

# Do Punishment Schemes Induce Fairplay? Exploring the Hidden Decision Science of Booking in Spanish Premier League

Safaeduzzaman Khan\*

Department of Economics, University of Barisal, Barisal, Bangladesh.

## Abstract

Economic analysis of sport suggests punishment schemes are designed to induce fairplay. We provide an alternative explanation to that by assessing the law of market efficiency which holds across most of the social interactions involving the inherent mechanisms of exchange. To accomplish the claim, we construct an alternative method of testing our idea through borrowing and simulating a logically analogous setting from professional football leagues. We perform a lucid econometric analysis of our model with pertinent data. Then, we make further usage of statistical and quantitative techniques to throw more light on the problem. In line with our evidence, we argue that with presence of product sophistication and service superiority, market is left best without regulatory schemes. Similarly and finally, we alliterate that punishment scheme does not necessarily stimulate the emergence of fair-play.

**Keywords:** Incentives, Spanish Premier League, punishment schemes, fairplay.

**DOI:** 10.7176/EJBM/11-24-16

**Publication date:** August 31<sup>st</sup> 2019

## 1. Introduction

It is a fundamental notion of libertarian market economy that markets mechanize themselves with efficient allocation of resources. The opposite dichotomy lies on the dictum that markets are prone to failure, and thus they need a good “dose” of regulation which should suffice them against mis-coordination of demand and supply. In response to this restrictive procedural tenets, the proponents of markets argue that, it is still the market which ensures are all constellations of supply and demand still meet if it is armed with decent property right assignments. Consequentially, the debate moves on. But these articulated and simple notions help us to dissect diverse social incentive compatibility problems with economic insights.

As related to our interest, in almost all of the professional sports we observe that punishment schemes such as deduction of points, credits, profiles or scores, are existent so that players do not give in to unfair behavior and cheating. But ambiguous results exist on evaluating their theoretical and empirical underpinnings. In contrast, we think the main problem that slips into the boundary is the way of testing the conjecture. Both ambiguous and paradoxical philosophy to attack the problem only within the existing frameworks makes the problem either too much down to earth or too abstract and artificial. We provide an economic analysis to the stated conundrum to solve it on bases of economic insights.

To solve the problem with a new level of sanctity, we assist a lateral and novel approach with a quasi-experimental methodology. We suggest that law of market efficiency pervades through all forms of social interactions dominated by the notion of exchange. As implied, we forward the case of testing the conjecture from professional sports data. Hereby, we propose an alternative method to test the hypothesis: whether punishment in form of “regulation” really makes the “market” for sports matches function more “politely”.

## 2. Literature Review

Neale (1964) is one of the first papers to study the economic implications of professional sports. Most of the related literatures do not pay too much attention either to econometric specification issues or to the economic implications of the results. With respect to the latter point, not all the papers include the variables stated above as explanatory factors and not many take into account the interpretation of the punishment schemes involved. Rather, the studies try to rationalize the soporific effects of football match ticket prices with income and price elasticities, etc. Papers by Heilmann and Wendling (1976), Ferguson et al. (1991), Peel et. al. (1997), Marburger (1997) and Boyd and Boyd (1998) are among those which try, in the context of a profit maximizing behavior for the professional teams, as assumed in El Hodiri and Quirk (1971), to give a theoretical explanation for the usual empirical finding. This finding can also be explained in a context where teams have objective functions other than profits, as proposed by Sloane (1971). Demmert (1973) provides a systematic analysis of the similarity of a professional sports league as an economic organization.

