International Journal of Forecasting 35 (2019) 712–721

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Efficiency of online football betting markets

Giovanni Angelini [a], Luca De Angelis [b,*]

[a] Departments of Economics, University Ca' Foscari of Venice, Italy
[b] Department of Economics, University of Bologna, Italy

ARTICLE INFO

Keywords:
Market efficiency
Sports forecasting
Probability forecasting
Favourite–longshot bias
Betting markets

ABSTRACT

This paper evaluates the efficiency of online betting markets for European (association) football leagues. The existing literature shows mixed empirical evidence regarding the degree to which betting markets are efficient. We propose a forecast-based approach for formally testing the efficiency of online betting markets. By considering the odds proposed by 41 bookmakers on 11 European major leagues over the last 11 years, we find evidence of differing degrees of efficiency among markets. We show that, if the best odds are selected across bookmakers, eight markets are efficient while three show inefficiencies that imply profit opportunities for bettors. In particular, our approach allows the estimation of the odds thresholds that could be used to set profitable betting strategies both ex post and ex ante.

1. Introduction

The issue of the degree of efficiency is crucial to the analysis of markets, as market inefficiencies, if correctly predicted and measured, may create significant opportunities for profit. For financial markets, Fama (1970) introduced the renowned Efficient-Market Hypothesis, which, in its weak form, postulates that markets are efficient in the sense that current prices reflect all the information that is contained in historical prices, thus ruling out the possibility of achieving excess returns using technical analysis techniques. In general, informational efficiency requires that prices represent the best forecasts of the outcomes of future events. Therefore, investors cannot achieve a risk-adjusted return that is in excess of the market by trading on new information.

Market efficiency naturally applies to many different kinds of markets, including betting markets. Following the growth of the online betting industry over the last decade, a number of scholars have focused on betting markets in particular because they represent a sort of 'real world laboratory' where efficiency can be investigated in a straightforward way (see e.g. the seminal paper by Thaler & Ziemba, 1988, and the comprehensive review for both financial and betting markets by Vaughan Williams, 2005). In fact, unlike in financial markets, betting market participants are in general well-informed, motivated and experienced, and breaking news in sports is usually reported cleanly and in a way that is easy for the agents to share and process. In other words, there is very little room for the information leakages that affect the efficiency of financial markets. Forrest and Simmons (2000) show that private or semi-public information which might be possessed by professional English newspaper tipsters enhances the forecasts of match outcomes only slightly, and that there is no overwhelming evidence that the predictions of match results from regression models are inferior to those made by professional experts who claim to possess private (insider) information. Moreover, for the German Bundesliga, Spann and Skiera (2009) show a better accuracy of bookmaker odds than tipster predictions. Recently, though, Brown and Reade (2017) found that tipster picks do provide additional predictive information over and above bookmaker prices, and Brown, Rambaccussing, Reade, and Rossi (2017)

* Corresponding author.
E-mail address: l.deangelis@unibo.it (L. De Angelis).

https://doi.org/10.1016/j.ijforecast.2018.07.008
0169-2070/

showed that Twitter tweets contain information that is not incorporated in live (in-play) betting prices, especially immediately after goals and red cards. Furthermore, bets are characterized by a precise deadline after which their value becomes certain, thus making testing for market efficiency much simpler.

Betting market efficiency implies that market prices (i.e., bookmaker odds) reflect all relevant historical information and represent the best forecasts of the match outcome's probabilities. Therefore, after considering bookmaker commissions, bettors cannot pursue profit opportunities, as all of the information available is already reflected in the odds quoted. Nevertheless, for (association) football matches, Angelini and De Angelis (2017), Boshnakov, Kharrat, and McHale (2017), Dixon and Pope (2004), Goddard and Asimakopoulos (2004) and Koopman and Lit (2015), among others, show abnormal positive returns out-of-sample from betting strategies based on econometric approaches; specifically, Poisson, ordered probit, dynamic state space, bivariate Weibull count and Poisson autoregressive models, respectively. As these methods use information on past match results, their results imply betting market inefficiencies. Moreover, forecasting models that produce abnormal positive returns have been proposed for other sports, including American football (Boulier & Stekler, 2003; Glickman & Stern, 1998), tennis (McHale & Morton, 2011), horse racing (Lessmann, Sungb, & Johnson, 2010) and Australian Rules football (Grant & Johnstone, 2010; Rydall & Bedford, 2010).

The topic of betting market efficiency has been developed rather extensively in the literature, but there is mixed empirical evidence as to the degree to which betting markets are efficient. In particular, the efficiency of win-draw-lose match outcomes in football betting markets is still an open issue. To the best of our knowledge, the recent literature is lacking in contributions which formally test the efficiency of betting markets for online football wagers. The only studies that we know of that rigorously test for market efficiency in win-draw-lose match outcomes in football are those by Kuypers (2000) and Pope and Peel (1989). In particular, Pope and Peel (1989) proposed an approach based on linear probability and logit models for testing for betting market efficiency in the 1981/1982 football season in the UK. Their findings suggest that there is no bias in the bookmaker's odds-setting processes for home and away wins and therefore that no profitable betting strategy can be identified. Using OLS regression between the outcome probability and the probability implied by the odds, Kuypers (2000) concluded that there was no systematic bias in the odds and that the market was weakly efficient for the 1993/1994 and 1994/1995 seasons of the four divisions of the English football league. However, Kuypers (2000) found a rare occurrence of both inefficiency and profitable betting opportunities if an ordered logit model with publicly available information variables was employed. With respect to these studies, we propose an innovative approach where we model the bookmaker's forecast errors in order to formally test for market efficiency, after considering bookmaker commissions. Moreover, our analysis considers a larger sample size in terms of both the time span and the number of football leagues, as well as in terms of betting market coverage, as Kuypers (2000) and Pope and Peel (1989) considered only one country and one and four bookmakers, respectively.

This paper investigates the degree of efficiency in European football online betting markets by testing the predictability of football match outcomes based on the information contained in the odds offered in the market. In particular, we test for efficiency in the football online betting markets related to single European major leagues, in order to investigate possible differences in the degree of market (in)efficiency among national club competitions in Europe. We achieve this goal by considering a dataset that comprises the odds proposed by 41 international bookmakers on 11 leagues over the last 11 years (2006–2017), for a total of 33,060 football matches.

One well-known deviation from unbiasedness (and sometimes from efficiency) is the favourite–longshot bias, which claims that the odds on favourites are more profitable than those on longshots; i.e., bookmakers tend to underestimate (overestimate) expected winners (underdogs) (see Sorensen & Ottaviani, 2008 for a review). This bias is well-documented in racetrack betting markets (Griffith, 1949; Hurley & McDonough, 1995; Thaler & Ziemba, 1988), but not in other sports betting markets; for example, Woodland and Woodland (1994, 2001) show that the betting markets for Major League Baseball and National Hockey League matches reveal the opposite bias. Exploiting only information contained in the odds and not relying on any econometric model, Direr (2013) showed that systematically picking out the betting odds of overwhelming favourites (whose probability of winning exceeds 90%) leads to abnormal positive returns. His evidence appears to go against the market efficiency hypothesis and to be consistent with the literature that documents the presence of a favourite–longshot bias in betting markets. Conversely, using high frequency data provided by an online betting exchange, Croxson and Reade (2014) tested for (semi-strong form) efficiency considering the (half-time) price reaction to cusp goals that arrive within five minutes of the end of the first half of the match. Their results show that prices update promptly and completely so that the news of a goal is incorporated fully by the time the break starts.

