The Journal of Gambling Business and Economics 2021 Vol 14 No 1 pp 57-65

AN EFFECTIVE APPROACH FOR EXPLOITING THE INEFFICIENCIES OF THE ITALIAN FOOTBALL BETTING MARKET

Alesina Davide
PhD Student of Economics Faculty, National Research University Higher School of Economics, Pokrovsky Boulevard, 11, 109028, Moscow, Russian Federation.
+39 380 143 8910 / +39 349 082 5088,
dav.alesina@gmail.com / dalezina@hse.ru

ABSTRACT

In this paper, the author analyses the main inefficiencies of the Italian Serie-A betting market: in particular, using previous findings on the favourite-long shot bias, the team scoring home advantage and the league winner scoring advantage, a strategy aimed to outperform the market will be presented. By testing the betting scheme over the past 19 seasons, it was possible to obtain an average return per bet of 9.31%, with 16 of the 19 years resulting in positive financial returns. In order to verify the computations, different betting odds databases were used. The results obtained are particularly significant for two reasons: firstly all computations were performed on market average coefficients, leaving on the table an additional 3–4% of profit, which could be obtained by using best coefficients among bookmakers and secondly compared to market benchmarks and other betting strategies, the net profit is considerably higher. In particular other three strategies were used as a benchmark: the first one uses the favourite-long shot bias, the second one the Home factor and the third one the league favourite team advantage. All these betting schemes performed more poorly, with the second best strategy scoring on average 6% worse. After analysing possible future improvements, in the final section the author describes how the findings of this paper may be applied to other betting markets, such as different football leagues, basketball, hockey and tennis.

Key words: Betting markets; Sports forecasting; Market efficiency; Favourite-longshot bias; Football betting market; Gambling market;

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

INTRODUCTION

The online betting market has become a $500 billion business, which offers every day new betting opportunities in various sport competitions all over the world. Nowadays, bookmakers are organized in teams of data science specialists, in order to build sophisticated models, which have the aim of predicting sport outcomes and maximizing their profits. In the past years many studies have been conducted, in order to test the betting markets efficiency and find strategies, which are able to guarantee long term profit. In particular the main proposed models can be divided into the following categories: (i) probability models aimed to use available data to outperform bookmakers predictions (Dixon & Coles, 1997; Vlastakis, Dotsis, & Markellos, 2008), (ii) arbitrage strategies and odds bias exploitation (Kaunitz, Zhon & Kreiner, 2017; Ashiya, 2015; A. C. Constantinou, Fenton, & Neil, 2013; A. Constantinou & Fenton, 2013), (iii) expert prediction models (Forrest, Goddard, & Simmons, 2005) and (iv) group forecasting models (García, Pérez & Rodríguez, 2016). These different types of approach, usually are tested on the most popular sports, such as football, hockey, basketball, which have a historical betting database, allowing scholars to analyze several years of betting data. In some cases the betting strategies produce a positive economic return, although many times these ideas are tested on only one or two years of betting data, which is not enough to categorize a strategy as a winning one, since it is possible to build several betting schemes, which in some particular years, by pure chance, will yield a positive return.

In this paper the focus will be on the Italian football betting market of the Serie A League, with the aim of proposing a betting strategy which outperforms the market across several years.

INEFFICIENCY OF THE ITALIAN FOOTBALL BETTING MARKET

A first study performed in 2018 (Angelini et al, 2018) proposes a model, which uses the bookmaker’s forecasting error as independent variable, regressing it over the bookmaker’s commission and the winning estimated probability (calculated as the inverse of bookmakers coefficients). Using this approach the authors were able to find possible market inefficiencies in different Leagues and to estimate the coefficient range in which the market inefficiency lies. The analysis suggests, that in Italian football betting market for Serie A is characterized by favorite-long shot bias and coefficients less than 1.67 may lead to positive returns.

