Economics Bulletin

Volume 41, Issue 3

On the efficiency of online soccer betting markets: a new methodology based on symbolic series

Salvatore Caruso
Zanasi & Partners

Giuseppe Pernagallo
Collegio Carlo Alberto, University of Turin

Abstract

We propose a new methodology based on symbolic series to test the weak form of efficiency for soccer betting markets. This purpose was accomplished by introducing a definition of returns for betting markets and then creating binary series using the notion of excess returns. We present an application of the method using Serie A and Premier League data.

Citation: Salvatore Caruso and Giuseppe Pernagallo, (2021) "On the efficiency of online soccer betting markets: a new methodology based on symbolic series", *Economics Bulletin* , Vol. 41 No. 3 pp. 1451-1460
Contact: Salvatore Caruso - salvatore.caruso@zanasi-alessandro.eu, Giuseppe Pernagallo - giuseppe.pernagallo@carloalberto.org.
Submitted: June 16, 2021. **Published:** September 17, 2021.

1. Introduction

Soccer betting markets have gained momentum in recent years (Vlastakis et al., 2009). Betting markets, especially European ones, are believed to be transparent (Oikonomidis et al., 2015) and suitable for many investors. Studying efficiency for this market is important because inefficient markets can create relevant profit opportunities. Many works investigated the issue using mainly regression models (e.g., Vlastakis et al., 2009, Oikonomidis et al., 2015, Angelini and De Angelis, 2019), here we propose an alternative and faster methodology based on binary series. We introduce a notion of returns for betting markets, then we propose a procedure to generate binary symbolic series for excess returns. Finally, using runs tests and entropy one can assess the randomness of the series and its efficiency. We tackle the weak form of market efficiency, so the idea is that current prices reflect all the information contained in historical prices, ruling out the possibility of achieving excess returns using technical analysis. We propose an application to Serie A (the Italian major league) and Premier League (the English major league) data finding evidence in favour of efficiency.

2. Methodology

2.1. Data generating process

In order to test for the efficiency of betting markets, we propose a parallelism with financial markets (for a definition of financial returns, see Pernagallo and Torrisi, 2019). In finance, one way to test for the weak form of efficiency is to dichotomise the series of financial returns and then apply statistical techniques for binary series (Pernagallo and Torrisi, 2020). In betting markets, market prices are represented by bookmaker odds. For the sake of simplicity, we consider the easiest bet possible. So, a gambler can bet on three mutually exclusive outcomes: home team wins (H), draw (D), away team wins (A). Let X_i be a random variable, which assumes value of one whenever the outcome of bet i is successful, for example the bet is on H and the match ends with a win for the home team. Following Angelini and De Angelis (2019), X_i can be modeled as a series of Bernoulli trials so that $X_i|I_i \sim Bin(1, p_i)$, where I_i denotes the hypothetical information set that contains all the information in the universe, and p_i is the probability of success of bet i. Then, the return function for the gambler, R_i, is given by

$$R_i(K) = \begin{cases} K * odds_i - K & \text{if} \quad X_i = 1 \\ -K & \text{otherwise} \end{cases} \tag{1}$$

where K is the capital invested in bet i, and $odds_i$ are the odds for the chosen symbol (H,D,A) in bet i. Note that R_i is a function of K only. Indeed, the gambler can only choose the amount of money to invest, the odds are exogenously determined by the bookmaker and are assumed, for simplicity, to be fixed.

Without loss of generality, assume that $K = 1$, so that the return from a bet is simply given by $odds_i - 1$, whereas a loss is given by -1. One major problem related to this definition of R_i is that the probability distribution of the series of returns is a mixture between a point mass at -1 and positive values on the real line, in other words the set of possible outcomes can take values on on $\{-1\} \cup [0, +\infty)$. This fact makes troublesome to test the weak form of efficiency with traditional tools (such as variance ratios, Hurst exponents or autocorrelations). We propose a simpler way by using binary series. The series of R_i can be dichotomised as follows. One implication of the weak form of efficiency is that investors (in our case gamblers) cannot obtain above-average returns (for simplicity, called excess returns in this work) (Pernagallo and Torrisi, 2020). We can use this definition to build the series

$$D_i(K) = \begin{cases} 1 & \text{if} \quad R_i \geq \mu \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where μ is the average return for the considered period. Once we have obtained the symbolic series, we can operate with statistical tools for binary variables. In this work we use runs test and entropy.

2.2. Data

Data were gathered on the website *Football-Data.co.uk*. We selected the odds for three clubs of Serie A and two clubs of the Premier League in a period of 10 seasons, from 2009/2010 to 2018/2019. We considered only Serie A and Premier League matches for pricing homogeneity (we excluded, for example, Champions League matches). In order to have heterogeneity in the sample, we chose a low-ranking club (Chievo), a medium-ranking club (Lazio), and a high-ranking club (Napoli) for Serie A. For the Premier League, we collected data for Chelsea (high-ranking) and Everton (medium-ranking), but we could not collect data on low-ranking clubs because they were relegated at some point during the 10 seasons time period. We collected the odds for home and away matches of these clubs from 12 European bookmakers. Many of these bookmakers offer similar odds, but we have included them as a robustness check. The analysis is conducted separately for home and away matches because home teams have an advantage in terms of supporters and this influences prices. Therefore, each season has 19 home matches and 19 away matches. The series have different lengths, from 190 to 57 observations, because some bookmakers did not quote all the matches.