A further strand of the literature focuses on the accuracy of betting odds-based probability forecasts both *per sé* (Štrumbelj & Šikonja, 2010) and in comparison with model-based probability forecasts (Forrest, Goddard, & Simmons, 2005; Štrumbelj & Vračar, 2012). The empirical evidence suggests that betting odds are the most accurate source of sports forecasts.[1] In line with this literature, we consider the implied probabilities provided by the online market odds as the 'best' available forecasts of the match outcomes and analyze the forecast errors to test for market-specific efficiency, under the null hypothesis of market efficiency.

Our main findings show that all of the markets are efficient and may allow extra profits for bookmakers only,

[1] Actually, evidence of superior forecasting performances for the English leagues is provided by prices in betting exchanges and prediction markets (Franck, Verbeek, & Nüesch, 2010; Reade, 2014). Nevertheless, the focus of our analysis is on the efficiency of online fixed-odds bookmakers.

if the mean market odds are used. Conversely, if the maximum odds offered by the market are considered, we find evidence of three European leagues where the favourite–longshot bias is sufficiently large to create profitable opportunities for bettors.

This paper is organized as follows. Section 2 outlines the testing approach for market efficiency. Section 3 analyzes the degree of market efficiency for 11 European leagues. Specifically, Section 3.1 describes the data, Section 3.2 presents the results for the market efficiency tests and Section 3.3 investigates the implications of market inefficiencies by presenting a simple but profitable betting strategy. Section 4 concludes.

2. Testing for efficiency of online football betting markets

Let y_i be a dichotomous variable which assumes a value of one if match i ends with the outcome under consideration, i.e., home team win, draw, or away team win. The variable y_i is then distributed as a Bernoulli with (true) probability π_i, i.e. $y_i|\Omega_i \sim Bin(1, \pi_i)$, where Ω_i denotes the hypothetical information set that contains all the information in the universe. A strand of the literature on sports forecasting agrees on the empirical evidence that betting odds are the most accurate source of data for predicting the probabilities of the match outcomes (see, e.g., Štrumbelj & Šikonja, 2010; Štrumbelj, 2014). In this regard, the odds quoted on the online (fixed-odds) betting market represent the 'best' available (*ex ante*) forecasts of the likelihood of the outcome of match i. Let o_i be the bookmaker odd for a particular outcome of match i (e.g. home win) and $p_i = 1/o_i$ be the corresponding implied probability forecast. Hence, the bookmaker's probability forecast should be $p_i = E(y_i|\mathcal{I}_i)$ where $\mathcal{I}_i \subset \Omega_i$ is the (actual) information set available to the bookmakers on match i. However, bookmakers do not offer fair odds because the odds must also incorporate the bookmaker commission or margin, also known as the 'vig'. Therefore, the bookmaker's probability forecast that is *de facto* employed to set the odds offered in the market is $p_i = E(y_i|\mathcal{I}_i) + \kappa_i$, where $\kappa_i > 0$ is the bookmaker commission. As a consequence, the p_is are not actual probabilities, as their sum over all possible outcomes exceeds one. Occasional cases where the sum of the inverse odds is smaller than one may occur when considering the best odds offered by the market, and these cases provide arbitrage opportunities for bettors. Arbitrage opportunities are very rare if only online bookmakers are considered. Vlastakis, Dotsis, and Markellos (2009) find that fewer than one every 1000 matches allows for arbitrage opportunities on online betting markets. However, they also show that this ratio increases to a non-negligible 0.5% if both online and fixed-odds bookmakers are considered. Arbitrage positions can be also achieved by combining bets at exchange and online bookmaker markets (Franck, Verbeek, & Nüesch, 2013).

Since the bookmaker commission κ_i is not generally fixed and may change among matches, between bookmakers and over time, one popular way of circumventing this issue is through odds normalization, i.e., dividing the inverse odds by the sum of the inverse odds. However, this approach makes the implicit assumption that bookmakers add their margins proportionately across all possible outcomes. Moreover, even though more sophisticated methods of determining probability forecasts from betting odds exist (see e.g. Štrumbelj, 2014), Levitt (2004) shows that bookmakers set their odds so as to exploit the bettors' biases, and thus, the implied probabilities will be different from those expected, even after normalization. Specifically, for NFL American football, Levitt (2004) provides evidence of the bookmakers' ability to set odds and concludes that they are better than gamblers at predicting match outcomes. Nevertheless, the odds offered by the bookmakers deviate systematically from those expected, as they are set to exploit bettors' biases and thus earn extra profits. A similar argument is made by Kuypers (2000).

Let $\varepsilon_i = y_i - p_i$ be the bookmaker's forecast error for the outcome of match i. Under the null hypothesis of market efficiency, we have that, in general, p_i overestimates π_i, i.e., $p_i > E(y_i|\Omega_i)$, and, as a consequence, the conditional expectation of ε_i is not null but is equal to (minus) the bookmaker commission and possible price distortions resulting from bettor's bias exploitation, i.e., $E(\varepsilon_i|\mathcal{I}_i) = -\kappa_i$.

Thus, market efficiency for league $j = 1, \ldots, J$ can be evaluated by estimating the following model:

$$\varepsilon_{i,j} = \alpha_{1,j} + \sum_{t=2}^{T} \alpha_{t,j} d_t + \beta_j p_{i,j} + \nu_{i,j},$$
$$\nu_{i,j} \sim i.i.d.(0, \sigma^2_{i,j}), \quad i = 1, \ldots, N_j, \tag{1}$$

where N_j is the number of matches considered for league j and d_t is a dummy variable which assumes a value of one for season t and zero otherwise, for $t = 2, \ldots, T$, so that $\alpha_{1,j}$ captures the average bookmaker commission for the jth league in season 1 (as a reference) and $\alpha_{t,j}$, for $t = 2, \ldots, T$, captures the possible development over time of the bookmaker margins.[2] Since the regression coefficient β_j in Eq. (1) captures the possible effect of the probability $p_{i,j}$ on the forecast error $\varepsilon_{i,j}$, the market efficiency of league j can be evaluated by investigating its statistical significance. More specifically, once we account for the bookmaker commissions, which are measured by the α coefficients in Eq. (1), market efficiency would imply that the conditional expectation $E(\varepsilon_i|\mathcal{I}_i)$ is zero, so that a rejection of the null hypothesis $H_0 : \beta_j = 0$ would imply that market j is not unbiased. Note that the setup used in Eq. (1) for testing the efficiency of betting markets is akin to the standard forecasting efficiency testing framework by Mincer and Zarnowitz (1969); i.e., under the efficient market hypothesis and in the case of no bookmaker commission ($\kappa_i = 0$), $E(\varepsilon_i|\mathcal{I}_i) = 0$ regardless of the regressor belonging to the information set $\mathcal{I}_i$ that we might include in the model specification; see also Clements and Reade (2016). In addition to this, the 'efficiency curves' described

2 Thus, Eq. (1) implies that the bookmaker commission may vary over time and across leagues. In Section 3.2 we test whether these commissions are time-invariant. Future research could consider generalisations of Eq. (1) that are able to account for fixed effects in other dimensions, such as individual bookmakers, for example, or perhaps even relax the assumption that α is independent of match i, since bookmakers may adjust their commission to the probability of the match outcome.

in Section 3.2 allow us to evaluate whether this bias is large enough to overcome the bookmaker commissions and allow a profitable betting strategy based on past odds. If such is the case, then market j is inefficient. Moreover, the inclusion of the intercept dummy variables for each season in the model specification in Eq. (1) allows us both to test whether the bookmaker margin is time-invariant and to capture its possible evolution over time.

