Match Fixing and Sports Betting in Football: Empirical Evidence from the German Bundesliga

Christian Deutscher
University of Bielefeld

Eugen Dimant
University of Pennsylvania

Brad R. Humphreys
West Virginia University

This version: January 2017

Abstract

Corruption in sports represents an important challenge to their integrity. Corruption can take many forms, including match fixing by players, referees, or team officials. Match fixing can be difficult to detect. We use a unique data set to analyze variation in bet volume on Betfair, a major online betting exchange, for evidence of abnormal patterns associated with specific referees who officiated matches. An analysis of 1,251 Bundesliga 1 football matches from 2010/11 to 2014/15 reveals evidence that bet volume in the Betfair markets in these matches was systematically higher for four referees relative to matches officiated by other referees. Our results are robust to alternative specifications and are thus suggestive of potentially existing match fixing and corruption in the German Bundesliga.

Keywords: corruption, betting exchange, football, referee bias

JEL: D73, K42, L8, Z2

Deutscher: University of Bielefeld, Faculty of Psychology and Sports Science, Department of Sport Science. Postfach 10 01 31, D-33501 Bielefeld Germany; Phone : +49 521 106-2006; Email: christian.deutscher@uni-bielefeld.de

Dimant: University of Pennsylvania, Behavioral Ethics Lab, Philosophy, Politics and Economics Program. 311 Claudia Cohen Hall, 249 S 36th Street, 19104 Philadelphia, USA; Phone: 215-898-3023; E-mail: edimant@sas.upenn.edu

Humphreys: Department of Economics, West Virginia University. 1601 University Ave., PO Box 6025, Morgantown WV 26506-6025; Phone: (304) 293-7871; Email: brhumphreys@mail.wvu.edu

Introduction

Corruption in sports is a growing problem with new allegations of match fixing regularly reported in the media. Evidence of match fixing can often be found as unusual patterns in aggregated data from betting markets, since match fixers profit by placing bets on matches with pre-determined outcomes (Forrest & Simmons, 2003). We extend the approach of Wolfers (2006) and investigate the idea that evidence of match fixing can be found in available data from a popular international betting exchange market, Betfair.

Sports betting is a growing industry and has become an integral tool for profiting from fixed matches. According to the European Gaming & Betting Association, regulated betting accounts for some $58 billion yearly and is forecast to reach $70 billion in 2016, with football (soccer in North America) accounting for about 70-85% of the bets placed (EGBA, 2014).

Economic models of match fixing predict that referees are prime candidates for corruption, since they can exert a strong influence on match outcomes and receive relatively low levels of compensation. Betting exchange markets provide convenient, highly liquid markets where match fixers can profit from influence on outcomes in sporting events. Previous match fixing scandals contain evidence that referees are sometimes involved in match fixing scandals.

Matches can be fixed in numerous ways and could possibly involve many different individuals, including team managers, staff, players, and match officials. Match fixing by referees is an intentional form of referee bias and matches can be fixed on numerous margins including outcome (home win, draw, away win), goals scored, and other outcomes colloquially referred to as "proposition bets" in gambling markets (LaBrie et al., 2007). We focus on the role played by referees in conjunction with specific wagers on the total number of goals scored in football matches that can be linked to match fixing. This type of match fixing requires a small number of initiators, increasing the individual benefit for all parties involved. With their career at stake in case of detection, a referee takes on a huge risk when fixing a match.[1] Dohmen and Sauermann (2016) survey the large literature on this topic that circles around the idea of referees making advantageous

---

[1] Boeri and Severgnini (2011) emphasize the importance of expected future career earnings by referees in influencing their decisions to provide unbiased adjudication of play and point out that referees involved in the *Calciopoli* match fixing scandal in Italy were coerced into influencing match outcomes rather than bribed.

decisions toward the home team as a result from social pressure by the home fans. In this context, referees can be understood as intermediaries who engage in corrupt behavior (Dusha, 2015).

The source of the exchange betting market data, Betfair, bills itself as the "world's largest betting exchange" and reported 1.7 million active customers in 2015 and a turnover of 475.6 million British pounds (approximately $694 million). We capitalize on a unique dataset to analyze variation in the volume of Betfair wagers on the total number of goals scored in German Bundesliga 1 games in the 2010/11 – 2014/15 seasons. We posit that match fixing in terms of total goals scored is more likely to occur than match fixing in terms of win/loss/draw outcomes as it has a lower detection rate and hence is less risky to the parties involved in the fixing process. In addition, unbiased referees should not affect the expected number of goals scored in a match.

Regression models include control variables capturing specific match characteristics, including the identity of the referee for each match and unobservable home team-, away team-, matchday-, and season-level heterogeneity. Regression results indicate that Betfair betting volume on the proposition that games end with over or under 2.5 total goals scored was higher in games refereed by four specific Bundesliga referees over this period, even when controlling for team and match level observable and unobservable factors that might affect goal scoring. These results are consistent with the hypothesis that corruption might have influenced some Bundesliga 1 match outcomes over this period. The paper proceeds with presenting a literature review on corruption inside and outside of sports, followed by a presentation of the data and empirical estimations on the impact of referees on betting volume. It concludes with critical remarks and an outlook of research ahead.

Literature Review and Context

Although the extent and quality media coverage varies (see Di Tella & Franceschelli, 2011), corruption in general, but particularly in sports, is a ubiquitous issue in both amateur and professional settings, especially in the form of match fixing. Corruption imposes adverse effects on both the society and economy, thus signifying importance of gaining a better understanding of the underlying mechanisms that inform policy makers (Rose-Ackerman & Palifka, 2016; for a comprehensive discussion of the mechanisms of corrupt behavior as well as the empirical findings on the causes and effects, see Dimant & Schulte, 2016 and Dimant & Tosato, forthcoming). As a

subset of corruption, match fixing represents a substantial threat for the integrity of sports (Dimant & Deutscher, 2015). It is a widespread phenomenon and has been uncovered in sports such as Basketball, Cricket, Football, Sumo Wrestling, and Tennis. Such distortions create negative externalities not only on the individual but also on the aggregate level, such as loss of media interest. Additionally, they erode the inherent principle of fair and competitive sports. Scholars have proposed a number of mechanisms to deal with match fixing, some of which have been implemented by policy makers, although with differing success (Carpenter, 2012). Forrest, McHale and McCauley (2008) discuss the economic incentives that influence match fixing in sport from a more general view. They point out that the emergence of betting exchange markets like Betfair increases the incentives to fix matches since they provide enhanced opportunities to benefit financially from match fixing due to a quick and fairly anonymous exchange of money.