3. Results

The underlying null hypothesis for runs test is that the series is random. Table 1 and Table 2 show the results of the test for the whole series. Few series are significant at 5%, hence most of them can be considered random; therefore, it is not possible to predict excess returns. Table 3 and Table 4 show that the entropy of the series slightly differs from the entropy of a Bernoulli with probability of success 1/2 (0.693). Nonetheless, the percentage difference with respect to the entropy of a Bernoulli series is small for most of the series.

The majority of significant results are obtained for the runs test on Chievo home matches when the bets are ‘draw’ or ‘home team wins’. The reason could be twofold. First, low-ranking clubs generally receive better odds for their victory or a draw since their opponent are clubs of the same or higher tier; therefore, the return from betting successfully on H or D is generally above average. Second, one well-known phenomenon in the literature is the ‘home-field advantage’ (Jamieson, 2010), which originates from the fact that a team tends to have better performances in its home ground (Ferraresi and Gucciardi, 2021). Hence, betting on ‘draw’ or ‘home team wins’ when Chievo plays home may generate a predictable pattern of returns.

To assess efficiency in the short period we use a rolling window of 19 observations, which is the number of matches (home or away) in a season for each team. The first p-value is computed over the first 19 observations, then the second one is obtained dropping the first observation and adding the 20th observation, and so on. For the sake of synthesis, we report the results only for the bookmaker B365 and away matches. Figure 1 shows that p-values are consistently above the levels of significance. This evidences that the series are efficient even in the short-term. Figure 2 shows that in the short-term some of the filtered entropies differ from the value 0.693; however, large differences are only temporary. The same comments can be made for Figure 3 and Figure 4.

4. Conclusions

Research on betting markets efficiency is a longstanding theme in economics (for a review of the literature, see Sauer, 1998, Williams, 1999, Williams, 2011). In practice, informational efficiency in betting markets means that no bookmaker can operate at a greater margin than the others. In the same fashion, no bettor can systematically achieve above-average returns. A stream of literature, coherently with our findings, shows that these markets are efficient (Pankoff, 1968, Pope and Peel, 1989, Wolfers and Zitzewitz, 2004). On the other hand, many works have also documented inefficiencies in betting markets (Vlastakis et al., 2009, Goddard and Asimakopoulos, 2004, Dixon and Coles, 1997) and mixed results (Angelini and De Angelis, 2019). However, it is important to note that our work introduces a new methodology and comparisons with previous methods should be made carefully.

The present research can be extended in several directions. From a behavioural perspective, one of the most evident biases observed in betting markets is the ‘favourite–longshot bias’, discovered in racetrack betting markets (Quandt, 1986, Williams and Paton, 1997). This bias consists in the fact that bets placed on favourites usually yield higher returns than bets placed on longshots. This phenomenon has been identified in a variety of betting markets, including the European football betting market (Cain et al., 2000, Cain et al., 2003). Our methodology can be used to check for this bias and many others.

From an econometric perspective, many researchers focus on the use and development of econometric models. Just to name a few, Pope and Peel (1989) develop a linear probability model, whereas Cain et al. (2000) develop a model to estimate the mean number of goals scored by a team in a match by means of Poisson and negative binomial regressions. However, the empirical literature focuses mainly on regression methods, whereas our paper is based on symbolic series and information theory.

The methodology introduced here makes available a large set of tools to assess efficiency via binary series. Our method yields great advantages in terms of computational time and simplicity. Using this methodology, we found that the Italian and English soccer betting markets seem to be efficient. Future research could expand these results in several directions: using different statistical tools, using data from different soccer leagues or different sports, or introducing systems of bets or multiple bets, with a focus on low-ranking teams, for which we found some significant result.

Table 1: P-value of the runs test for Serie A clubs. * stands for 10% significance, ** stands for 5% significance, *** stands for 1% significance. The name of the bookmaker is followed by the symbols H, D or A to indicate, respectively, a bet on 'home team wins', 'draw', or 'away team wins'.