Ioannidis and Peel (2005) show that forecast errors can exhibit heteroskedasticity under the null of market efficiency. We account for this possibility by obtaining the estimation of Eq. (1) through Weighted Least Squares (WLS), where the $N_j \times N_j$ weighting matrix is diagonal with elements $\sigma^2_{1,j}, \ldots, \sigma^2_{N_j,j}$. In our setup, $\sigma^2_{i,j}$ can be approximated by $p_{i,j}(1 - p_{i,j})$. Moreover, as we consider home and away win odds jointly in Eq. (1), we avoid the issue of possible correlations between observations by considering the cluster-robust estimate of the covariance matrix of the WLS estimator, where the clusters consist of the (two) observations related to the same match; see Liang and Zeger (1986) for more details.

In the next section we investigate the extent of the efficiency in European online betting markets.

3. Empirical results

3.1. Data

The data used in this paper are taken from www.football-data.co.uk, a big database of European football match results and fixed odds, where the odds are recorded on Friday afternoons for weekend matches and on Tuesday afternoons for midweek matches. These data comprise the odds offered by the 41 international online bookmakers considered by the BetBrain portal (www.betbrain.com) for the football matches played in 11 major European leagues over the period from August 2006 to February 2017, for a total of 33,060 matches. The leagues considered in the analysis are as follows: English Premier League, Scottish Premier League, German Bundesliga, Italian Serie A, Turkish Super Lig, Portuguese Primeira Liga, French Ligue 1, Spanish Liga, Greek Super League, Dutch Eredivisie, and Belgian Jupiler League. For each match, we consider both the mean and maximum odds offered by the market. The sample sizes (N_j), which are reported in the last row of Tables 2 and 3 for each league, are rather large, so that the theoretical convergence to the normal distribution (of sums) of Bernoulli variables should be attained.

Pope and Peel (1989) show that the variability of draw probabilities is very low and that draw odds have no significant predictive content for the draw outcome. An analysis of draws that was performed on our dataset confirmed the findings of Pope and Peel (1989), revealing no significant relationships between draw odds and draw outcomes for all leagues considered. Therefore, in what follows, we model home and away win odds jointly, and do not consider draws.

Our analysis also focuses on the deviation from unbiasedness that is provided by the favourite–longshot bias, which is an empirical regularity documented in many sports betting markets, as was discussed in Section 1.

3.2. Efficiency of online European football betting markets

This section tests for the efficiency of the online betting markets for the 11 major European football leagues listed in Section 3.1.

If betting markets are efficient, then the conditional expectation of the forecast errors should be equal to minus the bookmaker commissions. Therefore, from the estimation of Eq. (1) for the jth league, we would expect (i) to find that the estimate for $\alpha_{1,j}$ may be (significantly) negative, as this parameter captures the bookmaker margin, and (ii) not to reject the null hypothesis $H_0 : \beta_j = 0$. The results for the estimation of the models in Eq. (1) are reported in Tables 2 and 3 for the mean and the maximum odds, respectively.

Before focusing on the estimation results, we first test the assumption that the true probability is a linear function of the bookmaker probability, as implied by Eq. (1).[3] In particular, we consider a set of Ramsey's RESET tests for functional form misspecification in Table 1. The results from these tests show that, overall, the models in Eq. (1) are specified correctly, and thus, it is not necessary to adopt a non-linear model specification. However, it must be noted that some evidence of non-linearity is found for England (squares and cubes) and Italy (squares only) at the 5% (but not the 1%) significance level.

The results in Table 2 show that, considering the mean of the odds offered by the 41 online bookmakers analyzed, we do not reject the null hypothesis of market efficiency for any of the leagues, except for Italian Serie A and Portuguese Primeira Liga at 5% significance level, and Greek Super League even at a 1% level of significance. Quite surprisingly, we find evidence of a negative slope in German Bundesliga and Dutch Eredivisie, though they are not significant. All of the other regression slopes (including the non-significant ones) are positive, implying that, on average, the bookmaker's forecast error tends to increase as the forecast probability increases. This is consistent with the renowned favourite–longshot bias, which we investigate below.

The results for the estimates of $\alpha_{1,j}$, reported in Table 2, show that, as expected, these are all negative. Focusing on $\hat{\alpha}_{t,j}$, for $t = 2, \ldots, 10$, we observe that some of these dummy variables are significant at least at the 5% level, and positive. This provides (mild) evidence that the bookmaker commission has decreased over the sample, which might be due to the increased level of competition in online betting markets. However, we can infer from the F-tests for the (joint) null hypotheses $H_0 : \alpha_{2,j} = \cdots = \alpha_{10,j} = 0$ reported in Table 2 that actually commissions have not changed significantly over time. The evidence for time-invariant commissions in our sample is in line with the findings of Forrest et al. (2005, Table 4), who found that, though bookmaker prices were becoming more accurate with the increasing commercial pressure, their take-out was relatively constant over the period from 1998 to 2003.

We therefore improve the power of the test by simplifying the model in Eq. (1) by imposing the restriction of

[3] It must be noted that a violation of the linearity assumption would invalidate the results, as β would not carry the intended meaning and its being different from zero might be due to a functional form misspecification of the model rather than any actual market inefficiency.

Table 1
p-values for Ramsey's RESET tests for the null hypothesis that Eq. (1) is specified correctly (no functional form misspecification) for all leagues considered.

RESET tests											
Powers	England	Scotland	Germany	Italy	Turkey	Portugal	France	Spain	Greece	Holland	Belgium
2	0.276	0.371	0.257	0.041	0.563	0.592	0.316	0.453	0.050	0.409	0.350
2–3	0.024	0.559	0.315	0.058	0.529	0.841	0.554	0.060	0.125	0.613	0.427
2–3–4	0.054	0.311	0.507	0.125	0.722	0.948	0.720	0.131	0.185	0.605	0.394

Table 2
Estimates of the models in Eqs. (1) and (2) when we consider the mean of the odds offered on the betting market.