ITALIAN FOOTBALL MARKET “HOME ADVANTAGE”

The second important information, which it is possible to use is the one, obtained by (Baio & Blangiardo, 2010), where a Poisson regression has been used in order to predict value for the Home advantage, team attack and defense. This study suggests that the Italian football betting market could be characterized by a high home factor and a great difference in probability score between league winner and the rest of the teams. This means that playing in the home stadium may have a tangible positive effect and the team, which win the National League can potentially has a scoring probability much higher than the rest of the teams.

THE MODEL

The theoretical information analyzed can be organized in order to build a logical strategy to constantly beat bookmakers in Italian Serie A betting market, in particular our model will be based on:

1. *Favorite-long shot bias.* This information suggests us to bet only on favorites, ignoring the other coefficients, since it was empirically demonstrated their non-convenience.
2. *Home advantage.* Knowing that playing in the home stadium may have a bigger impact than in other betting markets, it seems logical to try to build a strategy where we bet only on teams which play in their own Stadium.
3. *Championship winner attack advantage.* Since it seems that the Championship winner in Italian League has a significant edge, with respect to other teams, betting on this team only could lead to significant profit.

Logically we would like then to find a way of exploiting all these three points in our strategy. The first one asks to bet on favorite and this information can be easily acquired just by looking at bookmakers coefficients, since it has been proved how they represent the best probability estimators for games outcome. The second point it is even more easy to use, since it is immediate to check which team is playing in the home Stadium for each match. The third and last point is the most difficult to use: in fact it is impossible a-priory to know which team will win the Championship. In order to bypass this problem we had two ideas: the first one is to use bookmakers predictions over the League winner, before the Championship starts and the second one consists in using official ELO rating of the teams before the start of the League. Both options seem reasonable, although the first one was impossible to adopt, since we did not find a way to retrieve the hysterical data from bookmakers on League favorites. For this reason we opt to use ELO rating of the day before the League starts, since they are easily retrievable from the official ELO rating website (*http://clubelo.com/).* In any case, it is highly likely that the team with best ELO rating is also the bookmaker favorite for winning the League Title, and so the two methods actually bring the exactly same results.

APPLYING THEORY TO REAL DATA

In order to apply the strategy proposed, it is possible to follow these steps:

1. *Download historical data.* As in other work in this field, the historical data

from *Football-Data.co.uk* (https://www.football-data.co.uk/) have been used, since they collect different bookmakers odds across several seasons: in particular we used all data available on their site, from season 2000/2001 up to season 2018/2019.

2. *Download historical ELO rating.* From official ELO rating website (*http://clubelo.com/)* it has been possible to download ELO rating off all teams, which participated in each season of the Italian Championship.
3. *Select favorite team.* Using ELO ratings, it is possible to select the expected League favorite team, simply by finding out the maximum ELO score in each season. In table 1 we reported the favorite team for each year, with its corresponding ELO rating.
4. *Select matches according to the strategy.* It is possible then to select only home matches of the League favorite, giving to us 17 matches in seasons 2000/2001–2003/2004 and 19 matches in seasons 2004/2005–2018/2019. This difference is due to the fact that initially Italian Championship was played by 18 teams and afterwards it was decided to switch to 20 teams.

Table 1. Favorite team according to ELO rating system. For each of the 19 seasons analyzed the best team has been selected with its corresponding ELO rating.

Season	Team	ELO Rating	Season	Team	ELO Rating
2000/2001	Lazio	1885	2010/2011	Inter	1922
2001/2002	Juventus	1859	2011/2012	Milan	1773
2002/2003	Juventus	1870	2012/2013	Milan	1799
2003/2004	Juventus	1894	2013/2014	Juventus	1826
2004/2005	Milan	1857	2014/2015	Juventus	1892
2005/2006	Milan	1898	2015/2016	Juventus	1921
2006/2007	Milan	1943	2016/2017	Juventus	1926
2007/2008	Inter	1890	2017/2018	Juventus	1945
2008/2009	Inter	1890	2018/2019	Juventus	1959
2009/2010	Inter	1865			