A substantial literature exists on the economics of match fixing in sports. Preston and Syzmanski (2003) developed a game theoretic model of strategic interaction between bettors, bookmakers, and participants in sporting events. The model assumes that participants in sporting events may be susceptible to corruption, given sufficient monetary incentives, and shows that the likelihood of corruption increases as the legal compensation of the participants decreases. This prediction implies that referees are prime candidates for corruption because of their relatively low levels of compensation, especially in comparison to coaches and players (McHale and McCauley, 2008). Pohlkamp (2014) reports compensation rates for Bundesliga 1 referees at €3,800 per match with no base compensation in 2009. In comparison, Premier League referees earned base salaries of €38,500 per season and an additional €1,170 per match in 2009. Referees in Bundesliga 1 pursue other jobs besides refereeing, ranging from dentist to lawyer. Preston and Syzmanski (2003) point out that a referee can have a larger impact on match outcomes than most players can, which also makes them prime candidates for corruption. Since the expected returns of wrongful behavior are negative if said behavior is uncovered, increasing referee compensation can be interpreted as reducing incentives to cheat. Premier League referees who switch from short-term contracts to salaried contracts show improved performance relative to those who do not (Bryson, Buraimo and Simmons, 2011). In line with this argumentation, Forrest and Simmons (2003) developed a model to explain match fixing in sports based on the expected costs and benefits of match fixing. This model also suggests that referees are likely candidates for corruption, in that the probability that an individual will take actions to affect the outcome of a sporting event increases with the likelihood

4

that the actions succeed in affecting the outcome. The probability that the actions of an individual referee can influence outcomes exceeds the probability for coaches and almost all players.

Forrest (2012) surveyed a number of recent match fixing scandals in sports, including the infamous 2011 "Bochum trial" in Germany where evidence of match fixing in more than 300 European football matches, including 53 in Germany, was introduced. Several of the match fixing scandals discussed by Forrest (2012) involved referees. Feltes (2013) discusses two key cases of match fixing in Germany involving referee corruption: the 2000 Hoyzer case and the 2011 "Bochum trial" involving Ante Sapina. Robert Hoyzer, a Bundesliga referee, was found guilty of fixing 23 Bundesliga matches and convicted of fraud. He was caught after a number of egregious calls, including awarding two penalties to SC Paderborn in a surprising 4-2 win after trailing 0-2 in a 2004 match against Hamburg and ejecting Hamburg's star striker for misconduct. Hoyer implicated Sapina as the source of his bribes, but Sapina was not prosecuted until years later.

Ante Sapina, the leader of a betting syndicate, was found guilty of fixing 32 football matches in Germany. Both referees and players were involved in the Sapina match fixing scandal. Two referees implicated in this scandal were banned for life by FIFA and UEFA, including Bosnian referee Novo Panic. Among the details emerging from the Sapina trial, one most relevant to the current work was his betting on the number of goals scored in matches among other "proposition" bets like the number of free kicks taken in a match.

Following the Hoyzer case, the committee of control for the Bundesliga (DFB) reacted by taking a number of actions, most prominently reducing the time between the designation of the referee for a match and the playing of the match. Aimed at minimizing the time available to fix football matches, the League proposed, and failed to implement a planned two day notice before the match because that process was ruled impractical. Instead, following the Hoyzer case, referees are assigned to matches according to the following schedule: referees for Friday games are announced on Wednesday and the referee assignments for games played between Saturday and Monday follows on Thursday; the announcement for Tuesday and Wednesday games happens on Fridays. This procedure features a time span of between two and five days between the referee assignment and the match, depending on the day of the match. In terms of the identity of the referee, the experience of the referee determines the chances a referee is assigned to a certain game; decisive games at the end of the season are assigned to experienced referees. To assure unpredictability of

assignments, no clear referee assignment mechanism has been announced by the league. Possible limitations of this process for our analysis are described in the last section of the paper.

In another famous case, Boeri and Severgnini (2011) analyze referee participation in the *Calciopoli* match fixing scandal in Italy in 2006. Using evidence about specific episodes of match fixing in *Serie A* uncovered through phone taps and other methods used in a criminal investigation of football match fixing, Boeri and Severgnini (2011) demonstrate how club officials used threats to adversely affect the career, and future earnings, of Italian football referees to Juventus's direct or indirect benefit. Referee's actions included issuing red cards to key players in matches immediately before a team was scheduled to play Juventus (disqualifying said player from the next match), incorrectly ruling (or failing to rule) players offside or not ruling players offside, and other subtle actions. The corrupted referees did not take overt actions like assessing red cards to opposing players in matches involving Juventus or awarding Juventus penalty kicks in important matches.

Various factors motivate our analysis of variation explicitly in bet volume in the *Over 2.5* and *Under 2.5* markets on Betfair. First, the media and the public are likely less sensitized to questionable calls that result in additional goals compared to dubious calls that change the outcome of the contest. Second, when comparing benefits and costs of match fixing, taking actions that affect the total number of goals scored reduces the chances of a referee getting caught, thus increasing the attractiveness of match fixing. For referees not being bribed, there should, *ceteris paribus*, be no systematic differences in the amount of money bet on the number of goals scored in a football match.[2] Closely related is the case of Tim Donaghy and the 2007 NBA betting scandal, where the former referee bet money on the *over* (the proposition that more than a specific number of points would be scored) in games he refereed (Lookwood, 2008).