Mean odds											
	England	Scotland	Germany	Italy	Turkey	Portugal	France	Spain	Greece	Holland	Belgium
$\hat{\alpha}_1$	−0.0417*** (0.0007)	−0.0226* (0.0987)	−0.0066 (0.3475)	−0.0571*** (0.0000)	−0.0351** (0.0255)	−0.0799*** (0.0000)	−0.0453*** (0.0025)	−0.0134 (0.1549)	−0.0489*** (0.0005)	−0.0183 (0.1122)	−0.0488*** (0.0029)
$\hat{\alpha}_2$	0.0116 (0.4121)	−0.0085 (0.6410)	0.0062 (0.7127)	0.0173 (0.2532)	0.0081 (0.6185)	0.0220 (0.2567)	0.0024 (0.8801)	−0.0004 (0.9783)	−0.0212 (0.2139)	−0.0114 (0.4733)	0.0218 (0.1813)
$\hat{\alpha}_3$	0.0155 (0.2519)	−0.0173 (0.3442)	−0.0081 (0.6379)	0.0118 (0.4317)	−0.0195 (0.2617)	0.0144 (0.4447)	0.0235 (0.1304)	−0.0257* (0.0592)	0.0000 (0.9987)	0.0000 (0.9976)	0.0181 (0.3199)
$\hat{\alpha}_4$	0.0102 (0.4875)	0.0052 (0.7651)	0.0326* (0.0517)	0.0214 (0.1586)	0.0020 (0.9006)	0.0250 (0.1953)	−0.0180 (0.2621)	−0.0004 (0.9765)	0.0010 (0.9530)	−0.0075 (0.6276)	0.0025 (0.8855)
$\hat{\alpha}_5$	0.0277* (0.0564)	0.0019 (0.9169)	0.0031 (0.8556)	0.0058 (0.7098)	0.0038 (0.8124)	0.0409** (0.0217)	0.0071 (0.6535)	−0.0170 (0.2138)	0.0080 (0.6421)	0.0009 (0.9507)	−0.0026 (0.8809)
$\hat{\alpha}_6$	0.0067 (0.6365)	−0.0195 (0.3205)	0.0037 (0.8268)	0.0223 (0.1417)	−0.0035 (0.8364)	0.0276 (0.1374)	0.0071 (0.6549)	−0.0013 (0.9247)	0.0031 (0.8559)	−0.0130 (0.4259)	0.0210 (0.2344)
$\hat{\alpha}_7$	0.0395*** (0.0042)	0.0068 (0.6972)	0.0179 (0.2549)	0.0167 (0.2504)	−0.0006 (0.9718)	0.0338* (0.0723)	0.0074 (0.6332)	−0.0035 (0.8032)	0.0081 (0.5879)	−0.0269 (0.1011)	0.0218 (0.2094)
$\hat{\alpha}_8$	0.0295** (0.0394)	0.0188 (0.2861)	0.0008 (0.9616)	−0.0034 (0.8288)	−0.0086 (0.6094)	0.0246 (0.1535)	0.0318** (0.0360)	−0.0047 (0.7259)	−0.0062 (0.6956)	−0.0047 (0.7659)	0.0123 (0.4786)
$\hat{\alpha}_9$	0.0205 (0.1671)	0.0040 (0.8332)	0.0150 (0.3553)	0.0180 (0.2216)	0.0025 (0.8784)	0.0398** (0.0249)	0.0149 (0.3393)	−0.0133 (0.3298)	0.0001 (0.9944)	−0.0148 (0.3548)	0.0188 (0.2888)
$\hat{\alpha}_{10}$	0.0294* (0.0642)	−0.0262 (0.1722)	0.0092 (0.6195)	0.0317** (0.0397)	0.0038 (0.8423)	0.0357* (0.0641)	0.0279* (0.0964)	−0.0202 (0.2230)	−0.0177 (0.3708)	−0.0284 (0.1088)	0.0172 (0.3447)
$\hat{\beta}$	0.0067 (0.8007)	0.0087 (0.8021)	−0.0476 (0.1275)	0.0568** (0.0409)	0.0244 (0.5124)	0.0610** (0.0257)	0.0249 (0.4401)	0.0220 (0.3135)	0.0685*** (0.0058)	−0.0023 (0.9292)	0.0267 (0.4604)
F-test	0.1466	0.3947	0.4615	0.4627	0.9413	0.4752	0.0672	0.4620	0.7848	0.6043	0.8163
$\hat{\alpha}$	−0.0227** (0.0113)	−0.0253** (0.0284)	0.0014 (0.4528)	−0.0434*** (0.0000)	−0.0363*** (0.0056)	−0.0524*** (0.0000)	−0.0349*** (0.0017)	−0.0219*** (0.0052)	−0.0506*** (0.0000)	−0.0281*** (0.0021)	−0.0359*** (0.0055)
$\hat{\tilde{\beta}}$	0.0060 (0.8195)	0.0087 (0.8022)	−0.0476 (0.1272)	0.0567** (0.0411)	0.0244 (0.5108)	0.0597** (0.0288)	0.0239 (0.4582)	0.0221 (0.3104)	0.0683*** (0.0059)	−0.0023 (0.9293)	0.0265 (0.4649)
N_j	7320	4394	5886	7338	5810	4976	7356	7294	4868	5922	4954

Notes: *p*-values are reported in brackets. *F*-test denotes Wald tests for the restriction $H_0 : \alpha_2 = \cdots = \alpha_{10} = 0$ (*p*-values are reported). The *t*-test for $\hat{\alpha}$ is one-tailed ($H_1 : \alpha < 0$). The last row reports the number of matches N_j played in each league.

*** denote significance at the level 1%.

** denote significance at the level 5%.

* denote significance at the level 10%.

a time-invariant intercept, and re-estimate the following model:

$$\varepsilon_{i,j} = \alpha_j + \tilde{\beta}_j p_{i,j} + \tilde{\nu}_{i,j}. \tag{2}$$

We report the results for Eq. (2) in the lower panel of Table 2. We find that, on average, the bookmaker commission is significantly lower than zero, at the 5% significance level at least, for all leagues except for Germany, and ranges from 2.19% (Spain) to 5.24% (Portugal). The results of the tests for unbiasedness are not affected by the restricted model in Eq. (2) in terms of the significance of the regression coefficients and the implications of evidence of deviation from unbiasedness in Italy, Portugal, and Greece, which may imply market inefficiency, as we analyse below.

In principle and whenever possible, bettors tend to select the best price that the market has to offer when wagering. Therefore, it is of interest to evaluate the degree of market efficiency when considering maximum odds (instead of mean odds). Table 3 reports the results for the best odds available among the 41 bookmakers that we are considering. As in the case of mean odds (cf. Table 2), we do not find evidence of time-varying intercepts, and hence we consider the restricted model in Eq. (2). Compared to the case of mean odds, the results in Table 3 suggest that bettors can reduce the bookmaker commissions substantially when considering the maximum odds offered by the market, as only three leagues reveal significant (negative) estimates of α, namely Italy, Portugal and Greece. Again, these results are in line with the findings of Forrest et al. (2005, Table 4), who show that the commission is virtually eliminated when using the best available odds. For the same three leagues, we also find evidence of significant estimates of $\tilde{\beta}_j$ at the 5% significance level at least, which confirms that these markets are not unbiased, as was found in the case of mean odds.

We now evaluate the degree of market unbiasedness and whether any biases are large enough to provide profitable opportunities for bettors, which in turn would imply market inefficiency. In particular, we consider the fitted values from the estimation of the models in Eq. (2) for all possible probability values and, for the jth league, derive the following expression, which we call the 'efficiency

Table 3
Estimates of the models in Eqs. (1) and (2) when we consider the maximum of the odds offered on the market.