RESULTS

Finally the strategy described above has been implemented on the 19 seasons data, in parallel with three benchmarks strategies: the first one uses the findings presented in the first section on favorite-long shot bias and bets on all coefficients less than 1.67, the second one uses the Home advantage and bets on all teams playing in the home stadium, the third one uses the potential advantage of the favorite team and bets on all its matches. In order to compare these different strategies we decided to use the average bookmakers odds and not the maximum one, as in some other studies for two reasons: firstly using the average coefficient on the market will give a more realistic estimation of the real financial return, and secondly this paper is not based on arbitraging strategies, so it is much more important to see how a strategy behaves in comparisons with some benchmarks, than maximizing to the extreme its profit. Moreover using the best available coefficients on the market will lead to an improvement of 3–4% homogeneously, independently from the strategy.

For practical reasons, in the following lines the name S_1, will be used to describe the average return in a given period of the strategy proposed in this paper, S_F for the favorite-long shot bias benchmark strategy, S_H for the home advantage benchmark strategy, S_W for the winner benchmark strategy.

Computing the strategies profits yielded an average a profit of 9.31% for S_1, a loss of –0,67% for S_F, a loss of –8.64%% for S_H and a win of 3.57% for S_W. The strategy proposed in this paper is more volatile that the benchmarks, because it play on fewer games per year and so the random outcome of one match may affect the overall strategy output of the given year. Although the higher volatility, S_1 strategy performed better than all benchmarks in 13 of the 19 years of our analysis, generating losses in only 3 years. Below in table 2 the results of each strategy in the past 19 years.

Table 2. Strategies returns over the past 19 years. Comparison between average returns per years of strategies S1 in comparison to market standards.

	S_1	S_F	S_H	S_W
2000/2001	7.82%	–8.07%	–14.08%	1.94%
2001/2002	0.65%	–5.68%	–14.73%	–13.79%
2002/2003	13.41%	–1.22%	–9.86%	6.68%
2003/2004	1.29%	0.79%	–23.26%	–4.13%
2004/2005	–24.74%	–11.24%	–14.02%	–4.74%
2005/2006	30.00%	–2.72%	–12.52%	8.16%
2006/2007	–18.42%	3.01%	–14.08%	–27.87%
2007/2008	5.74%	–4.00%	–7.67%	–1.68%
2008/2009	6.37%	2.59%	0.34%	5.61%
2009/2010	10.00%	–1.87%	–2.40%	–3.61%
2010/2011	21.42%	–6.36%	–1.61%	2.18%
2011/2012	–13.63%	–3.93%	–3.95%	0.26%
2012/2013	9.68%	10.96%	–4.58%	–6.66%
2013/2014	32.32%	3.72%	–2.37%	26.50%
2014/2015	25.00%	–15.23%	–12.79%	17.08%
2015/2016	22.58%	–0.01%	–3.58%	20.50%
2016/2017	28.53%	5.33%	0.38%	7.58%
2017/2018	7.00%	6.86%	–17.54%	18.16%
2018/2019	10.42%	9.55%	–13.55%	11.00%
AVERAGE	***9.31%***	***–0.67%***	***–8.64%***	**3.57%**

COMPARISON RESULTS OBTAINED WITH A DIFFERENT BETTING DATABASE

Due to the high return of the proposed betting scheme, we decided not to give by granted the odds database used in the previous part for computing strategies average returns. In order to double-check our results we decided to test the strategy on historical odds available on Parimatch website (*https://parimatch.com/*). We will compare data from season 2012/2013 up to 2018/2019, since historical data of previous seasons are not available on Parimatch. As we can see from table 3 the computations show little different from the results obtained in the previous section: in fact the average betting return in the analyzed seasons is 17.24%, which is only 2.12% less than the result obtained using *Football-Data.co.uk* database in the same years (19.36%). This small difference is only due to slightly higher betting margin for Parimatch, moreover all singular year returning rates and singular coefficients are almost identical, proving the goodness of data used.

Table 3. Comparison between *Football-Data.co.uk* and *Parimatch* databases. Returns of strategy S1 in the past 7 seasons using different data sources. The results are very similar and show no significant difference.