## 3. Research Idea and Motivation

We see, that the fundamental assumption of market regulation states regulatory policies are some sort of

punishment schemes which are necessary as instruments of market fair-play. We construct an analogous model of football match statistics to enunciate the problem. Professional sports leagues stands as an equivalent way of simulating the actual engines of a market. We compare their intricate workings as follows:

In a “real-world” market there exists:

1. Profit maximization and cost minimization assumptions
2. Small number of sellers
3. Large number of buyers
4. Products differentiated at different degrees
5. Presence of regulatory authority and governments
6. Presence of regulatory mechanism

In professional sports leagues such as football leagues stand as a sound approximation to the “functionalities” of a real-world market. Analogously we can conceive our reasoning as the existence of:

1. Win maximization and loss minimization
2. Small number of club teams
3. Large number of audience
4. Each club team serving with significantly different quality of “product” in terms of football
5. Presence of regulatory bodies such as league committee, sports management and referees
6. Regulatory instruments in terms of bookings such as red card, yellow card, fine and suspensions

Thus, we let the Fair Play Point, which is aggregately and exogenously computed, as our main determinant of fairness, or put differently as the index of “market efficiency” of professional football leagues. Similarly, we divide the related effects into three parts:

Objective scheme variables: such as: matches won, number of goals, points. They stand as explaining the mechanical facts of the fair-play points. These variables should hold ambiguous a priori sign.

Punishment scheme variables: such as: number of yellow and red cards, number of suspensions etc. They explain the regulatory effects of the fair play point. These variables should exhibit inverse a priori sign i.e. they tend to contribute negatively to the explained variable.

Thus, we can identify all the related variables to build our model.

#### 4. Description of the Variables

In trying to explicate the relationship between the explained variable, Fair Play Point (fpp); a point awarded to each team for the fair spirit in their games, we have used the following explanatory variables:

1. wlr – Win Loss Ratio; which as the name suggests, is the ratio of a team’s total nos. of wins to losses in the season.
2. gf – Goals For; the total nos. of goals scored by the team.
3. ga – Goals Against; the total nos. of goals the team had scored against them.
4. gddmv – Goal Difference Dummy Variable; a dummy variable equal to 1 if the goal difference is positive and 0 if negative.
5. pts – Points; the total points scored by the team in the season.
6. sy – Single Yellow Card; the total nos. of single yellow cards penalized to the team during the season.
7. dy – Double Yellow Card; the nos. of occasions the team was punished with a double yellow card during the season.
8. r – Red Card; the nos. of red cards penalized to the team during the given season.
9. sdmv – Suspension Dummy Variable; a dummy equal to 1 if there were any suspensions (Players & Club’s Personnel) during the season and 0 otherwise.
10. abdmv – Audience Behavior Dummy Variable; another dummy equal to 1 if there were more than 2 “Mild” rating of audience behavior in the season and 0 if less than two.

Here, the variables denoted with 1-5 serial numbers are the objective scheme variables and the variables denoted with 6-10 serial numbers are the punishment scheme variables as justified previously.

#### 5. Data Collection

We assemble relevant data on the projected variables. With data from two seasons from the Spanish league i.e La-Liga seasons 2009-10 and 2010-11 we accomplish our ventures. In the season of 2009-10 La Liga, total 20 club-teams of significantly different performance ratings attended in 380 matches. Barcelona was the champion team and Xerez was the team with the poorest performance. In order to the team with highest Fair Play Point (fpp) to the lowest Fair Play Point (fpp) the ranks were: Real Madrid, Tenerife, Deportivo, Barcellona, Mallorca, Almeria, Atletico Madrid, Osasuna, Sporting de Gijon, Espanyol, Racing Santander, Villarreal, Athletic Bilbao, Getafe, Valladolid, Valencia, Sevilla, Xerez, Zaragoza, Malaga. In the season of 2010-11 La Liga, total 20 club-teams of significantly different performance ratings attended in 380 matches. Barcelona was still the champion team and this time Almeria was the team with the poorest performance. In order to the team with highest Fair

Play Point (fpp) to the lowest Fair Play Point (fpp) the ranks were: Barcellona, Mallorca, Racing Santander, Hercules, Real Sociedad, Deportivo La Coruna, Real Madrid, Villarreal, Almeria, Athletic Bilbao, Getafe, Sporting de Gijon, Atletico Madrid, Espanyol, Malaga, Osasuna, Sevilla, Levante, Valencia, Zaragoza. We collect all the data from the Official Spanish League websites and the respective Wikipedia web-pages. As we are performing an exhaustive modeling from all available meshes of the bulk data set, there exists no query of random sampling, data bootstrapping or problems involving missing and/or manipulated data. For computational purposes, we use the econometric package STATA with respective command lines and macros. The data sets are formatted and defragmented accordingly. Also, aggregating across the club-teams of different performance history and team-quality, ensures a sound “buffer process” to be encountered in the estimation process. So, we then expect any inherent diversion in performance to be mutually normalized and canceled out, through the interactions of high-performance and low-performance teams.