Max odds	England	Scotland	Germany	Italy	Turkey	Portugal	France	Spain	Greece	Holland	Belgium
$\hat{\alpha}_1$	-0.0224^{**} (0.0333)	-0.0008 (0.4802)	0.0146 (0.1861)	-0.0301^{**} (0.0150)	-0.0075 (0.3361)	-0.0499^{***} (0.0015)	-0.0206^{*} (0.0910)	0.0106 (0.2094)	-0.0209^{*} (0.0796)	0.0052 (0.3626)	-0.0218^{*} (0.0967)
$\hat{\alpha}_2$	0.0132 (0.3400)	-0.0071 (0.6986)	0.0042 (0.8058)	0.0143 (0.3414)	0.0112 (0.5112)	0.0209 (0.2909)	0.0013 (0.9333)	-0.0031 (0.8328)	-0.0253 (0.1379)	-0.0122 (0.4524)	0.0199 (0.2195)
$\hat{\alpha}_3$	0.0124 (0.3437)	-0.0188 (0.3004)	-0.0124 (0.4745)	0.0053 (0.7232)	-0.0209 (0.2341)	0.0077 (0.6835)	0.0209 (0.1790)	-0.0341^{**} (0.0124)	-0.0029 (0.8686)	-0.0087 (0.5598)	0.0145 (0.4306)
$\hat{\alpha}_4$	0.0095 (0.5203)	0.0000 (0.9985)	0.0266 (0.1184)	0.0142 (0.3475)	-0.0016 (0.9222)	0.0193 (0.3160)	-0.0204 (0.2055)	-0.0053 (0.7232)	0.0013 (0.9437)	-0.0136 (0.3924)	-0.0032 (0.8540)
$\hat{\alpha}_5$	0.0273^{*} (0.0666)	-0.0018 (0.9202)	-0.0048 (0.7753)	-0.0004 (0.9774)	-0.0019 (0.9077)	0.0299^{*} (0.0921)	0.0035 (0.8260)	-0.0241^{*} (0.0796)	0.0089 (0.6061)	-0.0043 (0.7817)	-0.0086 (0.6177)
$\hat{\alpha}_6$	0.0033 (0.8130)	-0.0178 (0.3818)	-0.0037 (0.8261)	0.0158 (0.3046)	-0.0105 (0.5445)	0.0183 (0.3124)	0.0036 (0.8197)	-0.0107 (0.4367)	0.0005 (0.9757)	-0.0163 (0.3299)	0.0158 (0.3742)
$\hat{\alpha}_7$	0.0370^{***} (0.0070)	0.0002 (0.9887)	0.0075 (0.6319)	0.0079 (0.5805)	-0.0090 (0.5918)	0.0276 (0.1390)	0.0036 (0.8194)	-0.0075 (0.6020)	0.0012 (0.9393)	-0.0316^{*} (0.0523)	0.0161 (0.3586)
$\hat{\alpha}_8$	0.0270^{*} (0.0591)	0.0148 (0.4046)	-0.0055 (0.7458)	-0.0076 (0.6239)	-0.0108 (0.5305)	0.0151 (0.3723)	0.0272^{*} (0.0728)	-0.0128 (0.3568)	-0.0074 (0.6429)	-0.0090 (0.5792)	0.0062 (0.7165)
$\hat{\alpha}_9$	0.0177 (0.2326)	0.0028 (0.8836)	0.0054 (0.7424)	0.0078 (0.5898)	-0.0076 (0.6469)	0.0340^{*} (0.0734)	0.0083 (0.5917)	-0.0239^{*} (0.0809)	-0.0043 (0.7927)	-0.0227 (0.1567)	0.0093 (0.5990)
$\hat{\alpha}_{10}$	0.0242 (0.1213)	-0.0352^{*} (0.0601)	0.0002 (0.9923)	0.0196 (0.1966)	-0.0034 (0.8572)	0.0246 (0.1974)	0.0212 (0.2055)	-0.0239 (0.2061)	-0.0178 (0.3709)	-0.0365^{**} (0.0374)	0.0087 (0.6386)
$\hat{\beta}$	0.0130 (0.6268)	0.0214 (0.5346)	-0.0342 (0.2750)	0.0658^{**} (0.0174)	0.0341 (0.3661)	0.0681^{**} (0.0121)	0.0328 (0.3052)	0.0319 (0.1495)	0.0753^{***} (0.0025)	0.0090 (0.7234)	0.0364 (0.3129)
F-test	0.2245	0.3885	0.6221	0.7845	0.9026	0.7504	0.1354	0.2059	0.7577	0.4436	0.8313
$\hat{\alpha}$	-0.0054 (0.2872)	-0.0062 (0.3094)	0.0162^{*} (0.0772)	-0.0228^{***} (0.0077)	-0.0126 (0.1771)	-0.0292^{***} (0.0010)	-0.0136 (0.1091)	-0.0043 (0.2996)	-0.0250^{***} (0.0021)	-0.0092 (0.1568)	-0.0138 (0.1455)
$\hat{\tilde{\beta}}$	0.0128 (0.6312)	0.0217 (0.5302)	-0.0338 (0.2809)	0.0660^{**} (0.0171)	0.0337 (0.3717)	0.0669^{**} (0.0139)	0.0316 (0.3226)	0.0327 (0.1398)	0.0750^{***} (0.0026)	0.0086 (0.7338)	0.0362 (0.3146)
N_j	7320	4394	5886	7338	5810	4976	7356	7294	4868	5922	4954

Notes: p-values are reported in brackets. F-test denotes Wald tests for the restriction $H_0 : \alpha_2 = \cdots = \alpha_{10} = 0$ (p-values are reported). The t-test for $\hat{\alpha}$ is one-tailed ($H_1 : \alpha < 0$). The last row reports the number of matches N_j played in each league.

*** denote significance at the level 1%.

** denote significance at the level 5%.

* denote significance at the level 10%.

curve':

$$\hat{G}_j(p_G) = \hat{\alpha}_j + \hat{\tilde{\beta}}_j p_G, \quad p_G \in (0, 1), \tag{3}$$

where $\hat{\alpha}_j$ and $\hat{\tilde{\beta}}_j$ are the estimates of the parameters in Eq. (2), and the related confidence bands are computed as

$$CI_j = \left[\hat{G}_j(p_G) - z_{\alpha/2}\, s.e.\left(\hat{G}_j(p_G)\right),\ \hat{G}_j(p_G) + z_{\alpha/2}\, s.e.\left(\hat{G}_j(p_G)\right)\right], \tag{4}$$

where $s.e.\left(\hat{G}_j(p_G)\right) = \left[\triangledown \hat{G}_j(p_G)'\, V_{WLS}\, \triangledown \hat{G}_j(p_G)\right]^{1/2}$, $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution, $\triangledown \hat{G}_j(p_G) = (1,\ p_G)'$ is the gradient and V_{WLS} is the variance of the WLS estimator.[4]

Figs. 1 and 2 plot the efficiency curves $\hat{G}_j$ in Eq. (3) for each league against $p_G \in (0, 1)$ for the mean and the maximum odds, respectively. For a fixed value of p_G, say $p_G^0 \in (0, 1)$, $\hat{G}_j(p_G^0) = 0$ implies market unbiasedness. Conversely, when $\hat{G}_j(p_G^0) \neq 0$, there is evidence of bias, and the sign of $\hat{G}_j(p_G^0)$ suggests which side might profit from this bias. In particular, $\hat{G}_j(p_G^0) > 0$ would imply market inefficiency, since bettors can achieve positive returns, whereas $\hat{G}_j(p_G^0) < 0$ would entail profits for bookmakers. Thus, we focus on the cases of $\hat{G}_j(p_G^0) > 0$ in order to investigate market inefficiency.

The efficiency curves depicted in Fig. 1 show that, for all leagues except for the German Bundesliga and Dutch Eredivisie, $\hat{G}_j$ tends to increase as the outcome probability p_G increases, since $\hat{\tilde{\beta}}_j > 0$. We therefore find evidence that the market probabilities of underdogs (favourites) overpredict (underpredict) their empirical probabilities on average. This implies that longshots are underpriced and that wagering on favourites is more profitable for bettors. We can interpret these results as evidence in support of the favourite–longshot bias. However, as can be seen from Fig. 1, all of the efficiency curves are below the zero line, except for the case of the largest values of p_G for Italy, Portugal and Greece. Likewise, the related confidence bands, which are depicted in Fig. 1 for the 95% confidence level, show that no significant positive values of $\hat{G}_j(p_G)$ can be achieved. This empirical evidence implies that bettors cannot systematically achieve positive returns and that bookmakers profit in the long run, especially from longshots. Conversely, in the German Bundesliga, bookmakers appear to profit from favourites and not from longshots. This empirical evidence (albeit not significant) suggests a sort of reversed favourite–longshot bias for Germany. Interestingly, for the Dutch Eredivisie, bookmakers seem to achieve significant

[4] Note that the efficiency curves are linear by construction in our framework, as they are based on the estimates of the parameters α_j and $\tilde{\beta}_j$ of the linear model in Eq. (2). However, a non-linear specification of these curves is also possible, for example using non-parametric logistic regression, and might even approximate the tails of the distribution better, and hence, capture the favourite–longshot bias better. This approach is postponed for future research.

