Chance or Ability? The Efficiency of the Football Betting Market Revisited

Gross, Johannes and Rebeggiani, Luca

FOM University of Applied Sciences, Fraunhofer FIT

June 2018

Chance or Ability? The Efficiency of the Football Betting Market Revisited

Johannes Gross
Luca Rebeggiani*

This Version: June 2018

Keywords: Gambling, Sports Betting, Market Efficiency
JEL-Classification: G14, L83, Z2

**Corresponding author*
Fraunhofer FIT and FOM University of Applied Sciences
Schloss Birlinghoven
D-53754 Sankt Augustin
luca.rebeggiani@fit.fraunhofer.de

Abstract

The extent of market efficiency induced by rational behaviour of market participants is central for economic research. Many economists have already examined sports-betting markets as a laboratory to better understand trading behaviour and efficiency of stock prices while avoiding to jointly test the hypothesis of a correct capital market model. The following paper will investigate whether the European football betting market fulfils the efficiency paradigm introduced by Fama (1970) with a unique dataset allowing for an investigation of the German betting market in view of its regulatory changes recently. The analysis contributes to the literature by conducting a variety of empirical strategy including rational expectation frameworks and an ordered choice model to stress the ex post market performance from a weak and semi-strong form perspective. In view of existing market distortions as taxes, switching costs of changing betting providers and limitation in competition, the results of the analysis are indicative of a rational market equilibrium surprisingly close to the efficiency benchmark.[1]

[1]The analysis is based on a dataset provided by Sportradar. We are grateful to the whole Sportradar team for their support. We also gratefully acknowledge comments on an earlier draft of this paper from participants of the 2017 ESEA conference in Paderborn and of the 2017 IASE conference in Shanghai.

1 Introduction

To what extent are market outcomes efficiently induced by rational behaviour of market participants? This question is still one of the most controversially discussed in economics, with application to many different market types. The following paper will examine one very particular application: We will investigate whether the European football betting market fulfils the efficiency paradigm introduced by Fama (1970).

The degree to which markets incorporate information is one of the most important questions in economic research, especially for economists dealing with financial markets.[2] Hence, the empirical literature tried to test the degree of market efficiency extensively within stock exchanges and other markets for financial assets (Fama (1991)). However, as discussed by Fama, it is only possible to test the efficiency hypothesis, e.g. whether information is properly reflected in asset prices, in the context of a correct pricing model, the so-called Joint Hypothesis Problem. More recently, several studies examined the performance of sports betting markets and pointed out the advantage of a fixed terminal value of its assets realizing within an ex ante fixed period of time. Although these markets constitute only a small part of most economies compared to other financial markets, the characteristic of a reduced pricing problem opens up significant opportunities for economic analysis (Sauer (1998)). Given other similarities between trading with financial assets and wagering markets, the betting market provides even a superior setting for testing market efficiency in general according to Thaler and Ziemba (1988).

This empirical study is based on an exclusive dataset of European online bookmaker prices provided by Sportradar, a multinational sports data provider and institution employed by the UEFA to monitor the football gambling market. The analysis focuses on matches of major European leagues since they constitute one of the most liquid and important betting markets (Vlastakis et al. (2009)). Precisely, the data contains closing odds of 14 online bookmakers, who report price adjustments to Sportradar. While in the betting market literature empirical studies are mainly conducted with publicly available online databases collected at a fixed point during the week,[3] this database enables to test informational efficiency in a more rigorous way. Furthermore, the sample contains some bookmakers mainly or only active in the German market and allows an analysis with a special reference to the German betting market, a market characterized by a state monopoly for betting services until 2012 and afterwards partially liberalized.

This empirical analysis will be divided in two parts. First, the informational efficiency will be analysed by examining the unbiasedness and neutrality of betting prices. This can be seen as a weak-form efficiency test while stressing the assumption that no abnormal returns can be achieved using just historical price information. In the second part, the

[2] See on this e.g. the overview by Williams (2005)

[3] The main source of historical odds used in the existing literature is www.football.data.co.uk.

efficiency of the betting market is challenged by a statistical forecasting model using publicly available information from which betting strategies are derived. The analysis will include the English betting market and the first two divisions within the English football league system as well as British licensed bookmakers.[4] The focus on this market is mainly justified with its size, its tradition and its liberal market form. In addition, the German market will be reviewed in detail, e.g. Bundesliga 1 and 2 matches and odds from bookmakers that are available for German costumers. While the British betting market was liberalized and experienced a reduction in betting taxes already before the first period of historical odds (2006 - 2016) in the dataset, the German market was officially characterized by a state monopoly until 2012 and was afterwards affected by a tax introduction on stakes, including private bookmaker services, following a partial liberalization. Therefore, comparing the efficiency of these two markets over time may also reveal some effects of a change in the market structure and legal framework of betting markets. However, a clear separation of the two markets will not be possible since we are examining online betting odds in this analysis and wagering on foreign divisions is not excluded. Nevertheless, bookmakers may anticipate closely the domestic betting demand for the English and German divisions respectively.

The structure of this analysis will be as follows: First, the concept of market efficiency will be outlined briefly in chapter 2. Subsequently, chapter 3 will provide a short introduction to the European betting market and a theoretical perspective on gambling. This will be followed by a literature review of efficiency analysis within betting markets (chapter 4). The empirical analysis presented in chapter 5 will be structured in two parts, beginning with tests of the weak-form market efficiency hypothesis, e.g. the informational efficiency of bookmaker prices. Following this, an ordered-probit model will be estimated to challenge the efficiency of the bookmaker market which is described with the semi-strong-form efficiency hypothesis by deriving forecasts based on publicly available information. Finally, the economic significance of different betting strategies will be examined and the results of the analysis will be summarized in a conclusion (chapter 6).

[4]There are parallel league systems in the UK but the main divisions are the Premier League and First League.

2 Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) has mainly been shaped by Eugene Fama. In one of his early speeches, he claimed the following:

In an efficient market, competition among the many intelligent participants leads to a situation where, at any point in time, actual prices of individual securities already reflect the effects of information based both on events that have already occurred and on events which, as of now, the market expects to take place in the future. (Fama (1965))

This implies that prices should reflect at any time all the relevant and available information affecting the underlying security. A more economically sensitive definition is given by Jensen (1978) where markets are characterized as efficient with respect to an information set σ_t if it is impossible to make economic profits by trading based on the respective information set σ_t. As a result, costly fundamental analysis of securities or expert knowledge about market expectations lose their economic rationale. However, Fama (1970) claims that the following sufficient conditions for market efficiency need to be assumed. Precisely, there are no transactions costs in trading securities and the information is costlessly available to all market participants. Additionally, it is assumed that all agents agree on their future expectations and implications of current information. In Grossman and Stiglitz (1980) the characteristics of an efficient market is examined with a noisy rational expectations model and endogenous choice of being informed. The informativity of the price system depends on the number of individuals who are informed and on several other critical parameters in the model. Precisely, these are the cost of acquiring information, how informative the price system relative to the market random noise components is, and how informative the information obtained by an informed individual might be. In conclusion, investors need to be attracted by some form of expected profit as normal returns to their investment, otherwise there would be no incentive to analyse the market and ensure its informational efficiency. Hence, an efficient amount of inefficiency needs to be in the market to ensure a stable equilibrium close to the efficiency benchmark.

Consequently, assigning markets strict efficiency characteristics has to be seen due to non-zero costs for transaction and information accumulation as well as heterogeneity of expectations in reality as an approximation or benchmark case. Therefore Fama (1991) suggests that any finding of market inefficiency is not surprising per se, but proves the importance of measuring the extent of inefficiency within the market.

Fama (1970) adjusted the EMH by defining different information sets incorporated by the market price whereupon the tautology of a general efficiency claims becomes empirically falsifiable:

- The **Weak-Form Efficient Market Hypothesis**, in which the information set σ_t is defined to be solely the historical price information set derived from past price history of the market at time t.
- The **Semi-Strong-Form Efficient Market Hypothesis**, in which σ_t includes all public available information at time t extending the historical price information set.
- The **Strong-Form Efficient Market Hypothesis**, in which σ_t incorporates all information, from public as well as private nature, at time t.

However, one of the fundamental problems of detecting inefficiencies in financial markets is the underlying asset-pricing model defining "proper" prices to securities, given the uncertainty about future pay-offs, even over an infinite horizon as in the stock market. The so called Joint-Hypothesis Problem (Fama (1991)) arises since empirical tests can fail either because one of the two hypotheses, the hypothesis that the pricing model is correct or the market is efficient, is false or because both parts of the joint hypothesis are incorrect (Jensen (1978)).

Efficiency Hypothesis Applied to Betting Markets

Since betting markets are defined by a fixed terminal value of bets, the markets' assets, realizing within an ex ante fixed period of time, the joint-hypothesis problem of other financial markets can be neglected. Many economists have already examined sports-betting markets as a laboratory to better understand trading behaviour and efficiency of stock prices while avoiding to jointly test the hypothesis of a correct capital market model (Williams (2005)).

Precisely, the betting prices should be proportional to the true probability of the underlying event in an efficient market, since otherwise competing bookmakers with fair odds would get the betting demand solely. Therefore, examining the informational efficiency of bookmaker prices can be seen as a clear testing strategy of weak-form market efficiency. Moreover, since higher potential payoffs in case of winning the bet need to be proportional to lower probabilities of the underlying outcomes, the expected betting returns should be identical, independently of the applied betting strategy. Consequently, in betting markets weak-form efficiency implies that neither bookmaker nor punters can achieve abnormal returns using historical price information, where abnormal returns are defined by divergence from the bookmaker market takeout (Kuypers (2000)). Translating the semi-strong-form efficiency hypothesis into the betting market setting leads to the claim that incorporation of publicly available information should neither improve the accuracy of outcome predictions based on bookmaker odds nor lead to superior returns. The strong-form includes also private information to that claim. However, in football it is common to assume that there is no great deal of private or insider information, since games are usually played in front of large audiences and the conditions of a team and their individual players are extensively discussed in the media.[5] Therefore, the rest of the analysis is exclusively focusing on weak and semi-strong-form market efficiency in case of the football betting market.

3 The Sports Betting Market

In its basic form, sports betting has been around for hundreds of years, as betting on athletic events dates back to the Roman empire (Sauer (1998)). The market for sports gambling has steadily gained popularity while the nature of traditional betting markets has been strongly affected by changes in information and communication technology (Vlastakis et al. (2006)). Nowadays, there are several market forms offering in coexistence betting services for sport events. Hence, their individual market characteristics can be assumed to be significantly affected by each other, as examined by Franck et al. (2010) and Vlastakis et al. (2009). This study will focus exclusively on online betting markets, one of the fastest growing service sectors in Europe with annual growth rates of up to 15% (Peren and Clement (2016)). Particularly, the European market is today the largest market for online gambling worldwide and occupies more than 47.6% per cent of the €34.6 bn online generated global gaming gross win (stakes minus winnings) in 2015 (EGBA (2016)). Figure 1 illustrates the gross gaming revenues of the European online and offline market in 2015 and their expected evolution until 2020. It is expected that the market share of online betting will increase from 17.5% to 22.5% while the overall market revenue will rise from €88.2 bn to €109.2 bn.

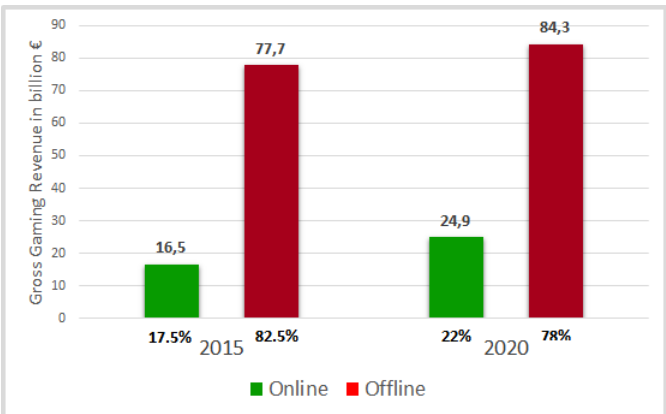

Figure 1: EU Gross Gaming Revenue
Source: EGBA 2016

In the following, a brief description of the two main forms of betting markets will be

[5]See e.g. Kuypers (2000). However, some evidence from studies on match-fixing (e.g. Rebeggiani and Rebeggiani (2013)) may suggest to be more careful on that.

given by highlighting their differences and the suitability of online betting markets for an efficient market analysis.

