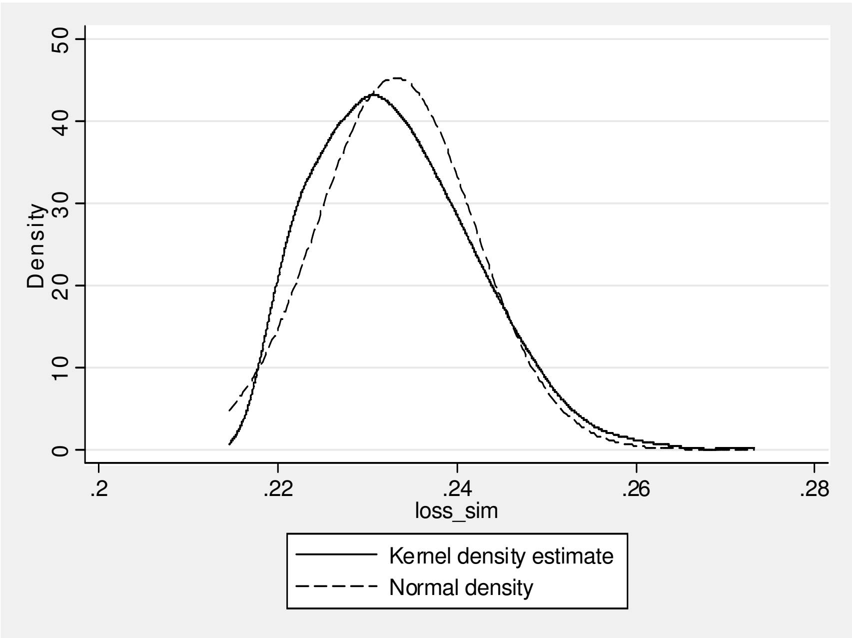

Figure 5: Monte Carlo Simulations: Densities for the Loss Probabilities of the Home Team and Normal Distribution

home team *and* less of the 90% percentile for the loss probabilities for the visitor.

2. The characteristics of these matches (in terms of timing, competitive power of the first two teams) should involve a corruption probability higher than 80 % according to the coefficients estimated in Table 5;

3. The matches belong to the grid A, , meaning that they are important for the career prospects of the referees.

We could only implement this methodology to the 2002-2003 and 2003-2004 Championships, as we have no data on matches in grid A as well as preclusions in previous Championships. Table 13 displays the two matches that could have been rigged according to the above criteria. Both of these events correspond to matches pointed out as potentially rigged by the Italian press.

A number of controverted referee decisions were made all in favour of Juventus.

Championship	Home	Visitor			Match Day	Total number of preclusions
2003/2004	Modena	Juventus	0	2	6	8
2003/2004	Juventus	Perugia	1	0	12	6

Table 15: Rigged Matches according to the Model

6 Final remarks

In this paper, we analyzed corruption in Italian soccer drawing on a simple model of allocation of "blackmail capital" and on detailed data on the competitive power of teams in the First Division and of career concerns of referees. Unlike previous studies on corruption, we could draw on matches being under judicial inquiry on the basis of hard evidence of pressure exerted by managers on the referees, notably tapped phone conversations. Guided by the theoretical model, this information allowed us to empirically assess the determinants of corruption. Unlike Duggan and Levitt (2002) we could then use these regression results in identifying potentially rigged matches in combination with the detection of outliers in regression prediction and the importance of the match in terms of the career of the referees.

Our results suggest that corruption could have been widespread well before the 2004-5 Championship.

Three implications of our results are potentially relevant beyond the case of Italian soccer, as they have a broader application.

The first implication is that it may be misleading to use only the outlier method in identifying potential corruption episodes. This is because often corruption targets rather balanced matched, where the deviation from the fair outcome is not too large.

The second implication is that career concerns make corruption less costly. Thus, it is important to monitor very closely the behavior of agents having strong career concerns and make as much as possible transparent the criteria followed in promoting them to international competitions.

The third implication is that more competitive balance is not necessarily a factor reducing corruption activities. We document that corruption increased just at times in which the competitive lead of Juventus was lower. We find instead some (admittedly preliminary) evidence that concentration in media power plays in favor of corruption.

Further work should possibly increase the number of observations and dig deeper into the effect of media concentration on potential match rigging. Our findings suggest that policies repressing corruption in industries characterized by "winner-takes-all" preferences should aim at increasing pluralism in the coverage of sport events and promote a more balanced allocation of TV rights. Another implication is that detection of corruption events should be concentrated on matches having an *ex ante* fairly balanced outcome.

References

Ades, A., and R. D. Tella (1997): "National Champions and Corruption: Some Unpleasant Interventionist Arithmetic," *The Economic Journal*, 107(5), 1023–1042.

Becker, G. S., and G. J. Stigler (1974): "Law Enforcement, Malfeasance, and the Compensation of Enforces," *Journal of Legal Studies*, III, 1–19.

Bernheim, B. D., and M. D. Whinston (1998): "Exclusive Dealing," *Journal of Political Economy*, 106(1), 64–103.