	2013	2014	2015	2016	2017	2018	2019
Football-D	9.68%	32.32%	25.00%	22.58%	28.53%	7.00%	10.42%
Parimatch	4.00%	29.53%	20.90%	20.26%	28.00%	8.90%	9.11%

DISCUSSION

The results obtained with the strategy presented in this paper naturally give raise to two additional questions: the first one investigates the possible reasons of negative economic returns in three seasons (2004/2005, 2006/2007, 2011/2012), while the second one focuses on the applicability of this strategy in other betting markets. In the following lines we will discuss about these two topics.

POSSIBLE REASONS, WHICH INFLUENCE STRATEGY OUTCOMES IN DIFFERENT YEARS

In this section will be analyzed deeper the method for choosing the League favorite, since by using ELO ratings and/or bookmakers coefficients, it is possible, not only to choose the favorite team, but also to understand if the chosen team is way higher rated than the rest of the concurrence, or if there is a more flat distribution among teams. In order to test this hypothesis we will introduce two measurement parameters: *D2* is defined as the rapport between the highest ELO rating and the second one in the field, while *D0* is defined as the rapport between the highest ELO rating and the average of all ELO rating among the clubs competing in the chosen season.

$$D2 = ELO_1 \;/\; ELO_2$$

$$D0 = ELO_1 \;/\; ELO_{avg}$$

In this way we can see if either the difference in rating between first and second club or the difference between the first and the average rating, has an influence on the strategy outputs.

Below we presented a table, which reports the values of D2 and D0 for all the 19 seasons analyzed, with corresponding graphs, plotting on the x-axis D0/D2 and on the y-axis the strategy output. From the data available there is no evidence of positive and/or negative correlation between the variables and so this additional information doesn't seem to help in the analysis.

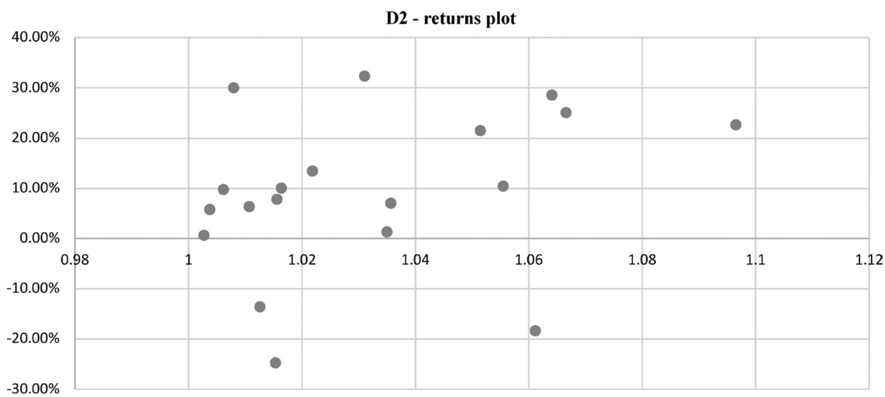

Figure 1. Correlation between D2 and strategy returns. The scatter plot show no significance correlation between the difference in ELO rating between the first and the second teams, with the corresponding yearly returns.