3.1 The Offline Fixed-Odd Bookmaker Market

The traditional bookmaker market is constituted by coupon betting in physical betting shops where prices are characterized as "fixed-odds", hence the betting quotes or so-called odds do not vary over time in response to changes in betting volumes. These offline services are still accountable for more than two thirds of the overall gambling market (EuropeanCommission (2012)). Dixon and Pope (2004) describe the typical fixed-odds pricing mechanism with a panel of professional odds compilers who meet on a weekly basis and use their subjective judgement and expert knowledge of the football teams' current state to set the odds on home team win, away team win and draw results one week in advance. Therefore, it can be assumed that the fixed-odd betting lines contain less information and more supply risk than betting prices that are dynamically adjusted until the match starts (Vlastakis et al. (2009)). This is the case for the rising online betting market on which we are focusing hereinafter. Another notable feature of the physical betting market is the anonymity of the betting audience. Usually, there is no requirement of registering in a betting shop to submit a wager, which is why bookmakers are not able to identify the betting behaviour of individual punters over time.

3.2 The Online Bookmaker Market

Online bookmaker services perform differently from the offline fixed-odd market, even though they provide identical assets and are mostly operated by the same service providers (Peren and Clement (2016)). Whereas punters remain anonymous in the offline market, at online betting platforms their identities are revealed through the requirement of a personal account. Therefore, a punter faces switching costs including the time and cost of money transfer by opening a new online account at a competitor's platform. Furthermore, online betting providers have the possibility to gather information about the trading behaviour of individual clients and reserve the right to limit the individual maximum stakes or to close an account in case of a history of extreme betting profits (Franck et al. (2013)). This characteristic potentially impairs the forces of informed traders to ensure an efficient market equilibrium and therefore has to be mentioned before conclusions across markets in general are drawn.

However, the main advantage of the online market is its dynamic quoting process in contrast to the offline services and thereby the ability of bookmakers to incorporate new information in their price adjustments. This includes, besides the new information about the contestants itself, primarily the information about the distribution of bets and thus the expectations of the betting public (Vlastakis et al. (2006)). Apart from this, the online market stands out with high price transparency, whereby bettors can easily compare the odds of various bookmakers and identify favourable price discrepancies at low searching costs[6] (Franck et al. (2011)). Being a relatively young and new market, the online gambling sector in Europe has been steadily developing, albeit more slowly in recent years due to increasing regulatory pressure (EuropeanCommission (2012)). In conclusion, while so far the literature was mainly examining fixed-odd betting prices or online odds that were collected in advance, the following examination of closing odds, which capture the final price of the online market's dynamic price adjustment process, may reveal new insights.

3.3 Regulation and Market Characteristics

In most European countries gambling markets are subject to strict regulation, in some cases they are even entirely controlled by the government (Rebeggiani (2009)). While the main justification of strict regulation and high tax burden for gambling markets is based on pathological gambling or addiction, it suits as well as "highly significant area for governments seeking continuous or compensatory revenue for the public purse" (Nikkinen (2014)). However, due to the rise of online platforms, which are based offshore but are available for domestic costumers, the national regulation and taxation schemes have recently come under pressure. In response to the threat of decreasing tax incomes and a rising grey market, the UK was the first country that reformed betting taxation by considerably lowering its effective level and declaring online betting as legal (Vlastakis et al. (2006)). As Forrest et al. (2005) note, the removal of betting taxes potentially improved the overall efficiency of betting markets by attracting well-informed professional bettors who are capable of identifying mispriced bets. Nowadays, the British market is the largest betting market in Europe and is mainly dominated by a few large service providers like WilliamHill and Ladbrokes (Vlastakis et al. (2006)).

Moving the focus to the German market, a different framework emerges. Until 2012, the state monopolist Oddset was the only legally accepted betting provider (Rebeggiani (2015)). In the same year, a new gambling state treaty was adopted to open up the way for private betting companies to enter the German market legally via state-issued licenses. However, none of the planned licenses has been granted to any company so far, leaving private bookmakers in a legally unregulated but tolerated grey area where companies are taxed but their legal status remains uncertain (Rebeggiani and Breuer (2017)). Peren and Clement (2016) examined the German online betting market in 2014 and estimate the number of active online betting portals for German costumers at round 130. Nearly 95 percent of total turnover of German sport bets is actually estimated to be generated in "grey" or completely illegal markets (Rebeggiani and Breuer (2017)). While black

[6] There are several platforms that compare bookmaker odds as betexplorer.com and oddschecker.com or identify arbitrage opportunities as surebet.com.

market providers want by definition to stay illegal, "grey market" companies normally wish to have a regular public licence or have already applied for one. Public revenues have significantly risen since the new state treaty collects starting from 2012 taxes even from not fully licensed betting providers.

All in all, these two national markets are characterized by different regulation schemes and some providers are only accessible for the betting public in one country . Nevertheless, both markets are affected by an increase in competition in view of a substantial grey market and therefore provide an interesting setting for an efficient market analysis.

3.4 A Theoretical Background of Betting Markets

From a theoretical perspective the outcome of a football game consists of three disjunct events ignoring the actual score result. A home victory (H), a draw (D) or an away victory (A): $\Omega = \cup H \cup D \cup A$. These outcomes provide the underlying events of the football result online betting market, in which we are interested in the following. By quoting odds for these individual events, bookmakers offer Arrow-Debreu securities[7] denoted by $(\sigma_H, \sigma_D, \sigma_A) \in \mathbb{R}^3_+$. In the European football betting market, decimal odds are common, determining the payout factor including the returned stake in case of winning for each of the match outcomes. Hence, within wagering markets, odds constitute prices supplied by competing bookmakers who want to attract the demand of the betting audience and act as market makers while the bet demand side is left with a take-it-or-leave-it decision, by either hitting the market quotes and decide about the amount of stake or refrain from participating. Therefore, bookmaker markets can be classified as quote-driven markets in analogy to similar financial markets (Franck et al. (2010)). Moreover, in this simple speculative market for future cash flows, transactions can be denoted as zero-sum transactions where the bookmaker takes the reverse position to the punter. Hence, bookmakers are confronted with a substantial risk in case of mispriced betting lines and in presence of bettors whose skills allow them to achieve positive expected profits (Levitt (2004)). The following will briefly provide some theoretical foundations how equilibrium prices are determined in a market with uncertainty and state contingent assets.

If bookmakers have deep pockets such that it can be assumed that they are able to honour their debt in each case, if they are risk neutral and if there is perfect competition in the market, then standard General Equilibrium Theory tells us that Arrow-Debreu securities are in perfectly elastic supply and by rational expectations correctly priced (Mas-Colell et al. (1995)), i.e. the asset price is the inverse of the probability of the event. The existence of these so-called rational expectation equilibria is described in Radner (1979) as generic and defined by an equilibrium that "*reveals to all traders the information possessed by all of the traders taken together*".

[7] State contingent assets which pay one unit of a numeraire, if a particular state of the world is reached and zero otherwise, see Mas-Colell et al. (1995).

Alternatively, one can analyse the market from a Game Theoretic perspective. Since the Arrow-Debreu securities are individually perfectly homogeneous, the market can be modelled as a Bertrand competition with N firms (see Mas-Colell et al. (1995)). By further assuming that all bookmakers share the same cost $(c_H, c_D, c_A) \in \mathbb{R}^3$ and subjective probability over the events $(p_H, p_D, p_A) \in [0,1]^3$, we can write the bookmakers' expected profit as:

$$E[\pi^{BM}] = \max_{\sigma_e} Q(\sigma_e, P)(1 - p_e\sigma_e - c_e) \tag{1}$$

where the punters' demand function is given by $Q(\sigma_e, P)$ depending on the bookmaker's price vector $(\sigma_H, \sigma_D, \sigma_A)$ and their own believes about the probability distribution of the match outcomes P.

Thus, the bookmaker's expected profit is the money at stake minus cost of provision and the expected payouts for each of the outcomes $e \in \{H, D, A\}$ where the punters receive the offered odds multiplied by their stakes $Q(\sigma_e, P)$. We know that in a unique Nash-Equilibrium, price equals marginal cost:

$$1 = p_e\sigma_e + c_e \Longleftrightarrow \sigma_e = \frac{1 - c_e}{p_e} \tag{2}$$

Hence, from this perspective we receive the same equilibrium price as we argued using the Rational Expectations Equilibrium but adding the cost-wedge c_e. The assumption of unanimity in the subjective probability of the event p_e is very strong, but there is also a reasonable rationale for it. In the real world, as mentioned above, online bookmakers constantly update their quotes ensuring the ability to learn from competitors' pricing behaviour. Since quoting "wrong" odds is a costly signal, i.e. bidders can take those odds, we are in a costly signalling game in which bookmakers learn their competitors' subjective probabilities and update their own. Hence, if all competitors are symmetric, this should aggregate to the same subjective probability across agents. There are analogous arguments for the unanimity in cost of provision since cost differences would exclude all but the cheapest firm from the market.[8]

Since bookmakers are bearing risk and transaction costs $\Longrightarrow c_e \neq 0$, they are charging commission for their services implying that the sum of the inverse odds $\sum_e \frac{1}{\sigma_e} > 1$ exceeds the probability of a sure event $\Omega = \cup H \cup D \cup A$. Assuming now the costs of providing bets are homogeneous across outcomes $c = c_H = c_D = c_A$, the implied commission or bookmaker margin is then given by:

[8] If agents had different beliefs and moved only simultaneously as in the fixed-odd market, there would be a winner's curse attached to the one with the highest odds. This bookmaker secures the whole market but will offer too much money similar to a common value auction. Thus, one should expect greater cost wedges there.

$$m = \sum_e \frac{1}{\sigma_e} - 1 = \sum_e \frac{p_e}{1-c} - 1 = \frac{1}{1-c} - 1 = \frac{c}{1-c} \tag{3}$$

Consequently, to receive the correct implied probabilities of the bookmaker market, we have to rescale the inverse odds by the commission charged,[9] precisely:

$$P_e^{IP} = \frac{1}{\sigma_e}\frac{1}{1+m} = \frac{1}{\sigma_e}\frac{1}{\sum_e \frac{1}{\sigma_e}} \tag{4}$$

This implied probability will be used in the following as a forecast instrument, e.g. the bookmaker prediction for outcome e, to assess the informational efficiency of the betting market.

The exact business strategy of bookmakers is not well known and outside the scope of this analysis. If it can be assumed that bookmakers are systematically at least as good as the gambling audience at predicting the outcomes of games and always set the correct prices in line with the true probability, the markets can be assumed to be efficient leading to a bookmaker return or takeout in the form of the margin m.

Since the betting market is characterized by zero-sum transactions as already mentioned, the betting audience is oppositely confronted with negative profits in expectation, which raises questions about the rationality of sports wagering at all. Even though the punters' participation in such an unfair game contradicts the axiomatic approach of expected utility maximization by Von Neumann and Morgenstern (1953), there are several explanations in the literature such as obtaining pleasure from gambling or adjusted utility frameworks where marginal utility is not universally declining but characterized by local risk preferences. However, for further information about gambling behaviour the reader is referred to the literature (e.g. Sauer (1998)). The focus will now return to the bookmakers' profit maximization problem, assuming the punters' decision to enter the market has already been made.

Besides the neutral pricing strategy, bookmakers may be aware of predictable patterns in betting behaviour by punters, including behaviour based on cognitive or judgemental biases (Dixon and Pope 2004). If on average punters display systematic biases exploited by the bookmakers, the market will contradict the efficiency hypothesis. Levitt (2004) briefly sketches some profit maximizing scenarios within the bookmaker market, leading to a deviation of an efficient market equilibrium. First, bookmakers are skilled in determining in advance the prices σ_e which equalizes the quantity of money wagered on the win and losing outcomes respectively, hence they balance the book. In this case, the bookmaker achieves a positive return equal to the margin charged, regardless of the match outcome realization. Consequentially, bookmakers do not need to be good at forecasting the underlying event outcome but rather good at forecasting bettors' demand or beliefs about

[9] In our analysis, we assume that the commission (overround) is equally distributed over the outcome probabilities, which is in line with the literature (see Forrest et al. (2005), Franck et al. (2010)).

the underlying probability distribution, which implies that in case of bettor irrationality the market implied probabilities may deviate systematically from the true probability distribution. However, in case of skilled bettors with superior forecasting performance, bookmakers face substantial losses if their prices do not include all relevant information and differ widely. Alternatively, bookmakers are not only efficient in predicting game outcomes but also anticipate bettors' preferences and systematically set the wrong prices in a manner that takes advantage of biased irrational behaviour of the betting audience. To sum up, the bookmaker odds are likely to be influenced by both, the true outcome probabilities and the betting audience's expectation and demand (Franck et al. (2010)).