## 6. Methodology

For the team-wise data set for the club-teams, attending the 2010-11 La Liga season, we compute the

1. Ordinary Least Square (OLS ) regression equation
2. Modified OLS regression equation for the variables with largest partial effects

Then, to analyze the models further in terms of statistical validity and soundness, we conduct the following tests:

1. Breusch-Pagan-Godfrey (BPG) Test for heteroskedasticity
2. White Test for heteroskedasticity
3. Ramsey’s RESET Test for functional form misspecification
4. Variance Inflating Factor (VIF) as a measure of multicollinearity

We repeat the same procedural steps for the team-wise data set for the club-teams, attending the 2009-10 La Liga season also. In latter part of our analysis, we combine the data from the two seasons to create a pooled data set and then carry out our regression analysis on the pooled data set. Eventually, we perform the Chow test to detect whether there exists any structural change across the regression patterns resulting from the 2009-10 and 2010-11 data.

## 7. Model

If we take Fair Play Point (fpp) as an aggregate computed index, we can imagine it to be a function of the following variables

$$fpp = ( wlr , gf , ga , gddmv , pts , sy , dy , r , sdmv , abdmv )$$

The OLS regression equation will take the additive form

$$fpp = \beta_0 + \beta_1wlr + \beta_2gf + \beta_3ga + \beta_4gddmv + \beta_5pts + \beta_6sy + \beta_7dy + \beta_8r + \beta_9sdmv + \beta_{10}abdmv + u$$

where  $\beta_0$  stands for the intercept coefficient. This is regression model to be run with the data sets, with the aids of computational package.

Adjusting our model according to the a priori expectation:

$$fpp = \beta_0 + \beta_1wlr + \beta_2gf + \beta_3ga + \beta_4gddmv + \beta_5pts - \beta_6sy - \beta_7dy - \beta_8r - \beta_9sdmv - \beta_{10}abdmv + u$$

We expect the regressed model to be resulting in the stated form We theorized the expected signs of the slope coefficients  $\beta_1$  ,  $\beta_2$  ,  $\beta_3$  ,  $\beta_4$  and  $\beta_5$  pertaining to the variables wlr, gf, ga, gddmv and pts are ambiguous i.e. it is dependent on the internal dynamics of the model whether they contribute to the explained variable proportionately or inversely. Contrastingly, the expected signs of the slope coefficients  $\beta_6$  ,  $\beta_7$  ,  $\beta_8$  ,  $\beta_9$  and  $\beta_{10}$  variables sy, dy, r, sdmv and abdmv are negative i.e. they are likely to be contributing inversely to the explained variable fpp. If our a priori regression form overlaps with our a posteriori regressed form significantly, we are on strong grounds to establish our posited claim.

## 8. Data Analysis and Results

### 8.1 Analysis for season 2010-11

The OLS results for regressing fpp on wlr, gf, ga, gddmv, pts, sy, dy, r, sdmv and abdmv are now provided below:

	fpp
wlr	0.931
	(1.19)
gf	-0.128
	(0.42)
ga	0.113
	(0.42)
gddmv	-5.080
	(1.30)

pts	0.175
	(0.38)
sy	-1.067
	(13.63)**
dy	-0.149
	(0.19)
r	-1.223
	(1.06)
sdmv	-5.919
	(2.68)*
abdmv	-15.938
	(5.16)**
cons	239.673
	(9.72)**
R <sup>2</sup>	0.97
N	20