Empirical Analysis

In order to develop evidence consistent with the presence of match fixing in the Bundesliga 1, we estimate reduced form models of the determination of betting volume on specific bets placed on

---

[2] If non-bribed referees systematically differ in their evaluation of fouls, systematic differences in the money bet on 'over' could occur. Image one referee is known to rather not call fouls and let the game continue in controversial situations. This knowledge could lead to statistically higher volume on over / under betting. This paper tries to control for this by using referee decision making as control variables (number of cards, penalties etc.)

Betfair. Match fixing occurs because some individual or organization wants to profit from sporting event outcomes by influencing the outcome of these events in predictable ways. We assume that matches with an unusually large bet volume, or with unusual patterns in said bet volume, could potentially be fixed in some way, reflecting bets made by the match fixers.

Match fixing, and profits from match fixing, could take many forms. For example, in sporting events with prizes for winners, match fixing could involve payoffs to participants to guarantee specific outcomes; in a foot race with a first prize of $1,000 and a second prize of $500, the two fastest runners could agree before the race to split the sum of the first and second prizes equally. However, a simpler way to profit from match fixing is to bet on some match outcome that has been determined in advance, or that some participant in the match has been paid or coerced to influence in a specific way. The Ante Sapina case discussed above revolved around profits earned by betting on fixed football matches in Germany.

An individual or organization attempting to profit from match fixing by betting on Betfair would need to place bets on an outcome that would be relatively straightforward to influence and relatively difficult to detect. The Appendix describes the types of betting markets available on Betfair. The largest Betfair football betting markets, in terms of bet volume, are the match odds markets (home win, draw, away win) and the over/under 2.5 goals markets. To profit on match odds betting, the match fixer would have to influence the outcome of the match, a more easily detected form of corruption due to the reasons described above. Exact match score markets have low volume and this outcome would be relatively difficult to fix. The over/under 0.5 and 1.5 goal markets also have low volume, so individual high volume bets would likely be identified as suspicious by market monitoring systems.

Based on these considerations, the over/under 2.5 goal market appears to be a likely candidate for match fixers to exploit to earn profits on fixed matches. In the 1,530 Bundesliga 1 matches played in the 2010/11 through 2014/15 seasons, the average number of goals scored in a match was 2.92. Two or fewer goals were scored in 44% of the matches and three or more goals were scored in 56% of the matches. A player or referee would not need to influence scoring in a glaringly obvious way to drive scoring over/under 2.5 goals.

Bet volume can clearly be influenced by factors unrelated to match fixing. Bettors may prefer to wager on more popular teams, on teams with star strikers, on teams playing opponents with weak

defenders, or simply prefer to wager in the *Over 2.5* market because they prefer matches with more scoring. In any event, if the *Over 2.5* and *Under 2.5* betting markets are weak form efficient, then all public information affecting match outcomes, including referee effects, should be reflected in betting odds.

Data

The data come from Betfair, an on-line betting exchange founded in the UK in 1999. Betting exchanges like Betfair allow bettors to both back (bet that an athlete or team will win a sporting event, or bet that some event will occur) or lay (bet that an athlete or team will not win a sporting event, or an event will not occur) on any sporting event. Traditional betting with a bookmaker involves the bettor backing and the bookmaker laying on each transaction. On a betting exchange, each wager must be matched: at least one backer and one layer must agree to wager a specific amount of money at stated odds on a specific event. Betfair quickly matches backers and layers. Sometimes multiple backers and layers are matched at stated odds, which allows for wagering both before and during ("in-play") sporting events.

We obtained data on Betfair betting prior to football matches in Bundesliga 1, the top football league in Germany, over the 2010/11 through 2014/15 seasons at the match level. The data set contains 1,251 football matches. Here, we only look at the bets made before the play started ("pre-play transactions"). The outcome variable of interest is the total volume of bets matched, in Pounds, for specific Betfair betting markets. We focus on two betting markets: bets that more than 2.5 goals will be scored in the football match (*Over 2.5*), and bets that fewer than 2.5 goals will be scored in the football match (*Under 2.5*).

We augmented the Betfair transactions data with information on match outcomes. We obtained the grade on a 1 to 6 scale (with 1 as the best performance grade and 6 as the worst) for the referee in each match from *Kicker*, a popular German football magazine. These grades represent assessments of the performance of each referee in each match as research suggests very good grades to increase nomination chances to succeeding games (Frick. Gürtler & Prinz, 2009). Additionally, we obtained the name of the referee in each match, and referee performance data (red and yellow cards given, penalty kicks given) from the German Football Association.
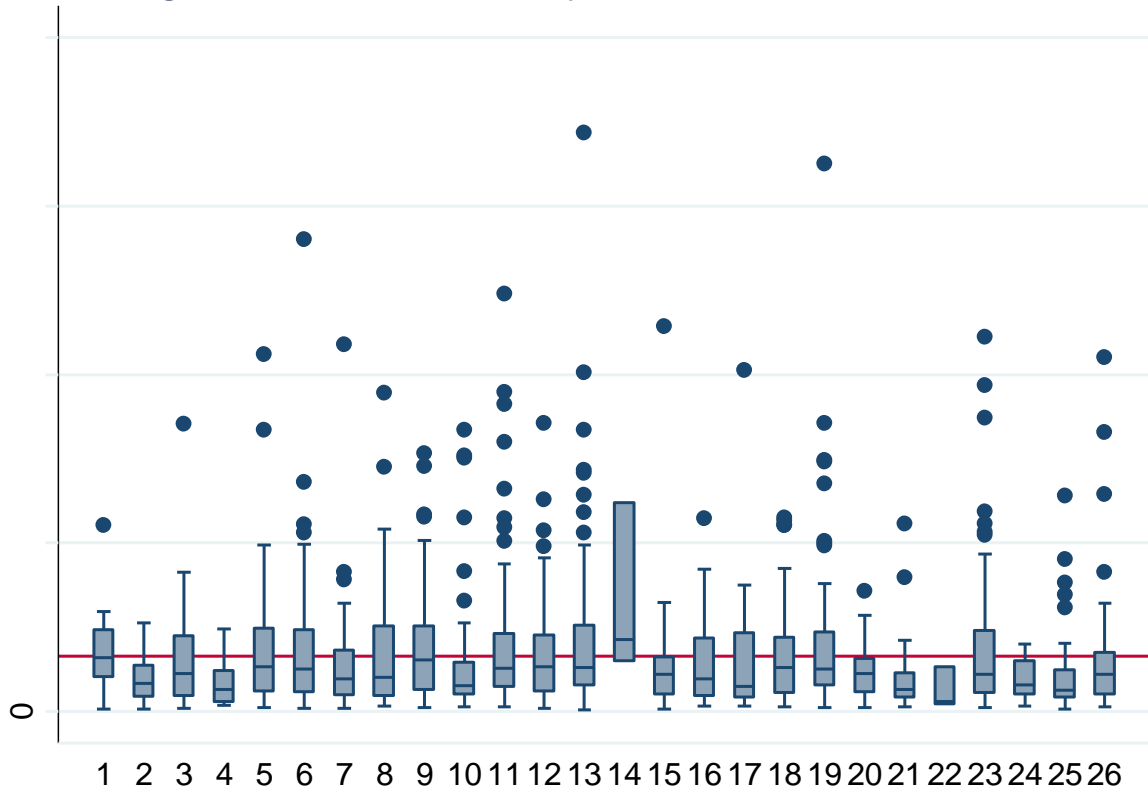
Table 1: Summary Statistics