	Home matches			Away matches		
Bookmaker	Chievo	Lazio	Napoli	Chievo	Lazio	Napoli
B365H	0.1078	0.2038	0.9389	0.3281	0.2173	0.9083
B365D	0.0194**	0.8212	0.8903	0.9724	0.1777	0.8903
B365A	0.3971	0.1113	0.1009	0.3040	0.4124	0.4660
BWH	0.1297	0.2038	0.9149	0.3281	0.2337	0.9083
BWD	0.0327**	0.8212	0.7568	0.9724	0.1856	0.8903
BWA	0.3511	0.1113	0.0998*	0.3040	0.4782	0.4660
GBH	0.0934*	0.6031	0.8651	0.2454	0.5819	0.9067
GBD	0.4569	0.5913	0.3582	0.5435	0.5913	0.6419
GBA	0.3390	0.4172	0.2294	0.2962	0.4094	0.1693
IWH	0.1297	0.2854	0.9389	0.4388	0.2635	0.9083
IWD	0.0076***	0.8068	0.8903	0.8399	0.1856	0.8903
IWA	0.3063	0.1492	0.1009	0.2932	0.4391	0.4660
LBH	0.0847*	0.1960	0.9036	0.5377	0.0622*	0.6601
LBD	0.0247**	0.7134	0.6219	0.8555	0.1956	0.8958
LBA	0.2601	0.0783*	0.1192	0.2615	0.2347	0.5833
SBH	0.1448	0.6031	0.9301	0.0778*	0.3979	0.9918
SBD	0.3224	0.5913	0.5911	0.3159	0.6187	0.7440
SBA	0.7734	0.4172	0.1992	0.1772	0.3863	0.6717
WHH	0.1464	0.2038	0.6875	0.3281	0.2173	0.9083
WHD	0.0328**	0.8212	0.8854	0.9724	0.1777	0.8903
WHA	0.3971	0.1113	0.1175	0.3040	0.4124	0.4660
SJH	0.0718*	0.6297	0.8161	0.5262	0.6660	0.5388
SJD	0.3728	0.7729	0.6176	0.9170	0.9560	0.5873
SJA	0.3669	0.7202	0.3415	0.7778	0.3326	0.0926*
VCH	0.1696	0.2038	0.9447	0.2976	0.2173	0.9083
VCD	0.0248**	0.8212	0.7741	0.9880	0.1777	0.8903
VCA	0.6644	0.1113	0.0987*	0.2932	0.4124	0.4660
BSH	0.1220	0.6031	0.8651	0.4070	0.5819	0.9067
BSD	0.3650	0.5913	0.3582	0.7143	0.5913	0.6419
BSA	0.6828	0.4172	0.2294	0.2767	0.4094	0.1693
PSCH	0.4168	0.1793	0.8545	0.9797	0.2705	0.8934
PSCD	0.0152**	0.9203	0.7949	0.3021	0.0746*	0.7288
PSCA	0.5297	0.0905*	0.2434	0.7383	0.6162	0.3667
PSH	0.4959	0.2221	0.8545	0.9797	0.2705	0.9147
PSD	0.0527*	0.8772	0.7949	0.3021	0.0746*	0.7477
PSA	0.6781	0.1591	0.2434	0.7383	0.6162	0.3285

Table 2: P-value of the runs test for Premier League clubs. * stands for 10% significance, ** stands for 5% significance, *** stands for 1% significance. The name of the bookmaker is followed by the symbols H, D or A to indicate, respectively, a bet on 'home team wins', 'draw', or 'away team wins'.

	Home matches		Away matches	
Bookmaker	Chelsea	Everton	Chelsea	Everton
B365H	0.3037	0.1599	0.5541	0.8354
B365D	0.5232	0.5332	0.6081	0.4933
B365A	0.2679	0.3616	0.9951	0.0479**
BWH	0.3037	0.1599	0.5541	0.8354
BWD	0.5232	0.5332	0.6081	0.4933
BWA	0.2679	0.3616	0.9951	0.0479**
GBH	0.4878	0.5469	0.2632	0.6457
GBD	0.9263	0.4410	0.8750	0.4695
GBA	0.9409	0.5947	0.8771	0.1461
IWH	0.3037	0.1599	0.5541	0.8354
IWD	0.5232	0.5332	0.6081	0.4933
IWA	0.2679	0.3616	0.9951	0.0479**
LBH	0.1437	0.2710	0.5421	0.2837
LBD	0.2036	0.5949	0.3120	0.7946
LBA	0.2572	0.5277	0.9421	0.1196
SBH	0.3053	0.3591	0.6128	0.7609
SBD	0.9179	0.7440	0.7352	0.2175
SBA	0.9532	0.7434	0.6644	0.0767*
WHH	0.3037	0.1599	0.5541	0.8354
WHD	0.5232	0.5332	0.6081	0.4933
WHA	0.2679	0.3616	0.9951	0.0479**
SJH	0.4047	0.4506	0.3394	0.8482
SJD	0.8011	0.6131	0.6058	0.5752
SJA	0.9723	0.7874	0.9595	0.1245
VCH	0.3037	0.1599	0.5541	0.8354
VCD	0.5232	0.5332	0.6081	0.4933
VCA	0.2679	0.3616	0.9951	0.0479**
BSH	0.4878	0.5469	0.2632	0.6457
BSD	0.9263	0.4410	0.8750	0.4695
BSA	0.9409	0.5947	0.8771	0.1461
PSCH	0.2341	0.2786	0.5791	0.6979
PSCD	0.3861	0.5784	0.2845	0.9149
PSCA	0.1366	0.3903	0.8183	0.3149
PSH	0.6239	0.2786	0.5791	0.6979
PSD	0.4191	0.5784	0.2845	0.9149
PSA	0.1509	0.3903	0.8183	0.3149