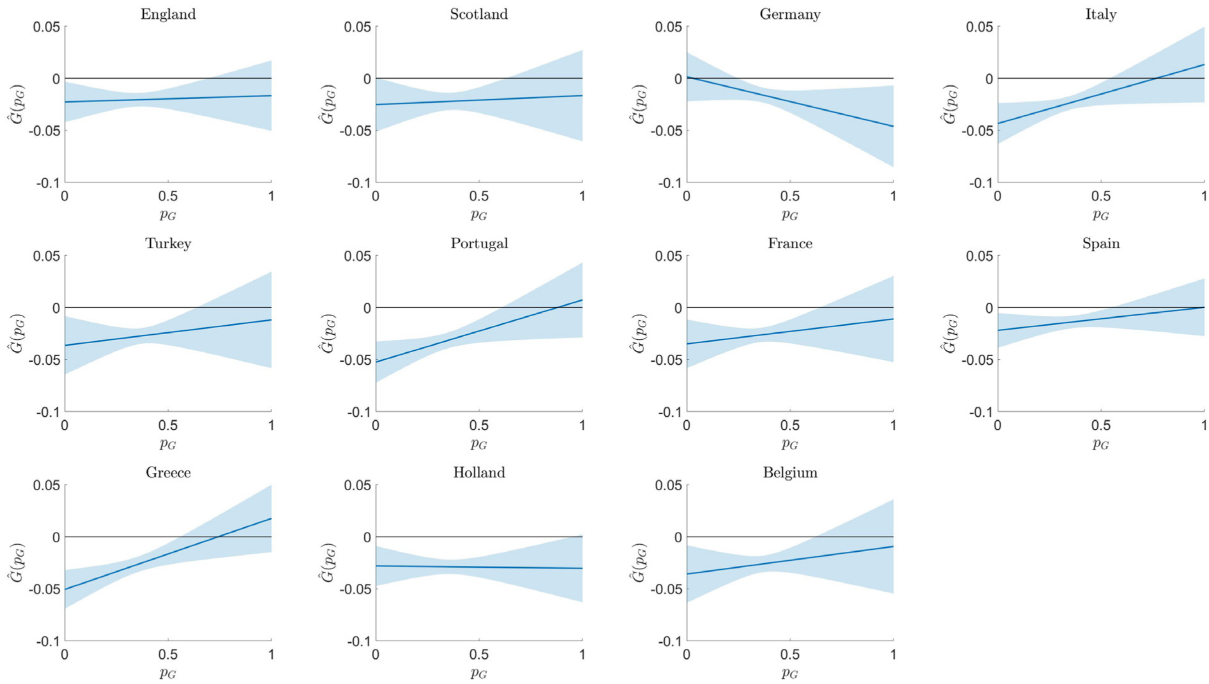


Fig. 1. Efficiency curves $\hat{G}_j(p_g)$ from Eq. (3) and related 95% confidence bands from Eq. (4), computed considering the mean of the odds offered by the betting market.

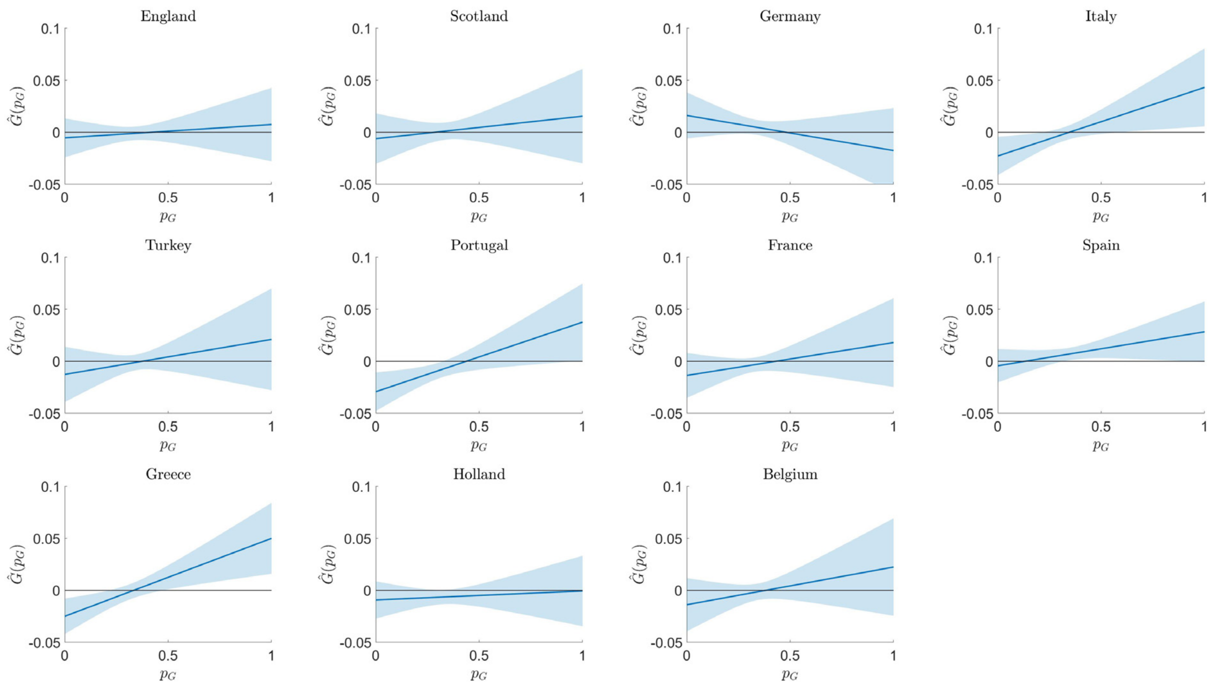


Fig. 2. Efficiency curves $\hat{G}_j(p_g)$ from Eq. (3) and related 95% confidence bands from Eq. (4), computed considering the max of the odds offered by the betting market.

profits for almost any value of p_G. Therefore, while there is evidence of biases, online betting markets are found to be efficient when mean odds are considered.

Fig. 2 shows the efficiency curves for the case of maximum odds. When the best odds offered by the market are considered, $\hat{G}_j$ suggests a lower profitability for bookmakers, as the values for $\hat{\alpha}_j$ are now close to zero (cf. Table 3). In particular, according to the 95% confidence bands depicted in Fig. 2, neither bookmakers nor bettors can achieve significant returns in the following leagues: English Premier League, Scottish Premier League, German Bundesliga, Turkish Super Lig, French Ligue 1, Dutch Eredivisie, and Belgian Jupiler League. Moreover, the Portuguese Primeira Liga shows profits for bookmakers but not for bettors. Hence, the online betting markets are found to be efficient for these eight European leagues if the best odds are considered, in the sense that any bias is too small to overcome the bookmaker commission and allow profitable opportunities for bettors. Conversely, Italian Serie A and Greek Super League allow significant returns for bettors when p_G is large (close to one), and significant profits for bookmakers for smaller values of p_G (close to zero). Therefore, in the case of maximum odds, we find empirical evidence in favor of the favourite–longshot bias, and there is room for profit opportunities for both bettors and bookmakers. Finally, an interesting case is given by the Spanish Liga, where picking the best odds on the market appears to deliver significant positive returns for bettors (for outcomes with probabilities from around 0.3 to around 0.9) but not for bookmakers.