4 Literature Review

There exists a considerable amount of papers dealing with market efficiency in sports betting markets. Detailed overviews are provided e.g. by Williams (2005) and Sauer (1998). In general most of the empirical research supports an efficiency paradigm for betting markets. However, some anomalies have been detected and repeatedly confirmed, though the statistical and economic significance of these inefficiencies have seemed to disappear in more recent years, which suggests an increase in betting market efficiency, potentially evoked by an increase in competition, price transparency and the rise of alternative betting market forms (Vlastakis et al. (2009)).

4.1 Weak-form Efficiency Testing

The first empirical studies examining market efficiency within wagering markets were mostly conducted in racetrack betting. This form differs from the bookmaker football wagering markets with ex ante fixed odds, since horse race betting is characterized by a pari-mutal betting mechanism, i.e. the potential payout is not fixed ex ante but the fraction of the total betting pool you will receive in case of winning. However, these studies revealed some interesting anomalies while providing mixed evidence for market efficiency in general. Several authors have studied the relationship between the objective probabilities of winning a race and the probability reflected by market odds (Quandt (1986)). Ali (1977) introduced the term "*favourite-longshot bias*" describing the well-documented tendency of punters to consistently overbet low probability outcomes (*longshots*) and underbet high probability outcomes (*favourites*) relative to their observed frequency of winning. This anomaly was confirmed also in other sports betting markets and summarized with the phenomena that lower odds (*favourites*) tend to be associated with higher returns and vice versa lower returns with higher odds (*underdogs*), which suggests that market prices anticipate the irrationality in betting demand (see Woodland and Woodland (1994), Snyder (1978) and Cain et al. (2000)). Similar biases were detected in financial markets for equity options by Rubinstein (1985) according to which shorter maturity options (*longshots*) tend to be overpriced. For the *favourite-longshot bias* phenomena there exist a variety of explanations such as risk attitudes in form of risk-loving utility of gamblers (Ali (1977), Quandt (1986)) or insider trading introduced by Shin (1991, 1993). Alternatively, behavioural theories as the Prospect Theory suggest that cognitive errors and misjudgements of probabilities may play a role in market mispricing (Kahneman and Tversky (1979)).

Examining more explicitly the relationship between objective probabilities of outcome occurrence and subjective probabilities implied in quoted odds, Pope and Peel (1989) investigate the efficiency of prices set by four UK high street bookmakers for fixed-odds betting on English association football. A simple test of the weak-form efficiency hypothesis is based on linear probability regression models of match outcomes against implicit bookmakers probabilities. The results indicate some deviations of the implied bookmaker probabilities from the axioms of rational expectations, though exploiting these inefficiencies didn't result into profitable returns. Several other studies were also not able to reject the weak-form efficiency hypothesis (Gandar et al. (1988), Sauer et al. (1988)). However, Golec and Tamarkin (1991) suggest that this simple form of regression tests may not reject market efficiency because of their weak statistical power and it is shown that economic tests in form of testing the profitability of different betting strategies reject the EMH more often. Examining the point-spread NFL betting market[10] by explicitly testing for home as well as favourite biases, Golec and Tamarkin (1991) detect significant deviations from the efficient market hypothesis in these two dimensions implying that bets on underdogs or home teams win more often than bets on favourites or visiting teams, albeit with disappearing magnitude over their sample period. Using a more general model accommodating the symmetries and interdependence of the principal team characteristics (favoured, underdog, home and away), Dare and MacDonald (1996) find little or no evidence against market efficiency in their subsequent analysis. Gray and Gray (1997) apply in a similar way a probit regression model and test for potential momentum effects including recent performance proxies. Their results indicate that the market overreacts to a team's recent performance suggesting that a profitable strategy involves betting on teams that have performed well over the season as a whole, but which have performed poorly in recent weeks. This result has a close analogy with financial markets, in which contrarian strategies are profitable, though the anomalies also seem to have been substantially attenuated in recent years.

Testing for sentiment bias, Avery and Chevalier (1999) examine the dynamics of point spread betting lines during betting periods and control for expert opinions, a hot-hand

[10] This is wagering on a range of match outcomes where the resulting payoff is based on the wager's accuracy.

bias as well as a bias for team prestige. The results from a sample of games from 1983 to 1994, indicate that each set of sentiment variables serves as a significant predictor of point-spread betting line movements but not of the actual game outcomes itself. Forrest and Simmons (2008) examine the weak-form efficiency of the Spanish online football betting market with respect to fan support bias besides the home-away and favourite-underdog dimensions. They use a multivariate analysis where a proxy is included to account for the relative numbers of supporters of the two teams. The odds appear to be influenced by the relative number of fans respectively, with supporters of the more popular team offered more favourable terms on their wagers. Focusing on an economic testing strategy, Vlastakis et al. (2009) assessed the international efficiency of the European football betting market by investigating the profitability of strategies based on combined betting and detection of arbitrage opportunities as well as betting strategies derived from regression models. Despite the increasing competition in the betting industry recently, the results of the empirical analysis in Vlastakis et al. (2009) suggest deviations from the standard weak-form market efficiency assumption. In particular, combined betting across two or three bookmakers simultaneously lead to highly profitable arbitrage opportunities. Subsequently, Franck et al. (2013) examine inter-market arbitrage by combining bookmaker odds and odds from a bet exchange market platform, receiving a guaranteed positive return in 19.2 % of the matches in the top five European soccer leagues. The authors indicate that bookmakers may set prices not only by myopic optimizing over a particular bet but also by taking the future trading behaviour of their customers into account.

4.2 Semi-strong-form Efficiency Testing

In testing the semi-strong form efficiency of financial markets, two main approaches have been adopted. The first one is to assess the direct impact of new public information on prices, the other is searching for opportunities to earn systematic abnormal returns based on publicly available information and identifiable market circumstances, so-called market anomalies (Williams (2005)). Translating this into the betting market setting, the semi-strong efficiency hypothesis implies that the incorporation of publicly available information should not improve the accuracy of outcome predictions based on bookmaker odds and not result in superior betting returns (Kuypers (2000)).

However, searching for market anomalies in betting markets including public information raises the question how we can incorporate any information for match outcome predictions in general. It seems evident that match results in association football are governed partly by chance and partly by skill (Hill (1974)) since teams are not identical and each one has its own inherent quality (Maher (1982)).

There are two main statistical approaches for modelling football match outcomes in literature. The first one models the goal scoring process of the two opponent teams by estimating a statistical model approximating the underlying stochastic process of a football match. Maher (1982) is one of the first papers applying a univariate and bivariate Poisson distribution[11] with means reflecting the attacking and defensive capabilities of the two opponents. Dixon and Coles (1997)) extend the independent Poisson distribution model with means modelled as functions of the respective teams' previous performances and are able to derive forecast predictions for UK football match score results. Consecutively, the probability forecasts of the match outcomes are obtained by aggregating the probabilities of different permutations of the home and away team goals scored. The economic value of these forecasts are assessed in Dixon and Pope (2004) later on. By using their advanced statistical forecasting model, the authors find that the market is inefficient as the model probabilities can be exploited to earn positive abnormal trading rule returns. In contrast to the *favourite-longshot bias* described above, they found less-than-fair odds for favourites and more-than-fair odds for longshots, hence surprisingly a reverse favourite-longshot bias.

The alternative approach to predict match outcomes in the literature involves modelling the match outcomes home win, draw and away win, directly using discrete choice regression models such as ordered probit or logit. Kuypers (2000) uses publicly available information covariates in his ordered probit framework including difference in teams' average points and goals scored in preceding games of the season. With a simple forecast-ratio based betting rule based on the received probability forecasts, he find evidence of market inefficiency within the UK fixed odd bookmaker market. Forrest et al. (2005) test odds for the English football seasons 1998 - 2003 with an extended ordered probit model incorporating a large number of quantifiable variables relevant to match outcomes. The principal team quality indicators were constructed by partitioning the teams' win ratios in components from the present season, the previous season and within 12 and 24 months respectively. Additionally, the model fit was improved by including recent match results, a proxy for the importance of the match for final season outcomes, cup involvement, the geographical distance between the team locations and a measure for the opponents' relative fan attendance. The findings suggest that the bookmakers' forecast accuracy is superior to the model performance while the odds-setters improved over the sample period relative to the statistical model. Nevertheless, evidence of statistically significant differences between the forecast probabilities in three out of five seasons were found and simulation of a selective betting strategy based on the forecast discrepancy translated into significant superior returns. In an alternative approach Hvattum and Arntzen (2010) assigned team ratings based on the ELO rating system, a rating system initially developed for assessing the strength of chess players, and used the rating difference as a single covariate in an ordered logit model specification. Although the model's forecasting performance ex-

[11]This is a discrete probability distribution that expresses the probability of a number of events (here: goal scored) occurring in a fixed interval of time.

ceed the previous approaches in literature, Hvattum and Arntzen also fail to outperform bookmakers forecast accuracy between seasons 2001-2008. The review of the forecasting approaches in the literature is completed by a comparison between the forecasting performance of the goal-based and result-based approach in Goddard (2005). Estimating a set of models based on an identical dataset but different modelling approaches, the difference in performance appeared to be small and no gain was achieved by using the more data and computational intensive approach of the Poisson distribution models.

All in all, the informational efficiency of the bookmaker market seems to be confirmed from a semi-strong form perspective, though several authors are able to simulate profitable betting strategies including their model forecasts.

5 Empirical Analysis

5.1 Data

The data used in the empirical analysis is combined from different sources, whereby the main part is originated from an exclusive dataset provided by Sportradar. The dataset covers all football fixtures of the English *Premier League*, the English second division *Championship* as well as the German *Bundesliga 1* and *Bundesliga 2*. The historical matches are recorded over the last ten years beginning in the second half of season 2006/2007 until February 2017[12] with 15.400 matches in total. Besides the match results, the data lists odds of 14 online bookmakers for each of the theoretical match outcomes $e \in \{H, D, A\}$. Including price changes of online bookmakers before the match starts, we are able to list the last pre-match price offers, the so-called closing odds, and it can therefore be assumed that all relevant pre-match information, including the betting demand by then, is incorporated in the examined bookmakers' pricing decision. However, as can be seen in Table 1, the database is not fully balanced and the sample of odds varies depending upon the individual bookmaker.

The 14 bookmakers in this sample are anonymised by their market focus. The sample contains 3 bookmaker with a primary focus on German speaking betting audience (German 1- 3), 6 leading bookmakers of the British betting market (UK 1- 6) and 5 bookmakers which are characterized by a more global/ European focus and denominated by EU 1- 5.

[12]For convenience the football seasons, which typically run from August to May/June are identified by its end year in the following.

The first part of the data will be extended to enable the analysis of informational efficiency with respect to the weak and semi-strong hypothesis. As already mentioned, examining the semi-strong form efficiency of the bookmaker market requires publicly available information with potential relevance for match result predictions. Since our forecasting strategy is structured with a model estimation based on a training sample including the five preceding seasons, we have to extend the historical match result series by 5 seasons up to season 2001. The information will be extracted from FootballData (2016), a public database of historical match results and bookmaker odds. In addition to the match results data, geographical distance between the opponents and relative attendance is included in the analysis. The geographical coordinates of the team's stadium location are collected from GoogleMaps (2016) and the respective distance between the opponents is calculated with the shortest path between two points on an ellipsoid. The information regarding fan support, e.g. the relative attendance of the competing teams, is collected from EFS (2016), whereby the average stadium attendance of the preceding season is used.