Bliss, C., and R. D. Tella (1997): "Does Competition Kill Corruption?," *The Journal of Political Economy*, 105(5), 1001–1023.

Brillinger, D. (2006): "Modelling Some Norwegian Soccer Data," *University of California, Berkley.*

Butler, J., and R. Moffitt (1982): "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model," *Econometrica*, 50, 761–764.

Deloitte, and Touche (2003): "Annual Review of Football Finance," .

Dobson, S., and J. Goddard (2001): "The Economics of Football," *Cambridge University Press.*

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

Frechette, G. (2001): "REOPROB command," *Stata Technical Bulletin*, 59.

Gambetta, D. (1993): "The Sicilian Mafia," *Harvard University Press.*

Garlando, L. (2005): "Nostra Signora del Dischetto," *Kowalski.*

Greene, W. (2004): "Econometric Analysis," *Pearson International, V Edition.*

Grossman, G., and E. Helpman (2001): "Special Interest Politics," *MIT Press.*

Karlis, D., and I. Ntzoufras (2003): "Analysis of Sports Data by Using Bivariate Poisson Models," *Statistician*, 49, 419–431.

Kloek, T. (1981): "OLS Estimation in a Model Where a Microvariable Is Explained by Aggregates and Contemporaneous Disturbances Are Equicorrelated," *Econometrica*, 49, 1–19.

Koning, R. (2000): "Balance in Competition in Dutch Soccer," *The Statistician*, 49(5), 419–431.

Kuypers, T. (2000): "Information and Efficiency: An Empirical Study of a Fixed Odds Betting Market," *Applied Economics*, 32(11), 1353–63.

Mauro, P. (1995): "Corruption and Growth," *Quaterly Journal of Economics*, 90(2), 681–712.

Pregibon, D. (1980): "Discussion of paper by P. McCullagh," *Journal of the Royal Statistical Society Series (B) Methodological*, 42, 109–142.

Rose-Ackerman, S. (1975): "The Economics of Corruption," *Journal of Public Economics*, IV, 187–203.

Shleifer, A., and R. W. Vishny (1993): "Corruption," *The Quaterly of Economics*, 108(3), 599–617.

Szymanski, S., and T. Kuypers (1999): "Winners and Losers: The Business Strategy of Football," *Viking*.

Tomz, M., J. Wittenberg, and G. King (2001): "Clarify: Software for Interpreting and Presenting Statistical Results," *Journal of Statistical Software*, 8(i01), available at http://gking.harvard.edu.

Wolters, J. (2006): "Point Shaving: Corruption in NCAA Basketball," *American Economic Review*, 96(2).

Annex 1

Weeks to the end	*Match*	*Result*	*IVS H.*	*IVS V.*	*Rigged Episode*
34	Udinese-Brescia	1-2	51	49	Red card for Jankulovksi (Udin.)
33	Udinese-Juventus	0-1	23	77	Jankulovksi cannot play
32	Fiorentina-Siena	0-0	51	49	Red card for Pasquale (Siena)
31	Siena-Juventus	0-3	31	69	Pasquale cannot play
30	Chievo-Sampdoria	0-2	48	52	-
30	Juventus-Roma	2-0	65	35	Penalty for Juve when Ibra's in offside
30	A.C. Milan-Atalanta	3-0	72	28	-
29	Juventus-Chievo	3-0	65	35	Juve helped by invented offsides
29	Messina-Reggina	2-1	61	39	-
29	Siena-Bologna	1-1	60	40	-
28	Fiorentina-Inter	1-1	46	54	-
28	A.C. Milan-Roma	1-1	54	46	-
28	Reggina-Juventus	2-1	42	58	Referee against Juve
27	Brescia-A.C. Milan	0-0	37	63	-
27	Siena-Lecce	1-1	56	44	-
26	Cagliari-Inter	3-3	34	66	-
26	A.C. Milan-Siena	2-1	63	37	-
25	Atalanta-Reggina	0-1	46	54	-
25	Messina-Fiorentina	1-1	57	43	-
25	Siena-Roma	0-4	22	78	-
24	Parma-A.C. Milan	1-2	31	69	-
24	Reggina-Brescia	1-3	43	57	-
24	Roma-Sampdoria	1-1	54	46	-
23	Bologna-Juventus	0-1	32	68	Referee favorable to Juventus
23	Chievo-Palermo	2-1	51	49	-
23	Lazio Rome-Lecce	3-3	56	44	-
22	Juventus-A.C. Milan	0-0	36	64	Penalty refused to A.C. Milan
22	Palermo-Cagliari	3-0	61	39	-
22	Roma-Parma	5-1	65	35	Red card Pisanu (Parma)
22	Siena-Livorno	1-1	50	50	
21	Parma-Juventus	1-1	38	62	Pisanu cannot play
21	Atalanta-Fiorentina	1-0	62	38	-
21	Brescia-Bologna	1-1	64	36	Yelow cards for Brescia
21	Reggina-Palermo	1-0	56	44	-
20	Inter-Sampdoria	3-2	66	34	-
20	Lecce-Reggina	1-1	62	38	-
20	Messina-Brescia	2-0	68	32	A lot of Brescia players suspended
19	Atalanta-Siena	1-1	60	40	-
19	Cagliari-Juventus	1-1	46	54	Offsides against Cagliari
19	Livorno-Messina	3-1	63	37	-