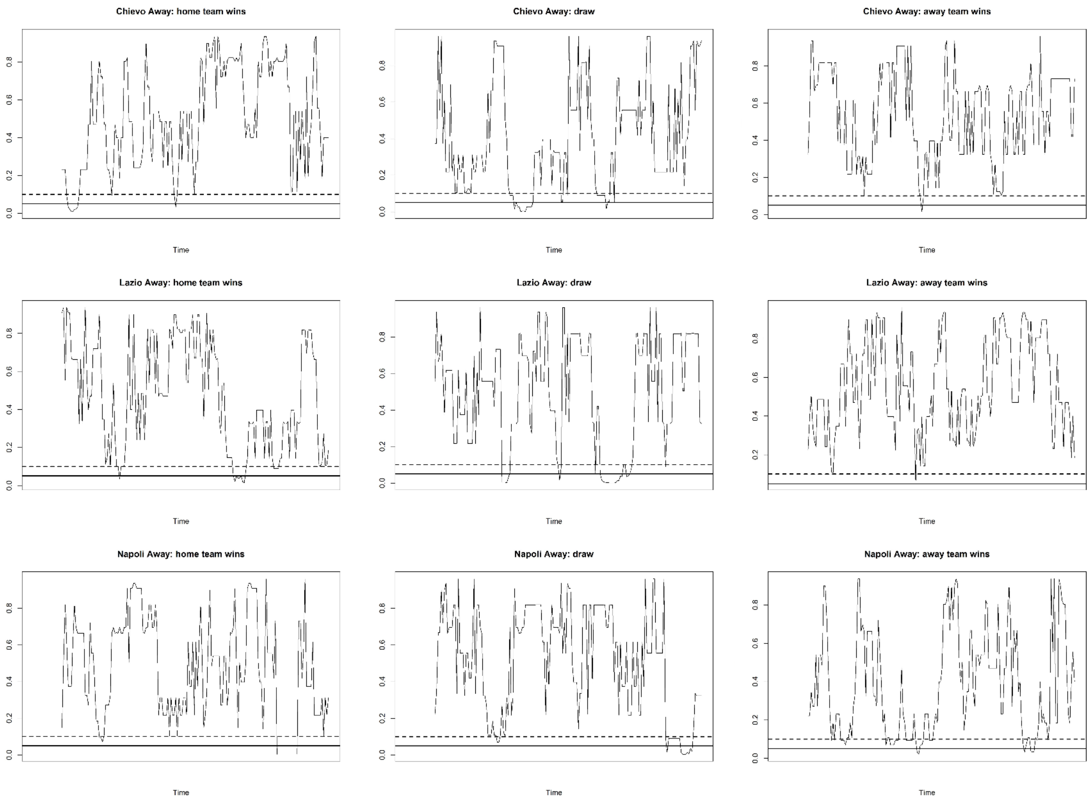

Figure 1: P-values of the runs test for betting returns using a rolling window (width = 19) for the Serie A odds of B365. P-value=0.05 (bold line), p-value=0.10 (dashed line).