5.2 Descriptive Analysis

In the following, the descriptive statistic of the bookmaker data will be examined. The first panel in Table 1 lists the individual implied bookmaker margins m and their standard deviations. In the previous section it was mentioned that no bookmaker should be able to operate at greater margins than their competitors. Surprisingly, as it can be seen, there is a significant heterogeneity in bookmakers' average implied margins. Especially, the German betting provider German 1 shows a very high margin by 18.50% compared to the average bookmaker margin of 7.94 %. Among the other bookmakers, the takeouts are more comparable, though the global oriented provider EU 1 stands out with a very low average commission by 2.46 % combined with the lowest standard deviation. According to their webpage, the firm intentionally follows a reduced margin pricing model to attract higher turnovers by reducing betting limits to safeguard against high-skilled punters. Despite taking the standard deviations as proxy for the riskiness of the bookmakers' pricing strategy in general, the statistic reveals no clear picture. The highest coefficient of variation[13] is occupied by German 3, a bookmaker with moderate average margin followed by UK 3 and EU 2. Therefore, from the descriptive analysis it is not possible to conclude that higher margins can be seen as a compensation for higher risk taken by the bookmaker.

To examine the margin behaviour in more detail, Figure 2 plots the implied margins by monthly averages based on Bundesliga 1 and 2 matches. First of all, there is a clear picture that margins have decreased over time where most of the bookmakers showed

[13]Coefficient of variation = $\frac{sd}{mean}$.

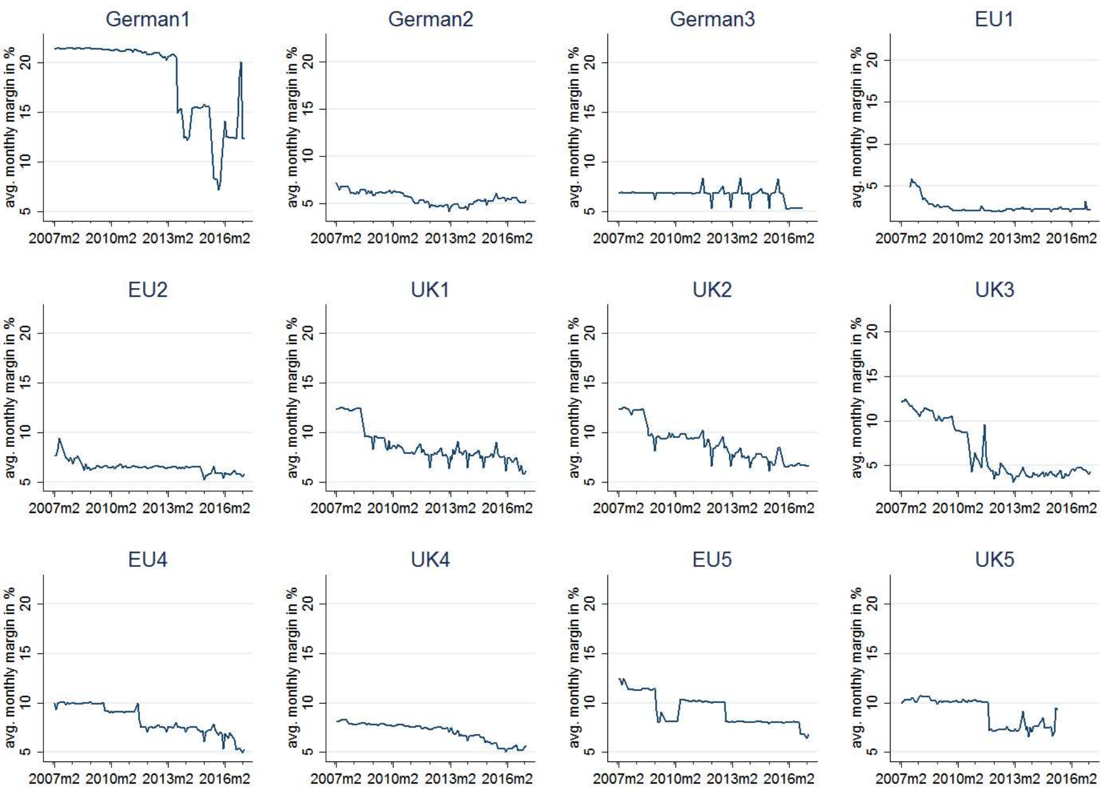

Figure 2: Evolution of bookmaker margins - German market

margins close to or higher than 10 % at the beginning of the observation period in 2007, besides only German 2, German 3, EU 1 and UK 4 with lower margins. Moreover, the plot reveals several structural breaks in the time series of monthly margins. The partial liberalization of the German market seems to affect the bookmaker German 1 in 2014 by a sharp decrease from roughly 21 % to below 15% whereby EU 4 and 5 as well as UK 3 and 5 reveal sharp decreases already at the beginning or mid of 2012, the year of the introduction of the new German interstate gambling treaty. All in all, the picture seems to support decreasing market profits within the bookmaker business, potentially induced by an increase in competition and market liberalization.

The next panel of Table 1 examines the distribution of implied probabilities derived from the market odds as described in section 3 relative to the true probability of the respective outcomes. The chance of the outcome home victory as the final match result is 44.75% and vice versa 26.64% for away victory. This is in line with the so-called home-field advantage in association football (Hill (1974)), which implies that the home team exhibits a significant higher probability of winning than visiting teams. Comparing the bookmakers' average implied probabilities, the market odds coincide with the underlying probability distribution quiet well. The home win implied probabilities P^{IP}_{Home} range from 43.27% to 44.75% comprising the true probability closely. However, comparing the numbers for away victory P^{IP}_{Away} and draw P^{IP}_{Draw}, the market odds tend to overestimate the first one, where P^{IP}_{Away} ranges from 26.23% to 29.73%, while underestimating the chance of a draw result, though the deviations are only marginal. Hence, there is no clear indication that the home-field advantage is consistently overestimated, as frequently mentioned in the literature for example by Vlastakis et al. (2009), but the relative proportions of draw and away results appear not to be correctly represented by the implied bookmaker probabilities on average. Examining the probability distribution with respect to favourite and underdog winning probabilities sheds some light on the magnitude of a potential *favourite-longshot bias* in bookmaker pricing. Teams are defined as favourites if the implied probability of winning exceeds the opponent's odds and for underdogs vice versa. Since there is not always an agreement about the assignment of the favourite role between the bookmakers, the true probabilities are defined if at least 5 betting providers agree on the favourite/underdog status of a team. Comparing the average values of P^{IP}_{Fav} by 48.02% and P^{IP}_{Und} by 23.83%, the bookmakers seem to slightly underestimate the true odds in the favourite-underdog dimension on both sides. This raises the question whether the deviation from the true probability distribution is continuously characterized in the relationship between implied and actual chance of winning or in form of kinks at the boundary of the probability space in an otherwise neutral relationship (Cain et al. (2003)).

	GER 1	GER 2	GER 3	EU 1	EU 2	UK 1	UK 2	Uk 3	EU 3	EU 4	UK 4	EU 5	UK 5	UK 6	True Prob.
$\varnothing\ m$	18.50	6.29	7.22	2.46	5.42	7.54	7.62	5.80	7.67	8.43	6.87	10.28	8.98	8.10	
$SD\ m$	4.40	1.52	1.58	1.01	1.91	2.44	2.53	3.19	2.11	1.87	1.66	2.20	1.69	2.72	
P^{IP}_{Home}	43.27	44.33	44.26	44.75	44.53	43.86	44.25	44.43	44.15	44.17	44.30	43.64	44.54	44.22	44.75
P^{IP}_{Draw}	26.99	26.48	26.47	26.23	26.38	26.96	26.56	26.31	26.60	26.44	26.48	26.74	26.36	26.61	28.62
P^{IP}_{Away}	29.73	29.19	29.27	29.02	29.09	29.18	29.19	29.26	29.25	29.39	29.23	29.62	29.10	29.16	26.64
P^{IP}_{Fav}	47.09	48.48	48.13	49.75	49.09	47.57	48.12	48.79	47.92	48.22	48.52	46.12	48.69	47.83	49.58
P^{IP}_{Und}	24.48	23.55	24.06	23.70	23.79	23.70	23.76	23.91	23.87	23.94	23.92	24.09	24.17	23.58	24.29
N	10698	14927	14798	14571	15305	13360	14792	14272	15281	15348	15110	15372	12288	10921	

Note: Indications are made in percent

Table 1: Descriptive statistics of implied margins and implied outcome probabilities

Examining the relationship between bookmaker's implied probabilities and the underlying actual chance of the outcome more precisely, Figure 3 and 4 in the appendix use scatter plots for grouped bets on their actual chance of winning[14] categorized by their implied probabilities. In each plot the dashed diagonal line represents neutrality, or lack of any

[14]The categories are specified by a bandwidth of 2.5 percentage points for implied probabilities and at least 50 observations are required for each group.

bias, between the mean bookmaker probabilities per group and their actual percentage of winners. The solid line is estimated by a locally weighted polynomial regression (see Fan (1992) and Fan and Gijbels (1996)) of implied probability on actual chance of winning using cubic polynomials, the Epanechnikov kernel and the rule of thumb bandwith (ROT)[15], so that no functional form of the relationship has to be prespecified. The grey areas are defining the 95% confidence intervals, whereby the existence of a conventional favourite-longshot bias would be indicated if the regression line was significantly above the diagonal at low chance outcomes and below when the chances of winning approach unity implying that favourites win more and underdogs less often than their betting prices imply (Cain et al. (2000)). In case of home victory bets the regression line in Figure 4 tracks the diagonal line for the lower and middle part of the probability space closely with only vague deviations. Narrowing the focus to bets with implied probabilities between 10% and 40%, where a main part of the home victory bets are located, a slight local tendency to underestimate the true odds is revealed for EU 1, UK 1 and German 3. At the upper boundary of the probability space the implied probabilities seem to underestimate the true odds as well, particularly for German 1, EU 4 and EU 5, supporting the evidence from Table 1 of underestimated favourite as well as underdog probabilities in general. Figure 5 plots the relationship for away victory bets, whereby interestingly the lower boundary lacks of any deviations from the diagonal except a slight heightened tendency in case of German 1, EU 4 and EU 5. The biases at the upper boundary of the away victory probability space is hard to examine due to very few observations in the dataset, in which a heterogeneous picture EU 1 overestimating high probable away victories while German 1 and EU 5 display a reverse tendency similar to Figure 3. All in all, the analysis of the relationship between the implied probabilities and the actual chance of winning revealed no clear bias in line with the so-called *favourite-longshot* or *home-field bias* but some indications for slight deviations at high and low probable outcomes which are differently characterized for home and away win bets. In the following, we will assess whether these slight deviations from neutrality in bookmaker prices may be exploited for superior expected returns.

Trading rules like always betting on the favourite, underdog, home or away team should yield in expectation the same return according to the efficient market hypothesis. Table 2 states the average returns for these simple trading rules. The first line lists the average returns a punter would have achieved by betting randomly on an individual bookmaker's odds in the sample. Obviously, the average returns should reflect the implied margins from Table 1, hence we get an heterogeneous picture as well, whereby the returns are close to the negative average overround estimate m in Table 1. The next two rows display the

[15]Weighted least square regression using at each point of interest x the Epanechnikov kernel as weighting function for points within the bandwidth neighbourhood of x. The results are robust to other kernel, bandwith or polynomial specifications.