\* p<0.05; \*\* p<0.01

The model has a high overall R<sup>2</sup> of 0.97 or 97% and when we adjust it for the number of explanatory variables we get an adjusted R<sup>2</sup> of 0.95 or 95%. Despite the high R<sup>2</sup> value, looking at the results we can see that only the coefficients on the variables sy, sdmv, abdmv are statistically significant at the conventional level of significance (5%), while all other variables fail to register a significant t-value. But nevertheless all the variables have their expected apriori signs with wlr, ga, and pts being positive and gf, gddmv, sy, dy, r, sdmv and abdmv being negative. These effects are as expected as increasing wlr, ga, and pts should result in fpp increasing while increasing

gf, gddmv, sy, dy, r, sdmv and abdmv should lead to the opposite. Of these the largest significant partial effect is of abdmv (-15.94), followed by sdmv (-5.92) and sy (-1.06). Although gddmv has a practically large partial effect of -5.08, but its effect is statistically insignificant. The coefficients on wlr (0.93) and r (-1.22) maybe regarded as being practically significant but not statistically. The constant term of 239.67 has no interesting meaning in our application since setting all explanatory variables equal to zero makes little sense as, except for the dummies, no other variables have values close to zero.

Therefore in explaining fpp we find abdmv, sdmv, gddmv, sy, r and wlr to have the largest partial effects and so we may modify our regression to include only these as the explanatory variables. Below are the results:

	fpp
wlr	0.769
	(2.63)*
gddmv	-4.973
	(2.62)*
sy	-1.069
	(17.83)**
r	-1.411
	(1.79)
sdmv	-5.573
	(3.40)**
abdmv	-16.226
	(7.70)**
cons	248.727
	(38.59)**
R <sup>2</sup>	0.97
N	20

\* p<0.05; \*\* p<0.01

Looking at the results we once again find that sy, sdmv and abdmv still continue to be statistically and practically significant. But now gddmv and wlr are statistically significant at the conventional level, although their coefficients have decreased to some extent and that on r have increased somewhat and is significant at the 10% level. Both the normal R<sup>2</sup> and adjusted R<sup>2</sup> continue to be high with being above 96%.

From the correlation table we can see that gf and ga are not only highly and significantly correlated with each other but also with wlr. At the same time these three variables are also highly correlated with pts. And so such a high degree of correlation among the explanatory variables might lead to gf, ga and pts being statistically

insignificant in our original regression model. Dropping these three variables leads to wlr being significant at the 5% level in our latter constrained model.

### 8.1.1 Tests for Heteroskedasticity

We now return to our original OLS model and look at its various issues. We begin by running few tests for heteroscedasticity.

We carry out the following tests.

The BPG Test:

$$\chi^2(1) = 0.00$$

$$\text{Prob} > \chi^2 = 0.9481$$

Failure to reject the null of constant variance leads us to the conclusion that heteroskedasticity may not be a problem in this data set. Still we carry out another test for heteroskedasticity.

The White Test:

$$\chi^2(19) = 20.00$$

$$\text{Prob} > \chi^2 = 0.3946$$

Again as we cannot reject the null, we find no evidence against homoskedasticity and so we work under the notion of homoskedasticity.

### 8.1.2 Test for Functional Form Misspecification

Next we check for functional form misspecification using Ramsey's RESET test:

Ramsey RESET test using powers of the fitted values of fpp:

$H_0$ : model has no omitted variables

$$F(3, 6) = 0.68$$

$$\text{Prob} > F = 0.5934$$

As before we fail to reject the null hypothesis, which in this case is that the functional form is correctly specified. And so find no problems with functional form we have chosen and continue with our present model.

### 8.1.3 Test for Multicollinearity

Now we can look at the extent of multicollinearity in our model. For this purpose we use as a measure the Variance Inflating Factor or VIF. The mean VIF of 14.90 indicates that some of the explanatory variables in our model are highly collinear. Of them pts, gf and ga exhibits VIF of above 10 and so they contribute the most to our multicollinearity problem. A primary reason of this high degree of multicollinearity might be our small sample size of 20 teams in the given season. Though, as a matter of deterministic measure, we computed the results for all the available meshes of the 20 teams. So here, multicollinearity may not be causing a noteworthy problem in our model.