One way of exploiting these empirical findings using a simple but profitable betting strategy is described in the next section in terms of both *ex ante* and *ex post* forecasting performances.

3.3. Implications of market inefficiency: A simple and profitable betting strategy

This section proposes a betting strategy which aims to exploit the market inefficiencies found in the previous section. Fig. 2 shows that the efficiency curves in Eq. (3) are significantly positive for three European online betting markets when picking the maximum odds offered by the market. Indeed, as was discussed in Section 3.2, the Italian and Greek leagues show positive values of $\hat{G}_j(p_G)$ associated with the largest probabilities p_G (the smallest odds), while for Spain we have $\hat{G}_j(p_G) > 0$ for central values of p_G. Our betting strategy for league j can be summarised as follows:

1. Estimate the model in Eq. (2) by WLS, as described in Section 2, considering the observations until season T^* as information set $\mathcal{I}_i$:

$$\varepsilon_{i,j} = \alpha_j^* + \tilde{\beta}_j^* p_{i,j} + \tilde{\nu}_{i,j}, \quad i = 1, \ldots, N_j^{T^*}, \tag{5}$$

where $N_j^{T^*}$ is the number of matches played in the jth league in seasons $t = 1, \ldots, T^*$.

2. Using the results obtained in step 1, compute the efficiency curve for league j up to season T^* as

$$\hat{G}_j^*(p_G) = \hat{\alpha}_j^* + \hat{\tilde{\beta}}_j^* p_G, \quad p_G \in (0, 1), \tag{6}$$

and the related confidence bands as

$$CI_j^* = \left[\hat{G}_j^*(p_G) - z_{\alpha/2} s.e. \left(\hat{G}_j^*(p_G)\right),\ \hat{G}_j^*(p_G) + z_{\alpha/2} s.e. \left(\hat{G}_j^*(p_G)\right)\right]. \tag{7}$$

3. Define the 'profitable probability range' as $P_j^* = \{p_G \in (0,1) : \underline{CI}_j^* > 0\}$, where $\underline{CI}_j^*$ denotes the lower bound of the confidence interval in Eq. (7), the threshold probabilities as the infimum and supremum of the profitable probability range, i.e., $p_{G,j}^{*(L)} = \inf_{p_G \in (0,1)} P_j^*$ and $p_{G,j}^{*(U)} = \sup_{p_G \in (0,1)} P_j^*$, and the related threshold odds as $o_j^{*(L)} = \left(p_{G,j}^{*(L)}\right)^{-1}$ and $o_j^{*(U)} = \left(p_{G,j}^{*(U)}\right)^{-1}$.
4. Systematically wager on all matches of league j whose odds are in the 'profitable odds range', $O_j^* = \left[o_j^{*(L)},\ o_j^{*(U)}\right]$, for either all seasons after T^* (out-of-sample forecast) or all seasons in the sample (in-sample forecast).

We adopt the above betting strategy for evaluating both in-sample (*ex post*) and out-of-sample (*ex ante*) forecasting performances. Note that our approach estimates the profitable odds range rather than choosing it arbitrarily as in previous analyses (see e.g. Direr, 2013). The results are reported in Table 4.

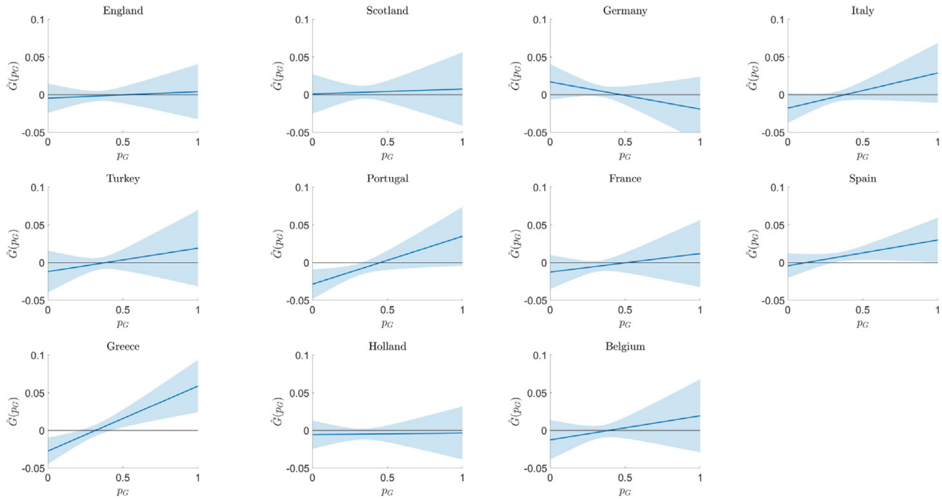
First we focus on the in-sample forecast, which considers the whole information set available, i.e., $T^* = 2016/17$. The results in the upper panel of Table 4 show that the betting strategy delivers positive mean returns for all three European leagues which were found in Section 3.2 to be inefficient. In particular, by systematically betting on odds inferior to 1.67 and 2.08, we achieve mean returns of 2.09% and 2.71% for the Italian Serie A and Greek Super League, respectively, while betting on matches played in the Spanish Liga with odds in the range 1.09–3.12 delivered a mean return of 2.12%.

Next, we investigate whether abnormal out-of-sample returns can be obtained using the proposed betting strategy. We evaluate the out-of-sample forecasting performance by fixing $T^* = 2015/16$ as our information set and using the 2016/17 season as our out-of-sample period. In this out-of-sample forecasting exercise, we have extended the sample period to now incorporate the end of season 2016/17 as well (i.e., the matches played from March to June 2017). The graphs displayed in the bottom panel of Table 4 depict the efficiency curves $\hat{G}_j^*(p_G)$ for all leagues, with their related 95% confidence bands CI_j^*, computed as in Eqs. (6) and (7), respectively. We note from these figures that, in line with the results obtained above, no profitable probability range is found for the eight *efficient* leagues or Italy when considering $T^* = 2015/16$ as the information set; i.e., the estimated lower bound of the confidence interval $\underline{CI}_j^*$ is below the zero line for all values of p_G. Conversely, for Greece and Spain there are values of p_G for which the condition $\underline{CI}_j^* > 0$ is satisfied. Hence, following steps 3 and 4 of our betting strategy, we compute the profitable probability range, P_j^*, and the corresponding profitable odds range, O_j^* (cf. the middle panel of Table 4).

Table 4
Upper panel: in-sample forecasting performance of the betting strategy described in Section 3.3. Middle panel: out-of-sample forecasting performance of the betting strategy described in Section 3.3. Bottom panel: efficiency curves $\hat{G}_j^*(p_g)$ from Eq. (6) (blue lines) and related 95% confidence bands used in the out-of-sample forecast, computed as per Eq. (7).