	GER 1	GER 2	GER 3	EU 1	EU 2	UK 1	UK 2	Uk 3	EU 3	EU 4	UK 4	EU 5	UK 5	UK 6
∅Return	**-17.25**	**-6.51**	**-7.40**	**-2.35**	**-5.48**	**-7.79**	**-7.99**	**-5.80**	**-7.93**	**-8.55**	**-6.91**	**-10.84**	**-8.94**	**-8.17**
∅Ret. Home	**-13.29**	**-4.96**	**-6.15**	**-2.11**	**-4.57**	**-5.45**	**-5.96**	**-4.57**	**-5.83**	**-6.60**	**-5.39**	**-7.88**	**-7.22**	**-5.38**
∅Ret. Draw	**-18.80**	**-6.50**	**-6.28**	**-1.07**	**-4.85**	**-8.21**	**-7.52**	**-4.25**	**-7.70**	**-7.66**	**-6.53**	**-10.35**	**-7.71**	**-8.11**
∅Ret. Away	**-19.66**	**-8.08**	**-9.78**	**-3.86**	**-7.01**	**-9.70**	**-10.49**	**-8.59**	**-10.27**	**-11.41**	**-8.81**	**-14.30**	**-11.90**	**-11.01**
# Bets	10698	14927	14798	14571	15305	13360	14792	14272	15281	15348	15110	15372	12288	10921
∅Ret. Fav	**-11.89**	**-5.44**	**-6.32**	**-3.61**	**-5.46**	**-5.77**	**-5.64**	**-5.57**	**-6.05**	**-7.32**	**-5.94**	**-6.48**	**-7.50**	**-5.87**
∅Ret. Und	**20.59**	**-7.34**	**-9.53**	**-2.29**	**-5.98**	**-9.08**	**-10.46**	**-7.37**	**-9.78**	**-10.53**	**-8.03**	**-15.05**	**-11.35**	**-10.20**
# Bets	10482	14613	14522	14506	15147	13024	14468	14076	14938	15045	14881	14717	12154	10620
∅Home.Fav	**-9.07**	**-4.04**	**-4.69**	**-2.47**	**-4.04**	**-4.23**	**-4.40**	**-3.82**	**-4.59**	**-5.30**	**-4.35**	**-5.07**	**-5.55**	**-4.25**
# Bets	7859	10952	11039	10771	11311	9816	10978	10517	11304	11316	11195	11370	9231	8058
∅Home.Und	**-4.20**	**-0.73**	**-1.38**	**0.43**	**-0.72**	**-1.19**	**-1.38**	**-0.62**	**-1.24**	**-1.25**	**-0.86**	**-2.53**	**-1.75**	**-1.03**
# Bets	2623	3661	3483	3735	3836	3208	3490	3559	3634	3729	3686	3347	2923	2562
∅Away.Fav	**-2.82**	**-1.40**	**-1.63**	**-1.14**	**-1.42**	**-1.54**	**-1.24**	**-1.74**	**-1.46**	**-2.02**	**-1.59**	**-1.41**	**-1.96**	**-1.62**
# Bets	2623	3661	3483	3735	3836	3208	3490	3559	3634	3729	3686	3347	2923	2562
∅Away.Und	**-16.38**	**-6.61**	**-8.15**	**-2.72**	**-5.27**	**-7.89**	**-9.08**	**-6.75**	**-8.53**	**-9.28**	**-7.17**	**-12.52**	**-9.61**	**-9.18**
# Bets	7859	10952	11039	10771	11311	9816	10978	10517	11304	11316	11195	11370	9231	8058

Note: Indications are made in percent

Table 2: Average returns for various simple betting rules

betting returns for betting exclusively on home, draw or away win outcomes respectively. Contrarily to the findings in the bookmaker's implied probabilities, the strategy selecting only home team winning outcomes performs better than choosing one of the other outcomes. From this ex-post return analysis, it seems that on average bookmakers underestimate the occurrence of a home victory while overestimating favourable outcomes for the visiting team. However, assessing betting strategies selectively on favourite or underdog status results in an inconclusive picture in the next panel. Predominantly, randomly betting can be slightly outperformed by betting exclusively on favourites while selecting underdogs results in an inferior ex post return. Only in case of the low margin bookmaker EU 1O there is a slight superior return by betting on underdogs. However, since home teams have a higher probability of winning due to the home-field advantage in football, home teams tend to be characterized as favourites more often. Therefore, a potential home-field and favourite-longshot bias may be interconnected (Vlastakis et al. (2009)) and is examined separately in the next panel. A different picture is now revealed. The selective betting strategies on home teams with underdog status performed significantly better than random betting or any of the other strategies. Noteworthy is that for bookmaker EU 1 who has already been characterized by very low margins the ex-post return changes to positive. Placing reverse bets on the same sample of matches, e.g. bets on favoured visiting teams, sharply improves the average returns compared to randomly betting as well and constitutes the best strategy for costumers of the two bookmakers German 1 and Eu 5, which are characterized by substantially superior favourite returns as mentioned above. In conclusion, there is some evidence of biased odds on both winning outcomes for games with a strong visiting team, which yields a slight positive average return in case of EU 1. Since both winning outcome occurrences seem to be underestimated for this selection of football matches, the implied proportion of draw results is upward biased here in contrast to the general conclusion from the analysis in Table 1.

5.3 Weak-form Efficiency Analysis

5.3.1 Methodology

The following section will outline the methodology used for testing the informational efficiency. In order to test the weak-form efficiency hypothesis, we will apply a rational expectation framework and assess the predictive quality of the implied match outcome probabilities derived from the individual bookmaker odds. In the literature the following general model is estimated by regressing the respective match outcome Y_{ie}[16] on implied probability $P^{ip}_{ie,j}$ derived from bookmaker j's market price $\sigma_{e,j}$ as described in section 3:

$$Y_{ie} = G\left(\beta_{0j} + \beta_{1j} P^{IP}_{ie,j} + \epsilon_{ie,j}\right) \tag{5}$$

where G(.) is assumed to be first a linear function (*linear probability model*) and in a second specification a standard normal cumulative distribution function (*probit model*). Note that the general model is assumed to be a pooled model including all available bookmaker prices in the dataset and presupposes no heterogeneity over time in $P^{IP}_{ie,j}$. For the linear probability model, a test for market efficiency would then be given by the joint null hypothesis stating that $P^{IP}_{ie,j}$ is an unbiased predictor for the match outcome e, hence no constant effect can be added to the bookmaker forecast that is characterized by a one-to one relationship to the underlying actual probability:

$$H0: \qquad \beta_{0j} = 0 \quad and \quad \beta_{1j} = 1$$

Considering the fact that outcomes e (home, win, draw) are not independent from each other but rather mutually exclusive as only one result can occur for each match i, estimating the equation for draw, home and away win outcomes independently fails to refer to the condition that the dependent variables have to sum up to 1 and for the interdependency between the outcomes e and e' resulting in a non-zero disturbance covariance $Cov(\epsilon^{e}_{ij}, \epsilon^{e'}_{ij})$.

To account for that, I will use a seemingly unrelated regression framework (SUR) introduced by Zellner (1962) and estimate the three match result probabilities per match as a system of equations employing the condition: $P^{IP}_{i,e=D,j} = 1 - P^{IP}_{i,e=H,j} - P^{IP}_{i,e=A,j}$[17]. The SUR framework uses a feasible generalized least square approach which estimates the

[16] $Y_{ie} = 1$ if match outcome e occurred, $Y_{ie} = 0$ else.
[17] This specification was inspired by Forrest and Simmons (2005).

coefficients by permitting non-zero covariances between the error terms ϵ_{ij}^{e} and $\epsilon_{ij}^{e'}$ while assuming strict exogeneity of the covariates and homoscedasticity[18]:

$$\begin{aligned}
Y_{e=H,i} = &\quad \beta_{0j} + \beta_{1j} P^{IP}_{i,e=H,j} + \epsilon^{H}_{ij} \\
Y_{e=A,i} = &\quad \alpha_{0j} + \alpha_{1j} P^{IP}_{i,e=A,j} + \epsilon^{A}_{ij} \\
Y_{e=D,i} = &\quad \gamma_{0j} + \gamma_{1j}(1 - P^{IP}_{i,e=A,j} - P^{IP}_{i,e=H,j}) + \epsilon^{D}_{ij} \\
= &\quad \gamma'_{0j} + \gamma'_{1j}(P^{IP}_{i,e=A,j} + P^{ip}_{i,e=H,j}) + \epsilon^{D}_{iDj}
\end{aligned} \tag{6}$$

Estimating model (6) with the SUR estimator is more efficient compared to a single equation ordinary least square (OLS) estimation, the greater the correlation of the disturbances is and the less correlation between the individual equation regressors can be assumed (Zellner (1962)). Since the regressors used in each equation vary and it can be expected that the disturbances across equation are significantly interconnected, this approach seems justified. The following joint null hypothesis of informational efficiency will then be tested for each bookmaker j:

$$H0: \quad \begin{aligned} &\beta_{0j} = 0, \beta_{1j} = 1 \\ &\alpha_{0j} = 0, \alpha_{1j} = 1 \\ &\gamma'_{0j} = 1, \gamma'_{1j} = -1 \end{aligned}$$

applying a test statistic analogous to the F-ratio in multiple regression analysis:

$$\hat{F} = \frac{1}{6}(R\hat{\hat{\beta}} - q)'[R\hat{Var}[\hat{\hat{\beta}}]R']^{-1}(R\hat{\hat{\beta}} - q) \sim F[6, n]$$

The statistic is only valid approximately similar to the Wald statistic, but tends to perform better in small or moderately sized samples (see Greene (2012)). In case of rejection of the null hypothesis and given that the marginal effects of the implied probabilities are significantly larger or smaller than unity, we can argue in favour of a bias leading to misperception of winning probabilities for the favourite or underdog team. However, since the SUR model assumes a homoscedastic error structure, it only allows for different covariances between equations and not within equation, the standard error estimates are inconsistent due to the heteroscedastic nature of linear probability models.[19] Therefore, as robustness test an alternative approach for model (6) will be applied using less-efficient single-equation OLS coefficient estimates but consistent heteroscedasticity as well as inter-equation dependency robust standard errors of sandwich/robust type[20].

Examining the pooled model assumption thereafter, model (6) will be estimated sea-

[18] The SUR model estimation procedure is described in the appendix.

[19] Since outcome variable is binary, it follows that $Var(\epsilon|x) = P(Y=1|x)[1 - P(Y=1|x)] = x'\beta\left(1 - x'\beta\right)$, hence depends on x, see Long (1997).

[20] The standard errors are obtained by the *suest* command in Stata.

sonally to be able to observe potential changes in the bookmaker market efficiency over time.

A further robustness check addresses the potential shortcomings of a linear probability model which is not constrained to estimate fitted values within the [0,1] range of the probability space and does not ensure consistent estimates of a non-linear relationship. This alternative approach is based on Forrest and Simmons (2008), who develop a multivariate analysis for the purpose of identifying separately home vs. away, longshot as well as sentiment bias extending model (5) by using a probit model. Hereby, the observation units are only bets that either the home or away team will win, excluding bets on draw results. The sample size doubles and each match observation i is represented twice from both team perspectives and combined with the forecast based on the respective $Bet_{i,e=(A,H)}$. The following single equation probit model is then estimated[21]:

$$Prob\left(Bet_{i,e=(A,H)} \quad wins\right) \quad = \Phi\left(\beta_{0j} + \beta_{1j} Prob^{ip}_{ie} + \beta_{2j} Home_e + \beta_{3j} \Delta Attend_i\right) \tag{7}$$

where the additional dummy variable $Home_e = 1$ if the bet e is conditional on home teams' victory identifying a potential home team bias. The second additional covariate $\Delta Attend_i$ serves as team support proxy and is defined as the difference between teams' preseason average home attendance. As Forrest and Simmons (2008) note, the number of active fans, e.g. people who are buying stadium tickets, is likely to be correlated with the number of passive fans, thus potential sentiment bettors. The null hypothesis of efficiency claims here as well that the odds quoted by the bookmaker reflect all available information relevant to the match outcome, hence the coefficients on $Home_e$ and $\Delta Attend_i$ should be zero and the marginal effect of the implied probabilities equal 1. Since the datasets consists of pairs of bets, home win and away win bet for each match, the error terms are correlated per match observation. Therefore, a clustered probit model with match observation cluster is estimated.

5.3.2 Results

The estimation results of the pooled model (6) based on the SUR methodology are displayed in Table 5 for each bookmaker individually. The SUR framework employs a FGLS approach incorporating the inter-equation dependency in this setting and is justified by a Breusch-Pagan test rejecting the null hypothesis of equation independence at all reasonable significance levels. Overall, the coefficients appear to be predominantly close to the null hypothesis of informational efficiency which states that the marginal coefficients of bookmaker forecasts are equal 1, or -1 in case of draw outcome predictions, and the absence of a constant bias captured by the intercept. From the 14 bookmakers within

[21] For $G()$ the standard normal cumulative distribution function $\Phi()$ is used.

the sample, the hypothesis is nevertheless rejected at a 5% significance level for German 1 and EU 5 as can be seen from the p-values in the last row of Table 5. This states that the test-statistic, characterized by an F(6,n)-distribution under H0, is significantly larger than the respective critical values. When examining the result with respect to general forecast unbiasedness represented by the constant terms, only the above mentioned providers reveal a significant negative coefficient for home and away victory bets accompanied by German 3, UK 1, 2 and 5 as well as EU 3 and 4 exclusively for the latter, indicating a systematic underestimation of the true probability despite a small magnitude.