## 8.2 Analysis for season 2009-10

Again after regressing fpp on wlr, gf, ga, gddmv, pts, sy, dy, r, sdmv and abdmv for the data of the season 2009-2010 we obtain the following results:

	fpp
wlr	-0.462
	(0.95)
gf	0.280
	(1.31)
ga	-0.020
	(0.07)
gddmv	-6.236
	(0.70)
pts	0.014
	(1.07)
sy	-0.929
	(6.83)**
dy	-2.693
	(2.07)
r	-3.161
	(3.31)**
sdmv	-15.434
	(2.66)*
abdmv	-9.555
	(0.96)
cons	237.915

	(10.56)**
R <sup>2</sup>	0.93
N	20

\* p<0.05; \*\* p<0.01

Both the normal R<sup>2</sup> and adj R<sup>2</sup> values of the model are high indicating a good fit of the explanatory variables. The variables sy, dy, r, and sdmv now appear to be statistically significant at the conventional level and all the key variables retain their original signs as with the regression for season10-11. But the coefficients on wlr, gf and ga are now reversed, contrary to 10-11, as we expect more goals scored and more matches won, both should result in more competitive performance and hence should affect fpp negatively. Nevertheless r and dy are significant now (they were insignificant for 10-11) and additionally sy and sdmv also significant now. The largest coefficient now appears on sdmv (-15.43) followed by abdmv (-9.55) and gddmv (-6.23) although the last two are now insignificant. The coefficients on dy (-2.69) and r (-3.16) are also large but that on sy (-0.93) is relatively small. As before the slope coefficient has no interesting meaning in our application and so we provide no explanation for it.

### 8.2.1 Tests for Heteroskedasticity

In the next part of our analysis we carry out the same diagnostic tests we carried out for 10-11 season data.

The BPG Test:

$$\chi^2(1) = 0.46$$

$$\text{Prob} > \chi^2 = 0.4999$$

White Test:

$$\chi^2(19) = 20.00$$

$$\text{Prob} > \chi^2 = 0.3946$$

In both cases high p-values does not allow us to reject the null of homoskedasticity.

### 8.2.2 Test for Functional Form Misspecification

Ramsey's RESET test:

$$F(3, 6) = 0.34$$

$$\text{Prob} > F = 0.7956$$

Again failure to reject the null hypothesis leads us to conclude that the model may not suffer from wrong functional form.

### 8.2.3 Tests for Multicollinearity

The mean VIF of 3.00 indicates that the extent of multicollinearity among the explanatory variables is relatively low in our present case as compared to the data for 10-11.

### 8.3 OLS Regression Equation for Pooled Data Set

In the next part of our analysis we combine the data from the two seasons to create a pooled data set and then carry out our regression analysis on the data set. The obtained results are now provided below:

	fpp
wlr	-0.550
	(1.77)
gf	0.227
	(2.15)*
ga	-0.114
	(0.87)
gddmv	-6.047
	(1.64)
pts	0.010
	(1.10)
sy	-0.957
	(13.57)**
dy	-2.183
	(3.17)**
r	-2.751
	(4.70)**
sdmv	-11.617
	(5.17)**
abdmv	-14.006
	(4.73)**
cons	244.485

	(20.49)**
R <sup>2</sup>	0.92
N	40

\* p<0.05; \*\* p<0.01

Now we can see that six of our explanatory, are now significant at the conventional level of 5% and additionally gf is significant at 10% level. We also find that gddmv, sy, dy, r, sdmv and bdmv have negative partial effects in all three of our regressions and like that from the 09-10 data set, wlr and ga have negative effects while that of gf is positive (which do not contradict with our apriori expectations as mentioned above). However, the coefficient on pts is insignificant in all the 3 our considered cases. Next we check if the data from the two seasons follow the same regression pattern i.e. to test if there is structural stability in the data, we perform the Chow test. In carrying out the Chow test, we first pool the data from the two seasons together and run our regression as before and obtain the RSS (p). From our individual regressions for the two seasons we obtain SSR (09-10) and SSR (10-11) and then carry out the Chow test as usual.

### 8.3.1 Test for Structural Change Detection

Chow test:  
 Calc F Value: 0.94  
 P Value: 0.528

Therefore we fail to reject the null hypothesis that the two seasons follow the same regression patterns and so the same regression model can be used to fit both seasons.