	Mean	Standard Dev
Bet volume - over 2.5 goals (£)	33,103	36,904
Bet volume - under 2.5 goals (£)	22,456	38,221
Kicker Referee Grade	3.219	1.140
Home yellow cards	1.566	1.148
Home red cards	0.045	0.211
Home penalties	0.154	0.376
Home corners	5.421	2.900
Away yellow cards	1.967	1.230
Away red cards	0.054	0.230
Away penalties	0.104	0.328
Away corners	4.230	2.430

Table 1 contains summary statistics. The average volume of bets matched in the *Over 2.5* market was about 33 thousand Pounds and the average volume of bets matched in the *Under 2.5* market was about 22 thousand Pounds.

The football matches in the sample were officiated by 26 different referees. Table 4 shows the number of games officiated by each referee in the sample. The paper focuses on analyzing variation in bet volume in the *Over 2.5* and *Under 2.5* markets by referee. Figure 1 summarizes the variation of interest using box plots of bet volume in the *Over 2.5* market for each referee. The box identifies the 75th and 25th percentile of the distribution for each referee, the interquartile range (IQR), and the line identifies the median value. The whiskers identify the smallest and largest values within 1.5 IQR of the nearest quartile, while the dots above the top whisker represent extreme values and the red line is the sample mean in the *Over 2.5* market.

As shown in Figure 1, quite a bit of variability exists in bet volume in the *Over 2.5* market on Betfair across these 26 referees. The median value for each referee is lower than the average, indicating skew in the distribution for each referee.

Figure 1: Bet Volume by Referee - Over 2.5 Goals

The average *Kicker* grade for a referee was 3.2 on the 1 to 6 German "school mark" grading scale where 1 is the best possible grade and 6 the worst possible grade. The Kicker grades are assessed in 0.5 unit increments. A referee received the highest grade (1.0 or 1.5) in only 4% of the matches. About 60% of the matches in the sample resulted in the referee getting a grade of between 2.0 and 4.0 ("good", "satisfactory" or "sufficient").

On average, home teams received about 1.5 yellow cards and away team received almost two yellow cards. Red cards were quite rare in the sample. Away teams also received more red cards on average. Penalty kicks were relatively rare, occurring in only about 1 in 10 matches. Home teams were awarded more penalty kicks on average, 0.154 per game, than away teams, 0.104 per game, which is in line with the literature on home bias in football (Dohmen and Sauermann, 2016).

Empirical Model

The aim of this study is to estimate the impact of referees on betting volume, controlling for game characteristics and visible referee statistics. The reduced form models of the determination of bet volume at the match level for the *Over 2.5* and *Under 2.5* markets on Betfair take the form

$$VOL_{ijrws} = \alpha_h HT_i + \alpha_v VT_j + \alpha_w MW_w + \alpha_s SEAS_s + \gamma_r REF_r + \beta X_{ijrws} + e_{ijrws} \qquad (1)$$

where VOL_{ijrws} is the bet volume matched in one of the two exchange markets on a football match between home team i and visiting team j refereed by referee r in match week w of season s. Football matches occur in settings with substantial unobservable heterogeneity, as teams face different constraints and incentives in terms of the effect of a draw or win on the team's expected success.

We control for unobservable match-level heterogeneity using a number of fixed effects in the regression model. HT_i is a vector of home-team specific indicator variables that reflect unobservable heterogeneity in home teams. These variables reflect team popularity, roster composition, pitch characteristics, fan characteristics, and other unobservable home-team specific factors that affect bet volume on Betfair. VT_j is a vector of visiting-team specific indicator variables, MW_w is a vector of match week specific indicator variables and $SEAS_s$ is a vector of season-specific indicator variables. Boeri and Severgnini (2013) show that the probability of match fixing changes systematically over the course of a season. X_{ijrws} is a vector of match-specific characteristics, including variables reflecting observed referee performance.

e_{ijrws} is a random variable that captures all other factors that affect the matched bet volume in the *Over 2.5* and *Under 2.5* markets for each football match on Betfair. This random variable is assumed to be mean zero and possibly heteroscedastic.