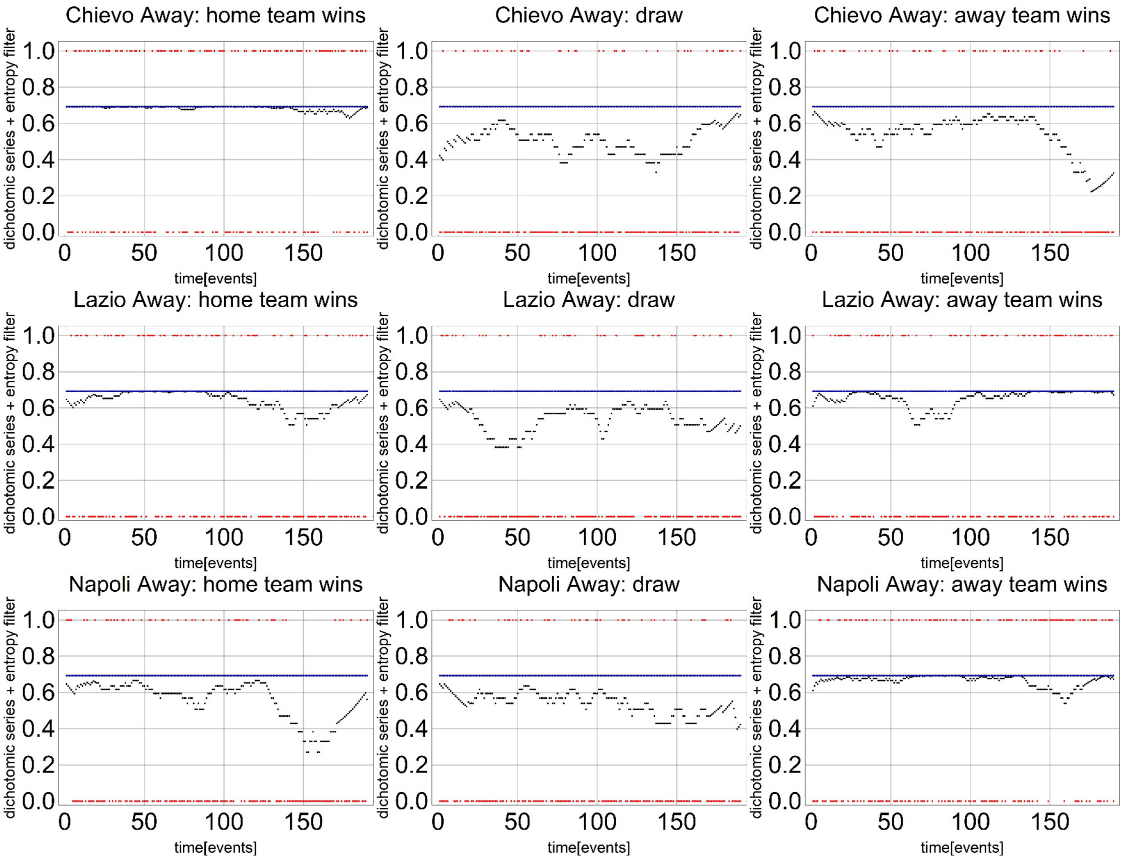

Figure 2: Binary series (in red), entropy for a Bernoulli series (in blue) and filtered entropy for the actual series using a neighbourhood of 19 observations. Serie A data for the bookmaker B365.

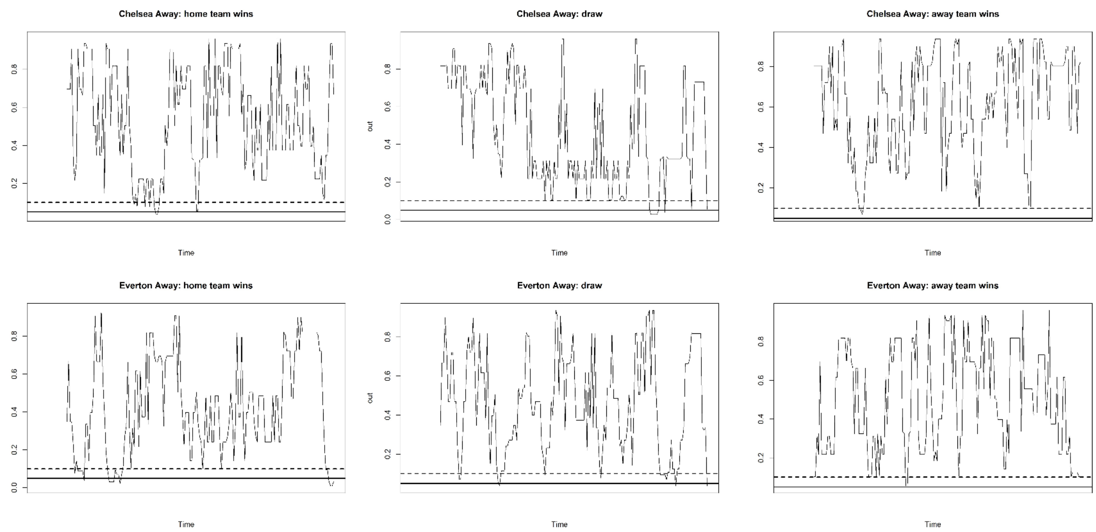

Figure 3: P-values of the runs test for betting returns using a rolling window (width = 19) for the Premier League odds of B365. P-value=0.05 (bold line), p-value=0.10 (dashed line).

Table 3: Entropy in Nats for the binary series of Serie A (percentage difference with respect to the entropy of a Bernoulli series in parentheses).