We conclude this section by summing up our results so far. The model we have chosen to explain Fair Play Point (fpp) captures most of the variation in fpp, leading to the high R<sup>2</sup> values obtained in our regressions. Although some of the explanatory are statistically insignificant at the conventional level but nevertheless the key variables are found to be significant. Additionally our data set does not suffer from problems such as heteroskedasticity or wrong functional form and the regression model follows the same pattern for the two seasons under consideration.

## 9. Conclusion

Analyzing our model in relation to the computation, we observe that, our a priori expectations on the respective variables significantly match with the a posteriori model outcomes. So, our original claim on economic insights into the professional sports model that: punishment schemes do not necessarily induce fairplay, holds.

We also gather insights on the following views:

1. Professional sports events can work as natural “laboratories” on analyzing economics issues concerning incentive compatibilities.
2. Fairplay is categorically related to team performance. The better a team performs the higher quality of fairplay it is likely to deliver.
3. Punishment schemes do not necessarily incite fairplay. But, the presence of punishment schemes can have a remarkable “threat-effect” on fairplay.
4. If punishment schemes are properly designed and kept adjustable to reviews, the more likely it is to contribute positively to fairplay.
5. Incentive compatibility is negatively related to punishment schemes.

Our model addresses and investigates on the economic foundations of the behavioral phenomena associated with professional sports events. We hope these results can aid us to formulate further interesting questions to be answered through economic way of thinking and deciphering similar scientific investigations decisively.

## References

- Baimbridge, M., Cameron, S. and Dawson, P. (1996) Satellite Television and the Demand for Football: A Whole New Ball Game?, *Scottish Journal of Political Economy*, 43, 317-333
- Boyd, D.W. and Boyd, L.A. (1998) The Home Field Advantage: Implications for the Pricing of Tickets to Professional Team Sporting Events, *Journal of Economics and Finance*, 22, 169-179
- Cairns, J.A., Jennett, N. and Sloane, P.J. (1986) The Economics of Professional Team Sport: a Survey of Theory and Evidence, *Journal of Economic Studies*, 13, 3-80
- Demmert, H.G. (1973) *The Economics of Professional Team Sports*, Lexington Books, Lexington, Massachusetts
- El Hodiri, M. and Quirk, J. (1971) An Economic Model of a Professional Sports League, *Journal of Political Economy*, 79, 1302-1319
- Falter, J.M. and Pérignon, C. (2000) Demand for Football and Intramatch Winning Probability: An Essay on the Glorious Uncertainty of Sports, *Applied Economics*, 32, 1757-1765
- Ferguson, D.G., Stewart, K.G., Jones, J.C.H. and Le Dressay, A. (1991) *The Pricing of Sports Events: Do Teams*



- Maximize Profit?, *Journal of Industrial Economics*, 39, 297-310
- Heilmann, L. and Wendling, W.R. (1976) A Note on Optimum Pricing Strategies for Sport Events, in *Management Science in Sports*, (eds.) R.E. Machol and S.P. Ladany, North-Holland Publishing Company, Amsterdam, pp. 91-99
- Marburger, D.R. (1997) Optimal Ticket Pricing for Performance Goods, *Managerial and Decision Economics*, 18, 375-381
- Neale, W. C. (1964) The Peculiar Economics of Professional Sports. *Quarterly Journal of Economics*, 78, 1-14
- Peel, D. and Thomas, D. (1997) Handicaps, Outcome Uncertainty and Attendance Demand, *Applied Economics Letters*, 4, 567-570
- Schofield, J.A. (1983) Performance and Attendance at Professional Team Sports, *Journal of Sports Behaviour*, 6, 196-206
- Sloane, P.J. (1971) The Economics of Professional Football: The Football Club as a Utility Maximizer, *Scottish Journal of Political Economy*, 17, 121-146
- Szymanski, S. (1997) The English Football Industry: Profit, Performance and Industrial Structure, *International Review of Applied Economics*, 11, 135-153
- Welki, A.M. and Zlatoper, T.J. (1994) US Professional Football: The Demand for Game-Day Attendance in 1991, *Managerial and Decision Economics*, 15, 489- 495