Table 16: Rigged Matches, Results and IVS: First Part

Weeks to the end	*Match*	*Result*	*IVS H.*	*IVS V.*	*Rigged Episode*
19	Sampdoria-Bologna	0-0	64	36	-
18	Inter-Chievo	1-1	54	46	-
18	Lecce-Atalanta	1-0	60	40	-
18	Messina-Parma	1-0	56	44	Referee favorable to Messina
17	A.C. Milan-Bologna	0-1	60	40	Referee favorable to Bologna
17	Parma-Udinese	1-0	62	38	-
17	Roma-Messina	3-2	63	37	-
17	Sampdoria-Siena	1-1	52	48	-
16	Inter-Atalanta	1-0	86	14	-
16	Juventus-Sampdoria	0-1	53	47	-
16	Lazio Rome-Brescia	0-0	50	50	Pressure on the referee
16	Messina-A.C. Milan	1-4	32	68	-
15	Cagliari-Lecce	3-1	67	33	-
15	Chievo-Messina	1-0	56	44	-
15	Palermo-Juventus	1-0	65	35	-
15	Roma-Bologna	1-1	62	38	-
14	Fiorentina-Parma	2-1	53	47	Referee favorable to Fiorentina
14	Lazio Rome-Atalanta	2-1	63	37	Referee chosen by Moggi
14	Siena-Messina	2-2	57	43	Referee favorable Messina
13	Palermo-Lecce	3-2	58	42	Red cards for Lecce players
13	Siena-Fiorentina	1-0	58	42	Red card for Siena players
13	Udinese-Inter	1-1	47	53	-
12	Inter-A.C. Milan	0-1	44	56	-
12	Lazio Rome-Parma	2-0	65	35	-
12	Lecce-Messina	1-0	57	43	Lecce players suspended
12	Reggina-Chievo	1-0	74	26	-
11	Atalanta-A.C. Milan	1-2	48	52	-
11	Messina-Lazio Rome	1-0	53	47	-
11	Roma-Juventus	1-2	42	58	Referee favorable to Juventus
11	Siena-Brescia	2-3	54	46	-
10	Atalanta-Parma	1-0	73	27	-
10	Reggina-Messina	0-2	43	57	Referre favorable to Messina
9	Inter-Fiorentina	3-2	61	39	-
8	Fiorentina-Juventus	3-3	42	58	Fiorentina players cannot play
8	Udinese-Roma	3-3	61	39	-
7	Livorno-Fiorentina	2-0	46	54	-
5	Bologna-Fiorentina	0-0	44	56	-
5	Lazio Rome-Juventus	0-1	27	73	Referee favorable to Juventus
2	Messina-Cagliari	2-1	na	na	-
2	Udinese-Sampdoria	1-1	na	na	-

Table 17: Rigged Matches, Results and IVS: Second Part

Appendix 1: Teams' Friends & Enemies

ATALANTA:	Friends: - Enemies: Brescia Inter A.C. Milan Palermo
BOLOGNA:	Friends: Udinese Enemies: Juventus Roma Parma
BRESCIA:	Friends: A.C. Milan Enemies: Atalanta Inter Roma Lazio Rome Palermo
CAGLIARI:	Friends: Sampdoria Foggia Enemies: A.C. Milan Verona Napoli Palermo
CATANIA:	Friends: - Enemies: Palermo Messina
CHIEVO:	Friends: - Enemies: -
FIORENTINA:	Friends: Livorno Enemies: Juventus Roma Lazio Rome
INTER:	Friends: Lazio Rome Enemies: Juventus A.C. Milan Napoli
JUVENTUS:	Friends: - Enemies: Fiorentina A.C. Milan Inter
Lazio Rome:	Friends: Inter Enemies: Roma Fiorentina Palermo
LECCE:	Friends: Palermo Enemies: Reggina Roma
LIVORNO:	Friends: Fiorentina Enemies: Lazio Rome
MESSINA:	Friends: - Enemies: Catania Palermo Reggina
A.C. Milan:	Friends: Brescia Reggina Enemies: Inter Juventus Atalanta Roma Lazio Rome Sampdoria
PALERMO:	Friends: Lecce Roma Enemies: Catania Messina Reggina Brescia Atalanta Lazio Rome
PARMA:	Friends: Sampdoria Enemies: Juventus Bologna
REGGINA:	Friends: Roma A.C. Milan Enemies: Messina Palermo
ROMA:	Friends: Reggina Palermo Enemies: Lazio Rome Juventus A.C. Milan Inter Sampdoria Fiorentina
SAMPDORIA:	Friends: Parma Enemies: A.C. Milan
SIENA:	Friends: - Enemies: Fiorentina Livorno
UDINESE:	Friends: Bologna Enemies: -