	Home matches			Away matches		
Bookmaker	Chievo	Lazio	Napoli	Chievo	Lazio	Napoli
B365H	0.6108 (11.88)	0.6895 (0.53)	0.6424 (7.32)	0.6895 (0.53)	0.6608 (4.67)	0.5777 (16.66)
B365D	0.6401 (7.65)	0.5076 (26.77)	0.5474 (21.03)	0.5146 (25.76)	0.5474 (21.03)	0.5474 (21.03)
B365A	0.6562 (5.33)	0.5652 (18.46)	0.3364 (51.47)	0.5594 (19.30)	0.6685 (3.56)	0.6930 (0.02)
BWH	0.6148 (11.30)	0.6895 (0.53)	0.6224 (10.21)	0.6895 (0.53)	0.6619 (4.51)	0.5816 (16.09)
BWD	0.6314 (8.91)	0.5076 (26.77)	0.5474 (21.03)	0.5146 (25.76)	0.5488 (20.82)	0.5474 (21.03)
BWA	0.6584 (5.01)	0.5652 (18.46)	0.3364 (51.47)	0.5594 (19.30)	0.6670 (3.77)	0.6930 (0.02)
GBH	0.6511 (6.07)	0.6900 (0.45)	0.6760 (2.47)	0.6928 (0.05)	0.6844 (1.26)	0.6334 (8.62)
GBD	0.6282 (9.37)	0.5315 (23.32)	0.6016 (13.21)	0.5146 (25.76)	0.5315 (23.32)	0.5894 (14.97)
GBA	0.6230 (10.12)	0.5474 (21.03)	0.3637 (47.53)	0.5894 (14.97)	0.6424 (7.32)	0.6708 (3.22)
IWH	0.6148 (11.30)	0.6895 (0.53)	0.6424 (7.32)	0.6890 (0.60)	0.6573 (5.17)	0.5816 (16.09)
IWD	0.6352 (8.36)	0.5076 (26.77)	0.5474 (21.07)	0.5090 (26.57)	0.5488 (20.82)	0.5474 (21.03)
IWA	0.6554 (5.45)	0.5652 (18.46)	0.3369 (51.40)	0.5608 (19.09)	0.6694 (3.43)	0.6930 (0.02)
LBH	0.6337 (8.58)	0.6881 (0.73)	0.6443 (7.05)	0.6902 (0.43)	0.6592 (4.90)	0.5822 (16.01)
LBD	0.6391 (7.80)	0.4987 (28.05)	0.5508 (20.54)	0.4905 (29.24)	0.5590 (19.35)	0.5574 (19.58)
LBA	0.6418 (7.41)	0.5639 (18.65)	0.3364 (51.47)	0.5822 (16.01)	0.6652 (4.03)	0.6927 (0.06)
SBH	0.6365 (8.17)	0.6900 (0.45)	0.6892 (0.57)	0.6917 (0.21)	0.6858 (1.06)	0.6479 (6.53)
SBD	0.6460 (6.80)	0.5315 (23.32)	0.6236 (10.03)	0.5763 (16.86)	0.5245 (24.33)	0.5936 (14.36)
SBA	0.6259 (9.70)	0.5474 (21.03)	0.4053 (41.53)	0.5763 (16.86)	0.6645 (4.13)	0.6581 (5.06)
WHH	0.6122 (11.68)	0.6895 (0.53)	0.6424 (7.32)	0.6895 (0.53)	0.6608 (4.67)	0.5816 (16.09)
WHD	0.6365 (8.17)	0.5076 (26.77)	0.5474 (21.03)	0.5146 (25.76)	0.5474 (21.03)	0.5474 (21.03)
WHA	0.6562 (5.33)	0.5652 (18.46)	0.3364 (51.47)	0.5594 (19.30)	0.6685 (3.56)	0.6930 (0.02)
SJH	0.6413 (7.48)	0.6896 (0.51)	0.6712 (3.17)	0.6922 (0.14)	0.6886 (0.66)	0.6126 (11.62)
SJD	0.5996 (13.50)	0.5505 (20.58)	0.5840 (15.75)	0.5117 (26.18)	0.5411 (21.94)	0.5840 (15.75)
SJA	0.6601 (4.77)	0.5455 (21.30)	0.3767 (45.65)	0.5840 (15.75)	0.6236 (10.33)	0.6824 (1.55)
VCH	0.6052 (12.69)	0.6895 (0.53)	0.6424 (7.32)	0.6899 (0.47)	0.6608 (4.67)	0.5816 (16.09)
VCD	0.6352 (8.36)	0.5076 (26.77)	0.5411 (21.94)	0.5164 (25.50)	0.5474 (21.03)	0.5474 (21.03)
VCA	0.6615 (4.57)	0.5652 (18.46)	0.3364 (51.47)	0.5608 (19.09)	0.6685 (3.56)	0.6930 (0.02)
BSH	0.6333 (8.63)	0.6900 (0.45)	0.6760 (2.47)	0.6923 (0.12)	0.6844 (1.26)	0.6334 (8.62)
BSD	0.6511 (6.07)	0.5315 (23.32)	0.6016 (13.21)	0.5004 (27.81)	0.5315 (23.32)	0.5894 (14.97)
BSA	0.6230 (10.12)	0.5474 (21.03)	0.3637 (47.53)	0.5929 (14.46)	0.6424 (7.32)	0.6708 (3.22)
PSCH	0.5932 (14.42)	0.6849 (1.19)	0.6050 (12.72)	0.6829 (1.48)	0.6495 (6.30)	0.5429 (21.68)
PSCD	0.6365 (8.17)	0.4720 (31.90)	0.5045 (27.22)	0.4832 (30.29)	0.5602 (19.18)	0.5244 (24.35)
PSCA	0.6670 (3.77)	0.5684 (18.00)	0.3030 (56.29)	0.5517 (20.41)	0.6723 (3.01)	0.6883 (0.70)
PSH	0.5851 (15.59)	0.6892 (0.57)	0.6050 (12.72)	0.6829 (1.48)	0.6579 (5.09)	0.5450 (21.37)
PSD	0.6302 (9.08)	0.4905 (29.24)	0.5045 (27.22)	0.4835 (30.25)	0.5602 (19.18)	0.5265 (24.04)
PSA	0.6744 (2.70)	0.5763 (16.86)	0.3030 (56.29)	0.5517 (20.41)	0.6723 (3.01)	0.6890 (0.68)

Table 4: Entropy in Nats for the binary series of Premier League (percentage difference with respect to the entropy of a Bernoulli series in parentheses).