Equation (1) also contains a vector of indicator variables for the referee in each match. Our sample includes 26 different referees. Two of these referees officiated only 6 Bundesliga 1 matches in our sample. The average referee in the sample officiated 96 matches and some refereed more than 130 matches. The vector of parameters on these referee-specific indicators, γ_r, are the parameters of interest here. These parameter estimates reflect the average effect of each referee on the volume of bets matched in the over/under 2.5 goals markets on Betfair. On average, it would be expected that the betting volume on the proposition that more or less than 2.5 goals are scored in a football match would not be systematically related to the identity of the referee of a match, after controlling for

other match-related factors, including unobservable heterogeneity related to teams, match weeks, and seasons. If the estimated parameter on a referee-specific indicator is statistically different from zero, then the bet volume on matches worked by that referee are higher than average. This could reflect match fixing.

X_{ijrws} is a vector of match-specific variables. We include the Kicker grade for the referee, and various variables reflecting referee performance in each match, including the number of red and yellow cards issued and the number of penalty kicks and corner kicks awarded, in this vector. The α's, γ and β are unobservable parameters to be estimated.

Results

Table 2 shows results from estimating Equation (1) using OLS, with the estimated standard errors corrected for heteroscedasticity using the standard White/Huber "sandwich" correction. The dependent variable is bet volume in the Betfair *Over 2.5* market in each match.

The baseline specification, Column (1), sets X_{ijrws} equal to zero and includes only home team, visiting team, match week, season, and referee indicator variables. This model explains 22.6% of the observed variation in bet volume in the *Over 2.5* market. Two of the parameters on the referee-specific indicator variables are significantly different from zero at the 5% level in this model, suggesting that bet volume was above average in this market for matches refereed by these referees compared to the other referees in the sample.

Table 2: Results - Over 2.5 goal betting volume

	(1)	(2)	(3)	(4)
Referee Grade	--	-1071	--	--
		(-1.18)+		
Home yellows	--	--	-967	-1003
			(-1.05)+	(-1.09)+
Home reds	--	--	-3349	-3521
			(-0.68)+	(-0.72)+
Home penalties	--	--	-1777	-1750
			(-0.64)+	(-0.63)+
Away yellows	--	--	-412	-412
			(-0.47)+	(-0.47)+
Away reds	--	--	-6640	-6648
			(-1.48)+	(-1.44)+
Away penalties	--	--	-3604	-3599
			(-1.15)+	(-1.13)+
Home corners	--	--	---	-301
				(-0.81)+
Away corners	--	--	--	459
				(1.02)+
# Significant referees	2	2	3	3
Observations	1251	1251	1251	1251
R^2	0.216	0.216	0.221	0.222

t-statistics in parentheses; * p < 0.1; ** p < 0.05; *** p<0.01; + not significant. Significance refers to p<0.05.

Model specification (2) adds the *Kicker* referee grade variable to the model. This assumes that the *Kicker* grade reflects whatever actions each specific referee took during the match, and ultimately whether such actions were viewed as reasonable or unreasonable. Model specification (2), like specification (1) explains 21.6% of the observed variation in matched bet volume on Betfair in this market. The estimated parameter on the Kicker referee grade variable is not statistically different from zero. Again, two of the parameters on the referee-specific indicator variables are significantly different from zero at the 5% level in this model for the same two referees, suggesting that betting was above average in this market for the matches they refereed.

Model specification (3) replaces the *Kicker* referee grade variable with variables indicating the number of red and yellow cards issued, and the number of penalty kicks and corner kicks awarded, in each match. These variables reflect a different mechanism through which referee actions affect the outcome of football matches as players are potentially being sent off and thus directly affect effort and style of play of the affected team Model specification (3) explains 22.1% of the observed variation in bet volume. None of the parameters on the card or penalty and corner kick variables are statistically significant.

The parameter estimates on the two referee-specific indicator variables that were statistically different from zero in model specifications (1) and (2) are also statistically different from zero in model specification (3). In addition, a third parameter estimate on a referee-specific indicator variable is significantly different from zero (p-value 0.054) in this model specification. The p-value on the test of the null hypothesis that the parameter estimate for this third referee is equal to zero was 0.057 in model specification (1) and 0.071 in model specification (2). This represents weak evidence that betting volume in the Betfair *Over 2.5* market was higher for this referee as well.

Including both the Kicker grade and the card and penalty kick variables in the same model produces similar results. The model explains 22.1% of the observed variation in bet volume, none of the parameter estimates on the referee performance variables are statistically different from zero, and the parameter estimates on the same two referee-specific indicator variables are significantly different from zero. The p-value on the test of the null that the third referee has a positive association with bet volume is 0.065 in this model.

Table 3 repeats the analysis above using bet volume in the Betfair *Under 2.5* market as the dependent variable. Model specification (1) includes only the fixed effects variables to control for unobservable team, match week and season level heterogeneity. Model specification (2) includes the Kicker referee grade and model specification (3) includes the referee outcome variables (yellow and red cards and penalty kicks). Like on Table 2, none of the parameter estimates on the referee performance variables are statistically different from zero, and the models explain about 22% of the observed variation in bet volume in this market.

Similarly to Table 2, two of the parameter estimates on the referee-specific indicator variables are statistically different from zero at the 5% level. Bet volume in the Betfair *Under 2.5* market was above average for matches officiated by these two referees, compared to the other referees in the

sample. One of these referees is different from the three referees identified in Table 2. However, the second referee associated with higher bet volume in the *Under 2.5* market is the referee identified as associated with a higher bet volume in the *Over 2.5* market in model specification (3) on Table 2.