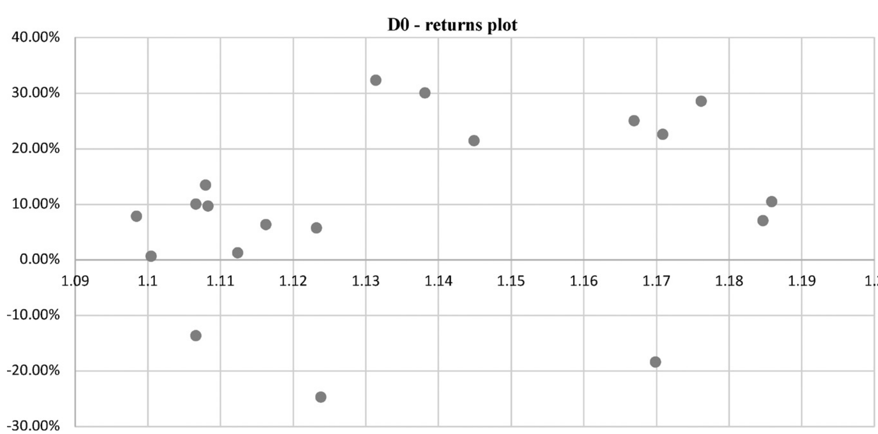

Figure 2. Correlation between D0 and strategy returns. The scatter plot show no significance correlation between the difference in ELO rating between the first team and the average ELO rating in the championship, with the corresponding yearly returns.

APPLYING THE METHOD TO DIFFERENT LEAGUES

In the future it will be interesting to apply the same, or similarly built strategies to other football leagues, focusing on the market efficiency analysis and the attack-defense-home advantage regression, in order to identify potential market inefficiencies to exploit. Strategies similar, but not identical to the one presented in this paper must be generated separately for each potentially inefficient market, since the inefficiency found in the Italian betting market of Serie A for *Home advantage* may work the opposite way in a different championship.

There are already some studies, which analyses the market efficiency, not only of football betting markets, but as well of others sports such as Tennis, Hockey, American Football, Basketball, and they potentially be a starting point for following investigations.

CONCLUSION

In this paper was presented a betting strategy for Italian football betting market for the Serie A championship, using data from 2000/2001 up to 2018/2019 season. The total matches in which the strategy bets are 353, with an average profit of 9.31%, using the average coefficient within bookmakers. This result is particularly significant due to the unconventionally high economical return across a great number of season, both in comparison to past published strategies and market benchmarks. Moreover the computations presented in this paper used only average bookmakers coefficients, and so they even don't account of an extra profit in the range of 3–4%, which could be added, simply by using the best available odds on the market.

In the future we plan to use a similar approach to investigate other football betting markets and as well experiment this methods with sports such as Hockey and Basketball.

REFERENCES

G Angelini and L De Angelis (2018) Efficiency of online football betting markets. 35(2) International Journal of Forecasting 712–721

M Ashiya (2015) Lock! Risk-Free Arbitrage in the Japanese Racetrack Betting Market. 16 (3) Journal of Sports Economics 322–330

G Baio and M Blangiardo (2010) Bayesian hierarchical model for the prediction of football results. 37(2) Journal of Applied Statistics 253–264

A Constantinou, N Fenton and M Neil (2013) Profiting from an inefficient association football gambling market: Prediction, risk and uncertainty using Bayesian networks. 50 Knowledge-Based Systems 60–86

A Constantinou and N Fenton (2013) Profiting from arbitrage and odds biases of the European football gambling market. 7(2) The Journal of Gambling Business and Economics 41–70

M Dixon and S Coles (1997) Modelling Association Football Scores and Inefficiencies in the Football Betting Market. 46 (2) Journal of the Royal Statistical Society. Series C, Applied Statistics 265–280

ELO raiting (2000–2019). Italy, retrieved from: *http://clubelo.com/IT*

Football-data-co (2000–2019). Italy, retrieved from: *https://www.football-data.co.uk/*

Forrest, Goddard and Simmons (2005) Odds-setters as forecasters: The case of English football. 21(3) International Journal of Forecasting 551–564

L Kaunitz, S Zhon & J Kreiner (2017) Beating the bookies with their own numbers - and how the online sports betting market is rigged. Online at: *https://www.researchgate.net/publication/320296375*

Parimatch historical data (2012–2019) retrieved from: *https://parimatch.com/*

L Perez, P Rodriguez and J Garci (2016) Forecasting football match results: Are the many smarter than the few? MPRA Paper No. 69687. Online at https://mpra.ub.uni-muenchen.de/69687/

N Vlastakis, G Dotsis and R Markellos (2008). Nonlinear modelling of European football scores using support vector machines. 40 (1) Applied Economics 111–118