	Home matches		Away matches	
Bookmaker	Chelsea	Everton	Chelsea	Everton
B365H	0.6237 (10.03)	0.6912 (0.29)	0.5969 (13.89)	0.6685 (3.55)
B365D	0.5004 (27.81)	0.5708 (17.64)	0.5348 (22.85)	0.6522 (5.91)
B365A	0.3585 (48.29)	0.5147 (25.75)	0.6929 (0.03)	0.5652 (18.46)
BWH	0.6237 (10.03)	0.6912 (0.29)	0.5969 (13.89)	0.6685 (3.55)
BWD	0.5004 (27.81)	0.5708 (17.64)	0.5348 (22.85)	0.6522 (5.91)
BWA	0.3585 (48.29)	0.5147 (25.75)	0.6929 (0.03)	0.5652 (18.46)
GBH	0.5894 (14.96)	0.6876 (0.80)	0.6131 (11.55)	0.6424 (7.32)
GBD	0.4362 (37.08)	0.6017 (13.20)	0.5763 (16.85)	0.6806 (1.81)
GBA	0.3638 (47.52)	0.4362 (37.08)	0.6845 (1.25)	0.5474 (21.03)
IWH	0.6237 (10.03)	0.6912 (0.29)	0,5969 (13.89)	0.6685 (3.55)
IWD	0.5004 (27.81)	0.5708 (17.64)	0.5348 (22.85)	0.6522 (5.91)
IWA	0.3585 (48.29)	0.5147 (25.75)	0.6929 (0.03)	0.5652 (18.46)
LBH	0.6191 (10.69)	0.6911 (0.30)	0.5880 (15.17)	0.6641 (4.19)
LBD	0.4820 (30.46)	0.5763 (16.85)	0.5440 (21.52)	0.6581 (5.05)
LBA	0.3725 (46.26)	0.5068 (26.88)	0.6930 (0.02)	0.5639 (18.64)
SBH	0.5763 (16.85)	0.6918 (0.20)	0.6237 (10.03)	0.6480 (6.52)
SBD	0.4362 (37.08)	0.5936 (14.36)	0.5936 (14.36)	0.6744 (2.70)
SBA	0.3365 (51.45)	0.4905 (29.23)	0.6744 (2.70)	0.5575 (19.57)
WHH	0.6237 (10.03)	0.6912 (0.29)	0.5969 (13.89)	0.6685 (3.55)
WHD	0.5004 (27.81)	0.5708 (17.64)	0.5348 (22.85)	0.6522 (5.91)
WHA	0.3585 (48.29)	0.5147 (25.75)	0.6929 (0.03)	0.5652 (18.46)
SJH	0.5707 (17.67)	0.6841 (1.30)	0.6009 (13.30)	0.6266 (9.61)
SJD	0.4307 (37.87)	0.5813 (16.14)	0.5595 (19.28)	0.6738 (2.79)
SJA	0.3318 (52.13)	0.4478 (35.40)	0.6905 (0.38)	0.5914 (14.68)
VCH	0.6237 (10.03)	0.6912 (0.29)	0.5969 (13.89)	0.6685 (3.55)
VCD	0.5004 (27.81)	0.5708 (17.64)	0.5348 (22.85)	0.6522 (5.91)
VCA	0.3585 (48.29)	0.5147 (25.75)	0.6929 (0.03)	0.5652 (18.46)
BSH	0.5894 (14.96)	0.6876 (0.80)	0.6131 (11.55)	0.6424 (7.32)
BSD	0.4362 (37.08)	0.6017 (13.20)	0.5763 (16.85)	0.6806 (1.81)
BSA	0.3638 (47.52)	0.4362 (37.08)	0.6845 (1.25)	0.5474 (21.03)
PSCH	0.6448 (6.97)	0.6909 (0.33)	0.5839 (15.76)	0.6754 (2.56)
PSCD	0.5430 (21.67)	0.5603 (19.17)	0.5045 (27.21)	0.6399 (7.68)
PSCA	0.3523 (49.18)	0.5244 (24.34)	0.6918 (0.20)	0.5684 (17.99)
PSH	0.6581 (5.05)	0.6909 (0.33)	0.5839 (15.76)	0.6754 (2.56)
PSD	0.5535 (20.15)	0.5603 (19.17)	0.5045 (27.21)	0.6399 (7.68)
PSA	0.3793 (45.28)	0.5244 (24.34)	0.6918 (0.20)	0.5684 (17.99)