Focusing on the size of the forecast coefficients, marginal effects significantly larger than unity imply that the actual winning probability increases disproportionately with the implied bookmaker probability and below unity vice versa. Hence, if we are able to detect coefficients significantly larger than one, we can state that the odds tend to underestimate high-probability outcomes (favourites) and overestimate low-probability outcomes (longshots) on average in line with the favourite-longshot bias phenomena (Franck et al. (2010)). First of all, surprisingly the forecast coefficients are identical in all 3 outcome equations similar to the results in Forrest and Simmons (2005) but counterintuitive to the results from the descriptive analysis. Even though the majority of bookmakers reveal coefficients larger than one, this individual difference is only significant in case of the betting provider German 1 and EU 5, the bookmakers already mentioned above in the examination of the constant terms. However, as noted in the methodology introduction, the seemingly unrelated regression model leads to valid inference only in case of an homoscedastic error structure which is not the case in linear probability models[22].

Therefore, Table 6 contains the results of an alternative specification using OLS estimation procedure for each equation independently combined with heteroscedasticity and inter-equation dependency robust standard error estimates. Here, we apply a Wald statistic characterized by a chi-squared distribution with 6 degrees of freedom when the null hypothesis is true[23]. The results for the first two equations are similar to the results of the SUR specification combined with only slightly larger standard errors. The null-hypothesis is now accepted only for 6 out of 14 bookmakers at a 5% significance level and a significant intercept for $HomeV$ and $AwayV$ is revealed again for German 1 and EU 5. While the system of equations approach in the first specification leads to identical coefficient estimates for the individual match outcome equations, estimating each equation independently now leads to substantially more variation in the estimates per bookmaker across match outcome. The strongest effect can be observed on coefficients of bookmakers' draw result forecasts, which results in substantial deviations from unity up to a coefficient estimate of

[22] This is, when the Breusch-Pagan Lagrange multiplier test for overall system heteroscedasticity is rejecting H0 for all bookmakers.

[23] Since the number of observations n is large, it follows that $F(6, n) \approx 1/JChi2(6)$ -> decision based on F-statistic nearly identical as based on Wald-statistic.

-1.423 for EU 5 combined with significant deviations from the in H0 assumed magnitude of the intercept. This is in line with the analysis of betting returns in section 5.2, where the probability of draw outcomes was downward biased and yielded significant superior returns for exclusively betting on ties. However, estimating the equation independently needs to be evaluated with the caveat of lower estimation efficiency.

All in all, the pooled model reveals mixed evidence for slight deviations from neutrality in bookmaker pricing, whereby the majority of betting suppliers displayed coefficients slightly higher than unity in the pooled model, which implies a tendency that the actual winning probability increases disproportionately with the implied bookmaker probability. By using the OLS approach and estimating the equation for the disjunct match outcomes independently, some deviations from neutrality were revealed for draw bets, which indicates that most bookmakers systematically underestimate draw results in their odds pricing.

Relaxing the assumption of no heterogeneity over time in bookmaker pricing, Table 7 lists the test-statistic values for the null hypothesis based on seasonal estimation of the above mentioned specification of model (6). Hence, the estimation sample is partitioned in yearly subsamples containing on average one-tenth of the observations respectively. Keeping in mind that the power of a test, e.g. the probability that it will correctly reject a false null hypothesis, decreases when the sample size is reduced (Greene (2012)), the scarcity of Chi^2-statistic values significantly larger than their respective critical values are nevertheless surprising. The null-hypothesis is now only consistently rejected for the subsample of the second half of season 2007 while for most bookmakers the test-statistics afterwards are predominantly of small magnitude. Only for the German betting provider German 1 the null hypothesis is rejected at a 5% significance level in seasons 2008-10 and 2013-2014, for provider EU 5 in seasons 2008, 2009 and 2014. Performing the alternative SUR specification seasonally yields similar results, even though the null hypothesis is rejected less often similar to the pooled model. The evidence for market inefficiency seems to be clustered, hence it cannot be concluded that market efficiency has improved over time in general but appears to fluctuate over the sample periods.

In order to complete the weak-form efficiency analysis, the results of the clustered probit model defined in equation (7) will be discussed now. The model specification includes a dummy variable for bets conditional on home victory $Home$ and a proxy for relative fan support $\Delta Attend$. The first row in Table 8 lists the marginal effects of the individual bookmaker forecasts. Allowing now for a non-linear relationship between the bookmaker forecasts and the actual underlying probability leads to coefficients primarily larger than unity except for EU 1. However, the deviations from the null hypothesis are negligible, as they range from 1.167 for German 1 to 0.989 in case of EU 1. In the next row, the coefficients for the dummy variable $Home$ are insignificant for all bookmakers excluding

UK 1 and UK 6 at 5% level. Hence we cannot unambiguously reject the null hypothesis of neutrality in the home-away dimension. Surprisingly, also the last row contains only small-sized coefficient estimates, which are only significant at 5% level in case of predominantly UK and EU oriented bookmakers.[24] Contraryly to the results in Forrest and Simmons (2008) and Franck et al. (2011), who found that posted odds of bets on popular teams are underestimated and vice versa for less popular teams, it is thus not possible to conclude that the pricing decisions and the implied bookmaker forecasts are affected by the relative fan base of the opponents. All in all, the evidence of inefficiencies with respect to the weak-form efficiency hypothesis is vague and differently characterised.

5.4 Semi-strong Efficiency Analysis

5.4.1 Methodology

The main approach to test betting market efficiency with respect to public information besides historical prices will be based on a discrete ordered choice model. The model, which is briefly outlined in the following, is mainly inspired by the model of Forrest et al. (2005) and Goddard and Asimakopoulos (2004) and will produce match result forecasts based solely on historical information that is publicly available before the start of each match. Therefore, comparing the forecast accuracy of the statistical model with the accuracy given by the individual bookmaker forecasts, separately for the English and German divisions, will provide a fair test of the market's informational efficiency. Furthermore, the estimated probability forecasts will be used later on to derive betting rules in order to assess the contribution of our fundamental analysis and its effect on ex-post returns. Since the semi-strong form efficiency hypothesis claims the absence of any sort of public information significantly affecting returns besides the market price, this constitutes a straightforward testing strategy.

As already outlined in the literature review, there are two main approaches to produce match result forecasts based on publicly available information in order to test the semi-strong efficiency hypothesis. While one strand uses computationally intensive statistical forecasting models based on the Poisson distribution to approximate the teams' underlying goal scoring processes, the ordered choice model approach estimates the match outcome directly by incorporating covariates with potential effects on the probability of final match result occurrence. A major advantage of this approach is, besides its computational simplicity, to avoid assumptions about the interdependence between the home and away team scoring process (Goddard and Asimakopoulos (2004)) and to enable the inclusion of different kinds of explanatory variables. The underlying ordered choice approach is justified by the ordinal nature of the discrete match result outcomes home win

[24]Alternative specifications with different support proxies for $\Delta Attend$ using normalized attendance figures and different subsamples yield similar results.

(1), draw (0.5) and away win (0), whereby a victory by either side is more connected to a draw outcome than to the reverse victory simply by the underlying goal difference (Constantinou and Fenton (2012)). Potential alternative strategies would have been based on individual binary forecasting models which estimate each of the three potential match results separately or an unordered choice model[25]. However, using independent forecasting models, we should expect a significant loss in estimation efficiency, while the multinomial unordered choice approach lacks the incorporation of the outcomes' ordinal nature. Hence, the application of an ordered probit model in the following seems reasonable.

The outcome of the underlying football match event between home team i and opponent j is now denoted by Y_{ij} and is assumed to depend on the unobserved continuous variable Y^*_{ij} that is subject to an i.i.d normally distributed random term ϵ_{ij}[26] representing the unsystematic or random element in the match results data generating process (DGP). The latent variable Y^*_{ij} can be interpreted as the home team's true capability of winning the match and hence cannot be observed directly but is assumed to depend linearly on n observable covariates x as stated in the following matrix notation.

$$Y^* = x'\beta + \epsilon \tag{8}$$

The following covariates x were tested in different model specifications, whereby the final forecasting specification is obtained by the stepwise model selection algorithm using the Akaike Information Criterion (AIC) introduced by Akaike (1974):[27]

- **Past Performance:** Score Results $\{1 = win, 0.5 = draw, 0 = loss\}$
  - X_i/X_j = Team i/j's past results moving average (last 30 games)
  - $X2_i/X2_j$ = Team i/j's past results moving average (last 20 games)
  - $Xsep_i$ = Team i's past home results moving average (last 20 games)
  - $Xsep_j$ = Team j's past away results moving average (last 20 games)

- **Past Goals Scored:**
  - Y_i/Y_j = Team i/j's moving average (last 30 games) of past goals scored
  - $Y2_i/Y2_j$ = Team i/j's moving average (last 20 games) of past goals scored
  - $Ysep_i$ = Team i's moving average (last 20 games) of past home goals scored
  - $Ysep_j$ = Team j's moving average (last 20 games) of past away goals scored

[25] Typical models for unordered choices are the multinomial logit or probit model, see Greene (2012).

[26] As robustness check, a logistic framework with a logistic distribution of the error term was estimated, but this produced similar results.

[27] The final forecasting model specification includes those variables that minimize the information loss given the selection of variables determined with *stepAIC* in R. The selection criteria is only valid on the same dataset, therefore all match observations with missing covariates have to be removed.

- **Recent Match Results:** $R1_{j,i}, R2_{j,i}$ = Teams' result of nth (1,2) last match
- **Last Result:** $Last_{ij}$ = Home team i's last result against current opponent j
- **Dummy Promotion:** $Prom_i/Prom_j$ = 1 if team i/j was promoted within the last year
- **Dummy Relegation:** Rel_i/Rel_j = 1 if team i/j was relegated within the last year
- **Relative Attendance:** $Attend_{ij} = \frac{\varnothing Attendance^i_{s-1} - \varnothing Attendance^j_{s-1}}{\varnothing Division.Attendance_{s-1}}$ of preceding season $s-1$
- **Log Distance:** $LnDist_{ij}$ = Logarithm of distance in km between teams' stadium locations

The first section of covariates captures the past performance of home team i and away team j respectively by using moving averages of each team's past results excluding the current match. It can be expected that the home team performance proxy has a positive effect on the home team's probability of winning while in case of the opponent's performance an adverse effect on home win occurrence should prevail. Experimenting with different time windows resulted in the above listed selection including the last 30 and 20 games prior to the match of interest. Alternatively, we will estimate a model specification which separates the past performance by including only past away results for the away team $Xsep_j$ and home results for the domestic team $Xsep_i$ to record team strength dependency on home and away grounds more precisely.

Additionally, to include more information about the team's offensive qualities, the next section of covariates includes moving averages of goals scored by team i and j. Even though these covariates are potentially closely correlated with the performance variables from above, but they nevertheless may contain information about the teams' offensive qualities not captured by historical score results completely. The window and separate specification choices are analogous to the result based variables. Following Forrest et al. (2005), the subsequent variables $R1_{j,i}$,$R2_{j,i}$ address the possibility of short-term persistence in match results by including the last and second last result individually for each team. For example Dobson and Goddard (2000) found in their analysis that sequences of consecutive results are subject to statistically significant negative persistence effects, which contradicts the widespread belief that a sequence of positive results improves confidence and moral. In the context of this discussion, the media also often refers to the performance in the teams' last meeting, which is included by $Last_{ij}$, the last match result against the visiting team from the domestic team perspective. The following covariates $Prom_i$ and Rel_i are capturing the home team's inter-division dynamic from past to current season and analogous for the visiting team j, whereby the dummy is $Prom_i = 1$ if the team experienced a promotion and in case of relegation $Rel_i = 1$. Besides the difference in financial strength and quality of the squad, it can be assumed that match outcomes are likely to be affected by heterogeneity in team incentives that are potentially correlated with inter-division dynamics.