Table 3: Results - Under 2.5 goal betting volume

	(1)	(2)	(3)	(4)
Referee Grade	--	-340.1 (-0.36)+	--	--
Home yellows	--	--	467.3 (0.49)+	451.3 (0.47)+
Home reds	--	--	-4197.8 (-0.83)+	-4462.9 (-0.88)+
Home penalties	--	--	1243.0 (0.43)+	1112.0 (0.39)+
Away yellows	--	--	-1340.5 (-1.49)+	-1326.2 (-1.47)+
Away reds	--	--	-8361.3 (-1.80)*	-8115.0 (-1.74)*
Away penalties	--	--	-3795.4 (-1.16)+	-3730.8 (-1.14)+
Home corners	--	--	--	-159.5 (-0.41)+
Away corners	--	--	--	404.7 (0.87)+
# Significant referees	2	2	2	2
Observations	1251	1251	1251	1251
R^2	0.216	0.216	0.222	0.222

t-statistics in parentheses; * p < 0.1; ** p < 0.05; *** p<0.01; + not significant. Significance refers to p<0.05.

Note that this result is quite plausible. Suppose that an individual has fixed a football match by inducing the referee to make calls that would lead to 3 or more goals being scored in the match. On a betting exchange like Betfair, a match fixer could either back bets in the *Over 2.5* market, or

lay bets in the *Under 2.5* market to profit from this outcome. This would tend to increase bet volume in either market.

The results on Tables 2 and 3 indicate that specific referees in Bundesliga 1 matches are associated with higher than average bet volume in the Betfair exchange markets for *Over 2.5* and *Under 2.5* trades. Since specific referees are identified as associated with higher bet volume in these markets, we next assessed the extent to which these referees perform differently than the other referees in the sample, in terms of measurable performance metrics.

Table 4 identifies the referees associated with higher than average bet volume in the Betfair markets by ID number. The two referees consistently identified as associated with higher than average bet volume in the *Over 2.5* market are Referee 11 and Referee 13. Recall that Robert Hoyzer was identified as a referee fixing matches, and subsequently convicted of fraud after making a number of high-profile, egregious calls, including awarding penalty kicks and issuing red cards to star players during a match between Hamburger SV and SC Paderborn. However, matches fixed by Ante Sapina, and matches fixed during the *Calciopoli* match fixing scandal in Italy, were only discovered to have been fixed long after they occurred, through sworn testimony during court cases or phone taps. None of the referees associated with unusually high over/under 2.5 goal bet volumes had particularly good or bad average *Kicker* grades over the period, so an objective assessment of their performance was that neither performed unusually well or unusually poorly. This is to be expected, since anyone attempting to influence match outcomes in sporting events must find ways to accomplish this goal without detection. The theoretical models of match fixing discussed above emphasize the important role played by the probability of detection in the decision to fix matches.

Note that Table 4 does not show the exact number of games refereed by each official, only an indicator for refereeing more or less than 100 matches in the sample. This is to avoid identifying specific referees.

Table 5 summarizes the propensity of the referees to issue red cards to players and award penalty kicks and corner kicks. In no case were more than 2 red cards or penalty kicks awarded to any team in a match in the sample.

ID #	N	# significant	Grade	Home yellows	Away yellows
		Table 4: Individual Referee Results – Means and Significant Parameters			
1	<100	0	4.12	1.23	1.38
2	<100	0	3.48	1.32	2.12
3	100+	0	3.42	1.86	2.14
4	<100	0	3.23	1.60	1.90
5	100+	0	3.29	1.61	1.79
6	100+	3 (Under 2.5)	3.25	1.61	1.89
7	<100	0	3.28	1.53	2.28
8	100+	0	3.21	1.68	2.15
9	100+	0	2.97	1.38	2.11
10	100+	0	3.50	1.79	2.15
11	100+	3 (Over 2.5)	3.36	1.62	1.80
12	100+	0	2.77	1.45	1.68
13	100+	3 (Over 2.5)	3.11	1.78	2.32
14	<100	0	3.33	1.33	1.67
15	100+	0	3.09	1.39	1.96
16	100+	0	3.30	1.35	1.88
17	<100	0	3.69	1.56	2.31
18	100+	0	3.27	2.05	1.93
19	100+	1 (Over), 3 (Under)	3.04	1.36	1.58
20	100+	0	3.19	1.45	1.78
21	<100	0	3.64	1.50	2.11
22	<100	0	3.83	0.67	1.00
23	100+	0	3.17	1.62	2.20
24	<100	0	3.34	1.32	1.87
25	<100	0	3.08	1.31	1.82
26	100+	0	3.09	1.73	2.01
Mean	96		3.22	1.57	1.97
+/- 1 Std dev.			4.36/2.08	2.71/0.42	3.19/0.75

Significance refers to p<0.05.

ID #	N	Red away	Penalty away	Red home	Penalty home	Corners away	Corners home
1	<100	0.00	0.23	0.08	0.15	3.38	6.15
2	<100	0.04	0.16	0.08	0.04	4.56	6.16
3	100+	0.02	0.10	0.00	0.04	3.92	5.45
4	<100	0.05	0.20	0.00	0.15	5.70	4.8
5	100+	0.03	0.18	0.05	0.16	4.90	5.68
6	100+	0.08	0.15	0.07	0.14	3.69	5.51
7	<100	0.09	0.11	0.13	0.17	3.89	4.79
8	100+	0.03	0.09	0.06	0.03	4.38	5.22
9	100+	0.04	0.18	0.04	0.03	4.28	5.00
10	100+	0.12	0.21	0.04	0.12	3.42	5.85
11	100+	0.05	0.08	0.02	0.08	4.28	5.66
12	100+	0.01	0.13	0.00	0.07	4.68	4.92
13	100+	0.03	0.12	0.07	0.07	4.62	5.85
14	<100	0.00	0.33	0.00	0.00	3.33	6.00
15	100+	0.04	0.12	0.07	0.07	3.96	5.42
16	100+	0.06	0.21	0.10	0.08	3.63	6.23
17	<100	0.06	0.19	0.00	0.25	3.69	4.88
18	100+	0.04	0.18	0.05	0.16	4.48	4.98
19	100+	0.06	0.22	0.06	0.10	4.00	5.09
20	100+	0.05	0.13	0.07	0.16	4.20	5.73
21	<100	0.00	0.19	0.00	0.08	4.67	4.44
22	<100	0.00	0.00	0.00	0.00	3.00	7.67
23	100+	0.01	0.17	0.07	0.12	4.86	4.94
24	<100	0.06	0.10	0.06	0.13	3.45	6.42
25	<100	0.04	0.22	0.06	0.18	3.45	6.1
26	100+	0.06	0.16	0.09	0.11	4.77	5.34
Mean		0.05	0.10	0.04	0.15		