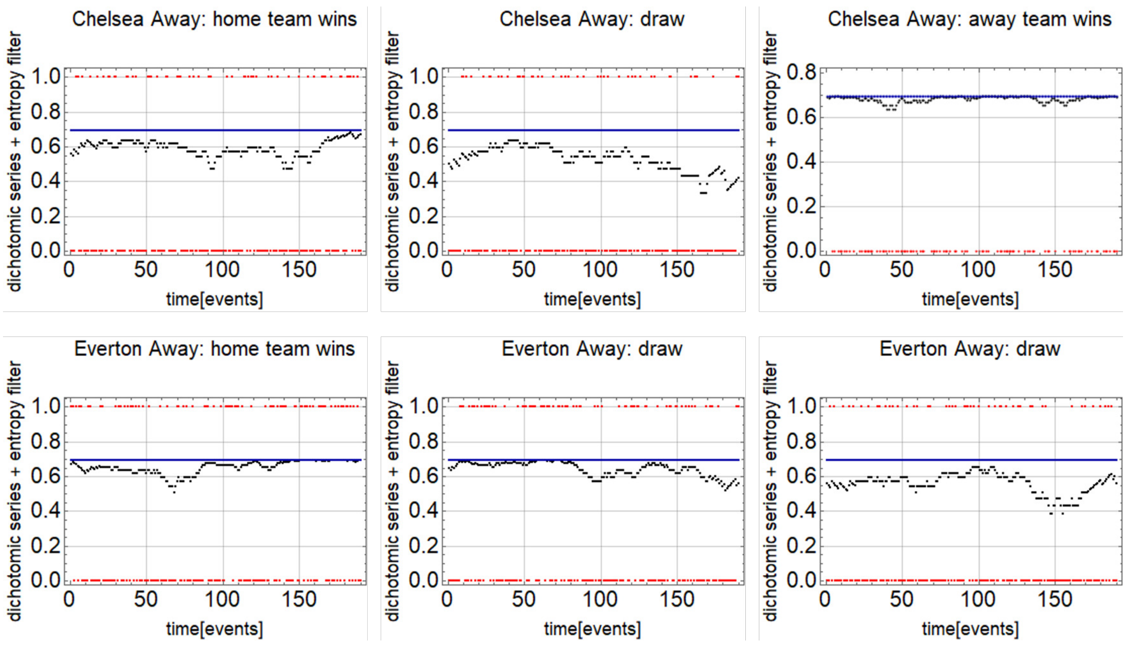

Figure 4: Binary series (in red), entropy for a Bernoulli series (in blue) and filtered entropy for the actual series using a neighbourhood of 19 observations. Premier League data for the bookmaker B365.

References

Angelini, G. and De Angelis, L. (2019) “Efficiency of online football betting markets” *International Journal of Forecasting* **35(2)**, 712–721.

Cain, M., Law, D., and Peel, D. (2000) “The favourite-longshot bias and market efficiency in UK football betting” *Scottish Journal of Political Economy* **47(1)**, 25–36.

Cain, M., Law, D., and Peel, D. (2003) “The favourite-longshot bias, bookmaker margins and insider trading in a variety of betting markets” *Bulletin of Economic Research* **55(3)**, 263–273.

Dixon, M. J. and Coles, S. G. (1997) “Modelling association football scores and inefficiencies in the football betting market” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **46(2)**, 265–280.

Ferraresi, M. and Gucciardi, G. (2021) “Who chokes on a penalty kick? Social environment and individual performance during Covid-19 times” *Economics Letters* **203**, 109868.

Goddard, J. and Asimakopoulos, I. (2004) “Forecasting football results and the efficiency of fixed-odds betting” *Journal of Forecasting* **23(1)**, 51–66.

Jamieson, J. P. (2010) “The home field advantage in athletics: A meta-analysis” *Journal of Applied Social Psychology* **40(7)**, 1819–1848.

Oikonomidis, A., Bruce, A. C., and Johnson, J. E. (2015) “Does transparency imply efficiency? The case of the European soccer betting market” *Economics Letters* **128**, 59–61.

Pankoff, L. D. (1968) “Market efficiency and football betting” *The Journal of Business* **41(2)**, 203–214.

Pernagallo, G. and Torrisi, B. (2019) “An empirical analysis on the degree of Gaussianity and long memory of financial returns in emerging economies” *Physica A: Statistical Mechanics and its Applications* **527**, 121296.

Pernagallo, G. and Torrisi, B. (2020) “Blindfolded monkeys or financial analysts: Who is worth your money? New evidence on informational inefficiencies in the U.S. stock market” *Physica A: Statistical Mechanics and its Applications* **539**, 122900.

Pope, P. F. and Peel, D. A. (1989) “Information, prices and efficiency in a fixed-odds betting market” *Economica* **56(223)**, 323–341.

Quandt, R. E. (1986) “Betting and equilibrium” *The Quarterly Journal of Economics*, **101(1)**, 201–208.

Sauer, R. (1998) “The economics of wagering markets” *Journal of economic Literature* **36(4)**, 2021-2064.

Vlastakis, N., Dotsis, G., and Markellos, R. N. (2009) “How efficient is the European football betting market? Evidence from arbitrage and trading strategies” *Journal of Forecasting* **28(5)**, 426–444.

Williams, L. V. (1999) “Information efficiency in betting markets: A survey” *Bulletin of Economic Research* **51(1)**, 1-39.

Williams, L. V. (2011) “Prediction markets: Theory and applications”, Routledge International Studies in Money and Banking.

Williams, L. V. and Paton, D. (1997) “Why is there a favourite-longshot bias in British racetrack betting markets” *The Economic Journal* **107(440)**, 150–158.

Wolfers, J. and Zitzewitz, E. (2004) “Prediction markets” *Journal of economic perspectives* **18(2)**, 107–126.