The last section of covariates contains information not directly linked to the team performance. The variable $Attend_{ij}$ serves as proxy for the relative fan support and is based on the difference in mean home attendance in the previous season between the home team and its opponent. Hereby, the variable may reflect the psychological influence of the crowd on the match outcome as well as some sort of material effects resulting from the ability of teams with larger fan base to spend more on their budget for new talents and players (Goddard and Asimakopoulos (2004)). To capture only the relative attendance differences within the division in which the match of interest takes place, the metric is normalized by average overall division attendance. Last but not least, the covariate $LnDist_{ij}$ controls for the dependency of the so-called home advantage on the geographical distance between the home towns of the teams participating in the match as detected by Clarke and Norman (1995). $LnDist_{ij}$ measures the natural logarithm of the shortest distance between the teams' stadium locations and may reflect the effect of increasing intensity in competition in local derbies (Goddard (2005)).[28]

The description of explanatory variables is completed by mentioning omitted variables in the analysis with potential effects on outcome probabilities. Variables accounting for player injuries, staff rotation or other information about the team constitutions are excluded due to the fact that constructing them requires accurate and time precise data which is not easily available. Furthermore, Forrest et al. (2005) include a variable highlighting matches which are characterized as important with regard to championship or relegation issues for one team but unimportant for the other. Even though their inclusion may provide improvements to the forecasts, their construction and incorporation is beyond the scope of this study.

The basic model structure will be explained in the following. Even though we do not observe Y^*, we do observe the outcome of the game and assume that the continuous variable Y^* can be linked to the outcome with some threshold parameter μ_{1s}, μ_{2s} in the following ordered probit model framework (Greene (2012), Forrest et al. (2005)):

$$
\begin{aligned}
&AwayWin: Y_{ij} = 0 \quad && if \quad && Y_{ij}^* + \epsilon_{ij} \leq \mu_{1s} \\
&Draw: \quad Y_{ij} = 0.5 \quad && if \quad && \mu_{1s} \leq Y_{ij}^* + \epsilon_{ij} \leq \mu_{2s} \\
&HomeWin: Y_{ij} = 1 \quad && if \quad && \mu_{2s} \leq Y_{ij}^* + \epsilon_{ij}
\end{aligned}
\tag{9}
$$

where Y_{ij}^* depends on the covariates x as stated above. The parameters μ_{1s} and μ_{2s} are the cut off values (percentiles) for the normal distribution controlling for the overall proportions of home wins, draws and away win results in the estimation sample s and

[28] An alternative specification using linear instead of logarithmic distance performed poorly. Hence, a non-linear relationship between distance and outcome probability seems more realistic.

therefore have to be estimated as well. The parameters are obtained by maximum likelihood estimation. The target log-likelihood function is given by the sum of logarithmized probabilities for each of the realized outcomes $y = (H(1), D(0.5), A(0))$:

$$LnL = \sum_{i,y=0} ln[\Phi\left(\mu_{1s} - x_i^{'}\beta\right)] + \sum_{i,y=0.5} ln[\Phi\left(\mu_{2s} - x_i^{'}\beta\right) - \Phi\left(\mu_{1s} - x_i^{'}\beta\right)] + \sum_{i,y=1} ln[1 - \Phi\left(\mu_{2s} - x_i^{'}\beta\right)] \tag{10}$$

whereby the normal distributed disturbance term ϵ is normalized without loss of generality by mean zero and variance one so that the standard normal cumulative distribution function Φ can be applied (Greene (2012)).

Estimating model (8) with a fixed estimation sample, we can derive out-of-sample fitted match result probabilities by the following rearrangement where $\hat{Y}_{ij}^{*}$ is the fitted value of the latent variable derived from the covariates x and the coefficient estimates $\hat{\beta}$ as well as the estimated cut-off values $\hat{\mu}_{1s}$, $\hat{\mu}_{2s}$:

$$\begin{aligned}
&Prob.\ AwayWin: \hat{P}_{ij}^{A} = Prob(\epsilon_{ij} \leq \hat{\mu}_{1s} - \hat{Y}_{ij}^{*}) = \Phi\left(\hat{\mu}_{1s} - \hat{Y}_{ij}^{*}\right) \\
&Prob.\ Draw: \quad \hat{P}_{ij}^{D} = Prob\left(\hat{\mu}_{1s} - \hat{Y}_{ij}^{*} \leq \epsilon_{ij} \leq \hat{\mu}_{2s} - \hat{Y}_{ij}^{*}\right) = \Phi\left(\hat{\mu}_{2s} - \hat{Y}_{ij}^{*}\right) - \Phi\left(\hat{\mu}_{1s} - \hat{Y}_{ij}^{*}\right) \\
&Prob.\ HomeWin: \hat{P}_{ij}^{H} = Prob\left(\epsilon_{ij} \geq \hat{\mu}_{2s} - \hat{Y}_{ij}^{*}\right) = 1 - \Phi\left(\hat{\mu}_{2s} - \hat{Y}_{ij}^{*}\right)
\end{aligned} \tag{11}$$

Finally, we are able to derive match result forecasts for the seasons 2006/2007 until 2015/2016 of the bookmaker sample. The forecasting procedure is structured as follows: At first the ordered probit model coefficients are consecutively estimated for each of the 10 seasons using the preceding 5 seasons respectively. Hence, the forecasting model for the first season 2007 of the forecasting sample is estimated based on a training sample including matches from season 2002 until 2006. Afterwards, the estimated models are used for out-of-sample predictions one season ahead, thus in the above example for all matches in season 2007. Consequentially, the coefficient estimates are adjusted after each season by ensuring a homogeneous sample size of preceding seasons, so that the statistical model is based exclusively on information that would have been available before each match of interest. Forrest et al. (2005) mention in their analysis that extending the estimation period up to 15 seasons would yield some benefits in terms of forecasting accuracy, but due to lack of data, we are restricted to 5 estimation periods in this analysis.

However, one main assumption of ordered choice models has to be mentioned and discussed further, as it is frequently violated (Williams (2009)). The model framework is based on a parallel regression or so-called *parallel odds assumption* claiming that the coefficients do not vary with the outcome category, hence the model has a common slope vector β (Long (1997)):

$$\frac{\partial P\left(Y \leq 0|x\right)}{\partial x} = \frac{\partial P\left(Y \leq 0.5|x\right)}{\partial x} = \frac{\partial P\left(Y \leq 1|x\right)}{\partial x} \tag{12}$$

If this assumption does not hold, the estimated model is likely to be inconsistent (Williams (2009)) and is therefore tested in the following by likelihood ratio tests allowing one or more regression parameters to vary with the ordinal categories e[29]. The parallel odds assumption was not rejected for any parameter at all usual significance levels. Hence, the parallel regression assumption is confirmed for the model setting.

Methods for Comparison

After estimating the match result probabilities, two approaches will be used to assess the relative efficiency of bookmaker and model forecasts. First, the forecast accuracy will be compared by statistical scoring measures for probability forecasts based on how close the probabilities are to the actual observed outcome. For comparison, we will make use of two different kinds of statistical scoring rules. First, the Brier score (Brier (1950)), which is defined as the mean squared difference between the actual outcome realization and predictions defined separately for each match result e (home, win, draw) :

$$BRS^{e} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_{ei} - P_{e}^{IP}\right)^{2} \tag{13}$$

where Y_e is the actual realization of outcome e and P_e^{IP} the implied forecasts by bookmakers or the benchmark model. The Brier Scores are averaged over N seasonal match observation, whereby the smaller the average squared distance appears, the better the forecast performs. Using this measure enables to examine the performance individually for the three match outcomes. However, Constantinou and Fenton (2012) note that the Brier score does not address the ordinality of football match results properly by scoring the individual forecasts separately. Therefore, they refer to a second measure called Rank Probability Score (RPS), developed by Epstein (1969), as a more appropriate measure for the evaluation of football forecasts, as it accounts for the ordinality by using the difference between the cumulative distributions of outcome forecasts $\sum_{j=1}^{i} P_{ej}^{IP}$ and their respective realizations $\sum_{j=1}^{i} Y_{ej}$ per match instance. Equation (14) states the RPS for a single match observation where r is the number of potential outcomes, i.e. $r = 3$ possible outcomes per match[30]:

$$RPS = \frac{1}{r-1}\sum_{i=1}^{r}\left(\sum_{j=1}^{i} Y_{ej} - \sum_{j=1}^{i} P_{ej}^{IP}\right)^{2} \tag{14}$$

[29] The test is implemented with command *nominal_test* in the *ordinal* package in R by Christensen (2015).

[30] Analogue to the Brier score, the measure will be averaged over N seasonal match observations and a lower average RPS indicates higher forecast accuracy.

Second, the information contained in the statistical forecasts will be examined relatively to the information contained in the bookmaker forecasts by including single explanatory variables for positive score results in an ordered probit regression[31]. Precisely, the model and bookmaker forecasts are compared with their maximized log-likelihood values of the regression models of match results on the implied result forecasts. Thereby, the log-likelihood values are obtained by using Y_{ei} as dependent variable and the implicit probabilities of bookmaker j for positive scores defined by[32]:

$$Pos^{IP}_{i,j} = 1 * P^{IP}_{e=H,j} + 0.5 * P^{IP}_{e=D,j} \tag{15}$$

as covariates in a version only including bookmaker market information and alternatively a version based on the model forecasts respectively by:

$$Pos^{M}_{i} = 1 * P^{IP}_{e=H,M} + 0.5 * P^{IP}_{e=D,M} \tag{16}$$

This approach allows to assess whether each summary measure (bookmaker probability Pos^{ip}_{ij} or model probability Pos^{M}_{ij}) contains relevant information that the other does not contain (Forrest et al. (2005)). Finally, including both probability forecast instruments in the regression model as explanatory variables and conducting a likelihood ratio test for the individual significance of the coefficients reveals the degree to which our statistical forecasting model contains useful information that is not already captured by the bookmaker's odds.

The final part of this analysis will be conducted by an economic testing strategy testing the forecast accuracy of the model indirectly via betting profits, which are derived from betting rules based on the discrepancy between bookmaker and model forecast.

5.4.2 Results

Statistical Benchmark Model

The coefficient estimates from several specifications of the ordered probit model are collected in Table 9. The selection of the best model specifications is based on a training sample including the full sample of historical matches within the German as well as English divisions from season 2002 until 2006 and is conducted via the AIC criteria. Column (1) lists the coefficient estimates for a parsimonious model only using the past performance, goals scored and recent match information covariates. The past performance coefficient signs are in line with our expectation by showing a positive effect of increasing past home team performance on the probability of a positive home result and vice versa a negative form for the opponent's performance. Interestingly, the recent match result coefficients

[31] This approach is based on Forrest et al. (2005).

[32] Note that the covariate is equal the expected score results conditional on the bookmaker forecast $Pos^{IP}_{i,j} = E[Result_i | P^{IP}_{i,j}]$.

indicate a positive persistence effect at least in the short run. Including now in column (2) the information about the teams' promotion or relegation histories as well as the non-performance related covariates $Attend_{ij}$ and $LnDist_{ij}$ increases the model fit significantly by a decrease in AIC from 13,204 to 13,145. The coefficient signs indicate a positive fan base effect for home teams with superior pre-season attendance and an increasing chance of home victory the larger the distance between the team's origin is, though by substantial lower magnitude compared to the performance variables. Column (3) uses the alternative moving average window of 20 prior match observations but performed slightly worse. Alternatively, a specification is listed in column (5) using the moving average windows conditional on home and away ground respectively by $Xsep$ and $Ysep$ but is also showing an inferior performance according to the AIC criteria. In the end, using the forward/backward algorithm including several other specifications, the model in column (6) is defined as the best specification given the selection of variables using a 30 match day window for past result performance, while reducing the moving average window for the scoring rates to 20 match days. Surprisingly, excluding the variable $Last_i$ improved the model fit slightly, which indicates that the last meeting of the contestants reveals no relevant information for future meetings. All coefficients are significant at a 1% significance level except for the recent match results. Nevertheless, including them improves the model fit. Due to lack of space, the threshold estimates are not listed in Table 9 but are given in model (6) for completeness by $\hat{\mu}_{1s} = -0.495$ and $\hat{\mu}_{2s} = 0.226$.