Table 5: Individual Referee Results – Red Cards and Penalties Called per match

Robustness Checks

The regression results above use individual referee fixed effects to identify referees in matches with unusually high bet volume in the markets for over 2.5 goals scored and under 2.5 goals scored in matches. In each regression model, one referee is omitted to avoid perfect collinearity among the fixed effect parameters. These individual referee fixed effects reflect the average bet volume in matches refereed by each referee relative to the volume in the matches refereed by the omitted referee.

Table 6: Robustness Checks – Omitting Different Referees from Regression Models

| | Over 2.5 Goals | | Under 2.5 Goals | |
| | 5% Level | 1% Level | 5% Level | 1% Level |
Referee ID	# Significant / 25	# Significant / 25	# Significant / 25	# Significant / 25
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	16	10
7	0	0	0	0
8	0	0	0	0
9	0	0	0	0
10	7	1	0	0
11	0	0	0	0
12	0	0	0	0
13	12	3	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	5	0	13	1
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	0	0	0	0
24	0	0	0	0
25	0	0	0	0
26	0	0	0	0

Since these parameters reflect relative match volumes, the referee fixed effect could be sensitive to the omitted referee. To assess the extent to which the results are driven by the omitted referee, we re-estimated Equation (1) with the full set of covariates 26 times, systematically using a different omitted referee in each iteration. This results in 26x25=650 different estimated referee fixed effects parameters.

Table 6 summarizes the results of this systematic re-estimation omitting different referees from each iteration. The parameter estimates are random variables, and some number of them are expected to be statistically different from zero by chance. At the 5% significance level, about 32 of the parameter estimates would be expected to differ significantly from zero. At the 1% significance level, about 6 parameter estimates would be expected to be statistically different from zero. If the parameter estimates were statistically different from zero by chance, then these significant parameter estimates would be expected to be randomly distributed across referees in the sample, and not concentrated in a small number of referees.

The first two columns of Table 6 show the results for the betting on over 2.5 goals scored markets. 24 of the estimated parameters are statistically different from zero, slightly smaller than the number that would be expected to be significant by chance. However, these significant parameters occur for only three referees, Referees number 10, 13 and 19. These are the three referees identified on Table 4 as associated with statistically significant individual fixed effects.

Put another way, each referee has an individual fixed effect estimated in 25 different regression models. At the 5% level, each referee would be expected to have a statistically significant fixed effect in one of these models by chance. A referee having a statistically significant individual effect parameter 7 or 12 times out of 25 has substantially more significant parameter estimates than would be expected by chance. At the 1% level, only 0.2 out of 25 parameter estimates would be expected to be significantly different from zero by chance. From the second column of Table 6, having 3 statistically significant individual parameter estimates out of 25 is also more than would be expected by chance.

The second two columns on Table 6 show the results for betting on under 2.5 goals scored market. 29 of the estimated referee specific fixed effects parameters are statistically different from zero, about what would be expected. However, these statistically significant fixed effect parameters are for only two referees, Referees number 6 and 19. At the 1% level, 11 referee fixed effect

parameters are statistically different from zero, well above the number that would be expected by chance, and ten of these statistically significant parameter estimates are for referee number 6. Again, both these referees were identified on Table 4 as refereeing matches with unusually high betting volume in the over/under 2.5 goals scored market.

Taken together, the results suggest that at least three referees, numbers 6, 13 and 19, were associated with Bundesliga matches with unusually high volume in the over/under 2.5 goals scored market. The results on Tables 4 and 5 reflect referees identified by chance because the estimated referee fixed effects parameters are random variables.

Conclusion

We performed an *ex post* analysis of betting volume in Betfair markets for over/under 2.5 goals scored in Bundesliga 1 football matches to determine if variation in this bet volume is consistent with the idea that some referees might have been engaged in match fixing. The empirical analysis identified a small number of referees that were associated with larger than average volume in both the *Over 2.5* and *Under 2.5* markets. The *Kicker* referee grades, and observed average propensity of these referees to issue yellow and red cards, and to award penalty kicks, did not appear to be different from other referees in the sample, suggesting no unusual behavior on the part of these referees.

While our results are compelling, we do not claim to have identified match fixing in these Bundesliga 1 matches per se. Rather, we have identified anomalous patterns in Betfair exchange betting data that are consistent with the story of match fixing. From a theoretical point of view, a repeated engagement, rather than a unique involvement, in deviant behavior such as match fixing is consistent with the slippery-slope literature (Gino & Bazerman, 2009) and anecdotal evidence from the uncovered cases of match fixing discussed previously. A similar exercise was carried out by Wolfers (2006) using point spread betting data from US college basketball games. Other explanations exist for the patterns in bet volume documented in these data. Alternative explanations for the results in Wolfers (2006) have also been proposed (Bernhardt & Heston, 2010).

Since match fixing can be difficult to detect, especially match fixing by referees or team officials, it is important to develop methods for detecting match fixing based on observed data in betting

markets. Some betting on fixed matches clearly takes place with traditional bookmakers. In general, only data on odds set by bookmakers on specific sporting events are publicly available; data on bet volume on specific sporting events at individual bookmakers are not easily available. This makes the Betfair data analyzed here of notable interest

The analyses have some limitations. Referees are not totally randomly assigned to games. The Bundesliga reacted to the Hoyzer scandal by reducing the time between the announcements of referees and the actual games, reducing the time available to fix a game prior to all games in this sample. Veteran referees are assigned to decisive games more often than rookie referees are. While our empirical work captures match week effects, there might still be some heterogeneity between the matches in our sample. Since betting volume cannot be disentangled on the day of each bet, this study includes overall betting volume on each possible outcome.