In-sample forecasting performance			
	Italy	Spain	Greece
Percentage mean return	2.09	2.12	2.71
Number of bets	896	3618	1451
Percentage of correct bets	70.76	55.20	65.75
Profitable probability range	$P^* = [0.5988, 1)$	$[0.3200, 0.9174]$	$[0.4808, 1)$
Profitable odds range	$O^* = (1, 1.67]$	$[1.09, 3.12]$	$(1, 2.08]$
Out-of-sample forecasting performance (up to season $T^* = 2015/16$)			
Percentage mean return	-	2.25	1.35
Number of bets	-	380	171
Percentage of correct bets	-	44.69	63.74
Profitable probability range	-	$P^* = [0.3105, 1)$	$[0.4348, 1)$
Profitable odds range	-	$O^* = (1, 3.22]$	$(1, 2.30]$

Efficiency curve (up to season $T^* = 2015/16$)



We observe that profit opportunities for bettors can be pursued in the Greek Super League and the Spanish Liga when wagering on match outcomes with odds inferior to 2.30 and 3.22 or, equivalently, implied probabilities larger than 0.4348 and 0.3105, respectively. The results reported in the middle panel of Table 4 show that the mean returns are positive for both leagues. In particular, in the out-of-sample period (i.e., the entire 2016/17 season), the betting strategy delivers mean returns of 1.35% and 2.25% for Greece and Spain, respectively.

4. Conclusions

Over recent decades, online betting markets have developed, evolved and thrived, and scholars' interest in investigating the characteristics of these markets has increased. This paper focuses on the degree of efficiency in the online betting market for European football using a large data set. Considering the mean market odds, we provide evidence that online betting markets are (weak-form) efficient and that any biases that are detected provide extra profits for bookmakers. However, thanks to a highly competitive market, bettors may choose between many bookmakers and pick the best odds offered by the market. Repeating the analysis using maximum odds, we find that the majority of the online betting markets are efficient, but we also find evidence of inefficiencies that can be exploited in order to define profitable betting strategies. In particular, our analysis shows that one of the most popular deviations from unbiasedness in betting markets, the favourite–longshot bias, is indeed present in three European football markets. We show that a simple betting strategy which exploits this bias leads to abnormal positive returns for bettors, after considering bookmaker commissions. Moreover, our results show that online bookmakers' commissions have not changed significantly over time over the period from 2006 to 2017, but appear to differ across leagues.

Acknowledgments

We are grateful to the Editor, Rob Hyndman, the Guest Editor, Ian McHale, and two anonymous referees for their helpful comments on the earlier versions of this paper. We thank Giuseppe Cavaliere, Luca Fanelli and Michele Costa for their helpful comments and suggestions and Iliyan Georgiev for the stimulating initial discussions about this paper. We are solely responsible for all errors.

References

Angelini, G., & De Angelis, L. (2017). PARX model for football matches predictions. *Journal of Forecasting*, *36*(7), 795–807.

Boshnakov, G., Kharrat, T., & McHale, I. G. (2017). A bivariate Weibull count model for forecasting association football scores. *International Journal of Forecasting*, *33*(2), 458–466.

Boulier, B. L., & Stekler, H. O. (2003). Predicting the outcomes of National Football League games. *International Journal of Forecasting*, *19*(2), 257–270.

Brown, A., Rambaccussing, D., Reade, J. J., & Rossi, G. (2017). Using social media to identify market inefficiencies: evidence from twitter and betfair. *Economic Inquiry*, *56*, 1748–1763.

Brown, A., & Reade, J. J. (2017). Combining prediction markets and forecasting contests. Available at SSRN: https://ssrn.com/abstract=2944675.

Clements, M. P., & Reade, J. J. (2016). *Forecasting and forecast narratives: The bank of England inflation reports. Technical report, Discussion paper*. Economics Department, University of Reading.

Croxson, K., & Reade, J. J. (2014). Information and efficiency: goal arrival in soccer betting. *Economic Journal*, *124*, 62–91.

Direr, A. (2013). Are betting markets efficient? Evidence from European football championships. *Applied Economics*, *45*, 343–356.

Dixon, M. J., & Pope, P. F. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, *20*, 686–700.

Fama, E. F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, *25*, 383–417.

Forrest, D., Goddard, J., & Simmons, R. (2005). Odds-setters as forecasters: the case of English football. *International Journal of Forecasting*, *21*, 551–564.

Forrest, D., & Simmons, R. (2000). Forecasting sport: the behaviour and performance of football tipsters. *International Journal of Forecasting*, *16*, 317–331.

Franck, E., Verbeek, E., & Nüesch, S. (2010). Prediction accuracy of different market structures — bookmakers versus a betting exchange. *International Journal of Forecasting*, *26*, 448–459.

Franck, E., Verbeek, E., & Nüesch, S. (2013). Inter-market arbitrage in betting. *Economica*, *80*, 300–325.

Glickman, M. E., & Stern, H. S. (1998). A state-space model for National Football League scores. *Journal of the American Statistical Association*, *93*, 25–35.

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

Grant, A., & Johnstone, D. (2010). Finding profitable forecast combinations using probability scoring rules. *International Journal of Forecasting*, *26*, 498–510.

Griffith, R. M. (1949). Odds adjustment by American horse-race bettors. *American Journal of Psychology*, *62*, 290–294.

Hurley, W., & McDonough, L. (1995). Note on the Hayek hypothesis and the favorite–longshot bias in parimutuel betting. *American Economic Review*, *85*, 949–955.

Ioannidis, C., & Peel, D. A. (2005). Testing for market efficiency in gambling markets when the errors are non-normal and heteroskedastic — an application of the wild bootstrap. *Economics Letters*, *87*, 221–226.

Koopman, S. J., & Lit, R. (2015). A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League. *Journal of the Royal Statistical Society, Series A*, *178*(1), 167–186.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, *32*, 1353–1363.

Lessmann, S., Sungb, M.-C., & Johnson, J. E. V. (2010). Alternative methods of predicting competitive events: an application in horserace betting markets. *International Journal of Forecasting*, *26*, 518–536.

Levitt, S. (2004). Why are gambling markets organised so differently? *The Economic Journal*, *114*, 223–246.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13–22.

McHale, I., & Morton, A. (2011). A Bradley–Terry type model for forecasting tennis match results. *International Journal of Forecasting*, *27*(2), 619–630.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*. New York: National Bureau of Economic Research.

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

Reade, J. J. (2014). Information and predictability: bookmakers, prediction markets and tipsters as forecasters. *Journal of Prediction Markets*, *8*, 43–76.

Rydall, R., & Bedford, A. (2010). An optimized ratings-based model for forecasting Australian Rules football. *International Journal of Forecasting*, *26*, 511–517.

Sorensen, P. N., & Ottaviani, M. (2008). The favorite–longshot bias: an overview of the main explanations. In D. B. Hausch, & W. T. Ziemba (Eds.), *Handbook of sports and lottery markets*. Amsterdam: North-Holland.

Spann, M., & Skiera, B. (2009). Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. *Journal of Forecasting*, *28*, 55–72.

Thaler, R., & Ziemba, W. (1988). Parimutual betting markets: racetracks and lotteries. *Journal of Economic Perspectives*, *2*, 161–174.

Štrumbelj, E., & Šikonja, R. (2010). Online bookmakers' odds as forecasts: the case of European soccer leagues. *International Journal of Forecasting*, *26*(3), 482–488.

Štrumbelj, E., & Vračar, P. (2012). Simulating a basketball match with a homogeneous Markov model and forecasting the outcome. *International Journal of Forecasting*, *28*(2), 532–542.

Vaughan Williams, L. (2005). *Information efficiency in financial and betting markets*. Cambridge: Cambridge University Press.

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

Štrumbelj, E. (2014). On determining probability forecasts from betting odds. *International Journal of Forecasting*, *30*, 934–943.

Woodland, . L. M., & Woodland, . B. M. (1994). Market efficiency and the favorite–longshot bias: the baseball betting market. *Journal of Finance*, *49*, 269–279.

Woodland, L. M., & Woodland, B. M. (2001). Market efficiency and profitable wagering in the National Hockey League: can bettors score on longshots? *Southern Economic Journal*, *67*, 983–995.