References

Bernhardt, D., & Heston, S. (2010). Point shaving in college basketball: A cautionary tale for forensic economics. Economic Inquiry, 48(1), 14-25.

Boeri, T., & Severgnini, B. (2011). Match rigging and the career concerns of referees. Labour Economics, 18(3), 349-359.

Boeri, T., & Severgnini, B. (2013). Match rigging in Italian professional soccer: the economic determinants of corruption. In *Match-Fixing in International Sports* (pp. 101-112). Springer International Publishing.

Borghesi, R. (2008). Widespread corruption in sports gambling: fact or fiction?. Southern Economic Journal, 74(4), 1063-1069.

Bryson, A., Buraimo, B., & Simmons, R. (2011). Do salaries improve worker performance?. Labour Economics, 18(4), 424-433.

Carpenter, K. (2012). Match-Fixing—The Biggest Threat to Sport in the 21st Century?. International Sports Law Rev, 2, 13-23.

Di Tella, R., & Franceschelli, I. (2011). Government advertising and media coverage of corruption scandals. American Economic Journal: Applied Economics, 3(4), 119-151.

Dietl, H. M., & Weingärtner, C. (2012). Betting scandals and attenuated property rights–How betting related match fixing can be prevented in future. Institute for Strategy and Business Economics University of Zurich Working Paper Series ISSN, 1660-1157.

Dimant, E., & Deutscher, C. (2015). The Economics of Corruption in Sports: The Special Case of Doping. Edmond J. Safra Working Papers, (55).

Dimant, E., & Schulte, T. (2016). The Nature of Corruption: An Interdisciplinary Perspective. German Law Journal, 17, 53.

Dimant, E., & Tosato, G. (forthcoming). Causes and effects of corruption: What has past decade's empirical research taught us? A survey. Journal of Economic Surveys.

Dohmen, T., & Sauermann, J. (2016). Referee bias. Journal of Economic Surveys, 30(4), 679-695.

Dusha, E. (2015). Intermediated corruption. International Economic Review, 56(3), 997-1017.

European Gaming & Betting Association (2014). Sports betting: Commercial and integrity issues. Link: http://www.egba.eu/media/Sports-Betting-Report-FINAL.pdf

Feltes, T. (2013) Match Fixing in Western Europe. In *Match-Fixing in International Sports*. (pp. 15-30) Springer International Publishing.

Forrest, D. (2012). The threat to football from betting-related corruption. International Journal of Sport Finance, 7(2), 99.

Forrest, D., McHale, I., & McAuley, K. (2008). "Say It Ain't So": Betting-Related Malpractice in Sport. International Journal of Sport Finance, 3(3), 156.

Forrest, D., & Simmons, R. (2003). Sport and gambling. Oxford Review of Economic Policy, 19(4), 598-611.

Frick, B., O. Gürtler and J. Prinz (2009): Men in Black: Monitoring and Performance of German Soccer Referees, in: Dietl, H., E. Franck and H. Kempf (Eds.): Football – Economics of a Passion, Schorndorf: Hofmann, 309-321.

Gino, F., & Bazerman, M. H. (2009). When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. Journal of Experimental Social Psychology, 45(4), 708-719.

LaBrie, R. A., LaPlante, D. A., Nelson, S. E., Schumann, A., & Shaffer, H. J. (2007). Assessing the playing field: A prospective longitudinal study of internet sports gambling behavior. Journal of Gambling Studies, 23(3), 347-362.

Lockwood, R. I. (2008). Best Interests of the League: Referee Betting Scandal Brings Commissioner Authority and Collective Bargaining Back to the Frontcourt in the NBA, The. Sports Law. Journal, 15, 137.Paul, R. J., & Weinbach, A. P. (2010). Investigating allegations of pointshaving in NCAA basketball using actual sportsbook betting percentages. Journal of Sports Economics, 12(4), 432-447.

Preston, I., & Szymanski, S. (2003). Cheating in contests. Oxford Review of Economic Policy, 19(4), 612-624.

Pohlkamp, S. (2014). The impact of referees on match outcomes in professional sports: Evidence from the German Football Bundesliga. University of Hamburg working paper.

Rickman, N., & Witt, R. (2008). Favouritism and financial incentives: a natural experiment. Economica, 75(298), 296-309.

Rose-Ackerman, S., & Palifka, B. J. (2016). Corruption and government: Causes, consequences, and reform. Cambridge University Press.

Wolfers, J. (2006). Point shaving: Corruption in NCAA basketball. The American Economic Review, 96(2), 279-283.

Appendix: Betfair Football Betting Markets

Betfair generally offers markets for several types of bets on football matches. General market types include the correct final score of the match (0-0, 1-0, 2-0, etc.) match odds (home win, draw, away win), over/under a specific number of goals in the match (0.5, 1.5, 2.5) and a proposition involving both match odds and number of goals scored.

The table below summarizes the number of markets and average volume, in Pounds, of the betting volume for the general betting markets offered on Bundesliga 1 matches over the 2010/11 to 2014/15 seasons.

The Correct Score market type contains a large number of specific markets. These markets include all possible final score outcomes between 0-0 and 3-3 (16 different markets) and a "any unquoted" market for any other final match score.

Market type	Number of Markets	mean	st dev
Correct Score	31586	1099	2059
Match Odds	5571	85107	165974
Match Odds and Over/Under	1134	53	221
Over/Under 0.5 Goals	2444	5956	12318
Over/Under 1.5 Goals	4904	6402	10928
Over/Under 2.5 Goals	3718	23168	33224

In terms of average betting volume matched (the value of back and lay exchanges matched), the match odds market is by far the largest, with more than 85 thousand Pounds wagered on the average Bundesliga 1 match over this period. The second largest market in terms of volume is the market on over/under 2.5 goals scored in the match.