Comparison of Forecasting Performance

After estimating our statistical benchmark model and being able to derive forecasts opposing the bookmaker predictions, this section will compare the resulting forecast accuracies. Figure 5 displays the average seasonal Rank Probability Scores (RPS) separately for the German and English league system. The blue line plots the performance of the median bookmaker forecasts active in the respective country and is opposed to the black line for the RPS of the statistical model. First of all, the optimized statistical model described above is not able to beat the median bookmaker accuracy for the entire sample period based on the RPS scoring measure. The median bookmaker RPS is consistently below the model RPS, which indicates a higher forecasting accuracy for odds setters, though in a comparable range to the statistical model. For the British divisions the accuracy appears to be more stable from season to season while the forecasting precision for German football matches is characterized by a anticlimax in 2007 and 2010-2011, where the bookmaker market and the model reveal a lower accuracy. Fortunately, the model follows a similar path as the bookmaker market in general, which implies that the model fit adopts the dynamic inter-season changes in relative team strength for the underlying divisions in a similar way as the betting market. In Table 10 the seasonal Brier scores (BRS) and

RPS are summarized for a selection of bookmakers and the benchmark model.[33] The first panel contains the BRS for home victories, whereby the average bookmaker BRS is given by 0.22 compared to a value of 0.23 for the benchmark model. Overall, the performance measures are very much alike across all bookmakers and there is only a slight variation over time. The next two panels list the BRS conditional on draw and away win forecasts respectively. It appears that the bookmakers as well as the model perform better in forecasting draw outcomes and away victories with individual average BRS of 0.19 than a successful outcome for the home team. Furthermore, there is no visible trend that would indicate an improvement of the betting market's forecasting accuracy in recent years. All in all, ranking the individual bookmakers according to the BSC or RPS criteria is not very successful since the performance is very homogeneous. Contrarily to the conclusion of section 5.3.2, there is no evidence for substantial fluctuation in the accuracy measures clustered in some periods.

Examining the difference between the model forecasts and the bookmaker predictions more precisely now, Figure 6 plots the average discrepancy between the model and the median bookmaker forecasts for different implied probability ranges of the bookmaker forecasts[34]. The first panel plots the discrepancy for home victories of the German and English match observations, while the second panel examines the difference for away victory bets separately for the national league systems. First of all, there is a negative relationship between the average discrepancy and the bookmaker market's implied probability range. If the implied chance for an outcome of a given match is low, the benchmark model predicts on average a higher probability for the respective outcome. This tendency is reversing the more the implied bookmaker forecasts approach unity. Hence, we can conclude that low likelihood outcomes (underdog victories) are in tendency higher estimated by the model and likely outcomes (favourite victories) lower in comparison to the bookmaker market. For home win outcomes the individual plots for Germany and England are very similar, while in the second panel high probable away win outcomes appear to be slightly more overestimated by the market relative to the model forecasts in case of German match observations. Highlighting the difference in the home-away dimension, the discrepancy at the lower boundary of the probability space is dominated by home victory outcomes and vice versa for the upper boundary part by the opponent's favourable outcome side. In conclusion, the model forecasts deviate from the bookmaker implied probability distribution with variation in the home-away or favourite-underdog dimension and therefore may capture potential inefficiencies in the market prices. The next section will assess whether the statistical model contains useful information that is

[33] Since RPS and BRS are very similar for all bookmakers, only 8 out of 14 bookmakers are listed in Table 10.

[34] The discrepancy is defined by: model forecast - median bookmaker forecast. The underlying match observations are grouped in probability intervals of 10 percentage points.

not captured by the market prices.

Informational Contribution of the Model

As already outlined in the methodology section, an additional ordered probit regression model will be applied in order to test the model's informational contribution. It includes the expected score results of the home team given the bookmaker and model forecasts respectively by $Pos_{i,j}^{IP} = E[Result_i|P_{i,j}^{IP}]$. The first panel of Table 11 contains the maximized log-likelihood functions which were obtained by ordered probit regressions using the match result as dependent variable and separately $Pos_{i,j}^{IP}$ of each bookmaker as the sole covariate. The sample is partitioned in two-yearly subsamples including the incomplete season 2007 and 2017 in the first and last subsample respectively to capture potential inter-temporal dynamics[35]. According to the maximised log-likelihood values, no bookmaker can be identified as consistently best performer . The maximized log-likelihood values of the ordered probit model using the model forecasts as covariates in regression (2) are consistently inferior similar to the comparison in Figure 5. However, a comparison of the log-likelihood functions including the model forecasts additional to each bookmaker forecast may reveal nevertheless the extent of useful information that is relevant for match outcome predictions which are not included in the bookmaker odds. The results of these specifications are listed in Panel (3). Since the semi-strong form efficiency hypothesis claims the absence of such information, we are able to test this hypothesis by likelihood ratio (LR) tests for significance of the model covariate $Pos_{i,M}^{IP}$ in model (3). The LR test statistics for the significance of the model forecast besides the individual bookmaker predictions are given in the subsequent panel. All in all, the analysis of the explanatory power of the model besides the betting prices reveales no indications for semi-strong form informational inefficiencies embodied by consistently low test statistic values throughout the subsamples. However, to be able to derive conclusions from the semi-strong form perspective, it is relevant to assess whether indications of informational inefficiency can be translated into economically significant differences in betting returns.

Economic Semi-strong Efficiency Test

The analysis will now be concluded with an economic evaluation of betting strategies based on the statistical forecasting model. According to the efficient market hypothesis, the expected return of betting at any identified odd should be identical (Williams (2005)). Hence, performing a fundamental analysis, e.g. incorporating publicly available information and evaluating the market prices in view of other forecasts, should not lead to higher or even positive returns in expectation. Building on the probability forecasts

[35]Each subsample contains at least two seasons and a selection of bookmakers with nearly complete match series to improve the balance of the number of matches per subsample since all match observations with missing bookmaker prices have to be excluded from the analysis.

from the ordered probit model now, a more sophisticated betting strategy than the simple trading rules like betting on home vs. away or underdog vs. favourite team respectively as descriptively examined in section 5.2 will be examined.

Precisely, the following betting rule is examined: Place 1 € on outcome e of a particular match only if the forecast of the statistical model, e.g. the fundamental analysis, exceeds the implicit probability of the bookmaker j, hence only if:

$$FR = \frac{P^{IP}_{e,Model}}{P^{IP}_{e,Bookmaker_j}} \qquad > \qquad X \tag{17}$$

The ex-post mean betting returns for different values of X are collected in Table 3. These strategies may lead to bets placed on more than one outcome in a particular game in case the model forecast exceeds the bookmaker probability not only for one outcome. However, alternative betting rules such as selecting only the most favourable outcome per match are not discussed, since these may not necessarily constitute superior rules as mentioned by Kuypers (2000). The first row in Table 3 contains betting returns of randomly betting on each bookmaker. The sample of match observations is slightly reduced compared to the return analysis in Table 2, as now we can only examine match observations for which the benchmark model was able to produce result forecasts. Nevertheless, the average returns are closely comparable.

Now, the betting strategy in the second row is given by selecting bets if the statistical model forecast exceeds the implied chance of the offered betting price by at least 10%, which reduces the number of placed bets significantly compared to the full sample. Comparing the ex-post returns with randomly betting by selecting bets according to this strategy, the returns improve for some bookmakers slightly up to an enhancement by 4.57 percentage points in case of German 1 while the returns of betting on EU 1 and UK 2 are only improved by less than 1 percantage point. Applying stricter values for X and selecting bets only if the model predicts a 20% higher chance of occurrence, the returns consistently improve for all bookmakers on average by 2.74 p.p.. Increasing the threshold X to 1.3 and 1.4 , a continuous improvement of returns is still achieved for most of the providers. Overall, the average negative market return of randomly betting by -7.96% seems to be improved by the selective betting strategies and even the most selective strategy (FR > 1.4) reaches an average market return by -1.72%.

In order to be able to evaluate the economic impact on seasonal returns over the sample period more closely, Table 4 lists the percentage returns of selection criteria FR > 1.4 for each season separately. Thus, a similar picture to the results in Table 10 and 14 is revealed. Remarkably, in season 2011 and 2014 positve ex post betting returns are clustered. Precisely, the betting strategy in season 2011 achieves a positive average market return

	GER 1	GER 2	GER 3	EU 1	EU 2	UK 1	UK 2	UK 3	EU 3	EU 4	UK 4	EU 5	UK 5	UK 6
Random Betting	**-17.25**	**-6.46**	**-7.36**	**-2.30**	**-5.45**	**-7.74**	**-7.95**	**-5.72**	**-7.89**	**-8.53**	**-6.87**	**-10.82**	**-8.93**	**-8.10**
# Bets	10.411	14.521	14.403	14.219	14.899	13.016	14.408	13.875	14.874	14.940	14.721	14.964	11.956	10.590
Betting Strategies based on Forecast-ratio FR > X:														
Bet if FR > 1.1	**-12.68**	**-5.07**	**-5.91**	**-1.61**	**-4.16**	**-6.38**	**-6.98**	**-4.69**	**-6.08**	**-7.38**	**-4.67**	**-8.42**	**-7.89**	**-6.69**
# Bets	6.976	9.873	9.422	10.040	10.198	8.913	9.518	9.467	9.873	9.959	9.902	9.445	7.906	7.122
Bet if FR > 1.2	**-6.91**	**-3.39**	**-3.41**	**-0.29**	**-3.16**	**-3.95**	**-3.16**	**-3.22**	**-3.53**	**-4.08**	**-2.57**	**-4.84**	**-4.05**	**-3.71**
# Bets	4.067	6.233	5.758	6.648	6.593	5.535	5.841	5.986	5.951	6.097	6.226	5.292	4.816	4.390
Bet if FR > 1.3	**-4.12**	**-2.32**	**-2.29**	**-0.15**	**-1.68**	**-2.92**	**-2.51**	**-1.79**	**-2.29**	**-3.46**	**-0.91**	**-3.48**	**-3.32**	**-2.56**
# Bets	2.319	3.924	3.477	4.390	4.257	3.449	3.613	3.833	3.670	3.775	3.952	2.973	2.956	2.752
Bet if FR > 1.4	**-2.44**	**-1.52**	**-1.47**	**-0.79**	**-1.48**	**-1.63**	**-2.42**	**-1.41**	**-1.61**	**-2.31**	**-0.51**	**-2.19**	**-2.37**	**-1.91**
# Bets	1.304	2.645	2.201	2.989	2.835	2.200	2.340	2.525	2.299	2.343	2.575	1.682	1.875	1.797

Note: Indications are made in percent

Table 3: Percentage Returns from Betting Strategies based on Model Forecasts

by 1.14%.

Finally, to be able to draw some statistical conclusions from the analysis of the above mentioned betting strategies, a simple bootstrapping procedure (see for example Efron (1979)) is used. By using bootstrapping methods the entire data sample is repeatedly re-sampled with replacement yealding Monte Carlo distributions of the estimated mean returns which can be used for the construction of return confidence intervals. Following Dixon and Pope (2004), it can be noted that as long as the trading rule is prespecified before the data is examined, any significant deviation from the random betting return can be interpreted now as representative, using this technique. Figure 7 plots the observed returns conditional on different forecast-ratios between 1.0 and 1.4 for 6 selected bookmakers out of the sample combined with their bootstrap confidence intervals. It emerges clearly, for all bookmakers there is a positive trend visible for increasing forecast-ratios.

The dashed line represents the random betting return of the respective bookmaker and is significantly outperformed by the observed returns in case of German 1 and 3 as well as EU 4,5 and UK 4 using ratios above 1.2. Especially the observed returns in case of the the first mentioned bookmaker demonstrate a striking effect of the fundamental analysis, revealing again the provider's lower relative efficiency. However, none of the bookmakers' odds provide the opportunity to achieve significant positive returns.

In conclusion, Figure 7 and Table 3 indicate that abnormal returns can be achieved by trading strategies based on a statistical forecasting model that incorporates publicly available information. However, it is not possible to achieve substantial positive returns throughout the sample by applying the above mentioned sophisticated betting strategies such that these inefficiencies could be exploited to earn positive returns in expectation. Furthermore, taking into account that Table 3 and Table 4 contains pre-tax returns,

	GER 1	GER 2	GER 3	EU 1	EU 2	UK 1	UK 2	UK 3	EU 3	EU 4	UK 4	EU 5	UK 5	UK 6	∅ Market
Season	Betting Return based on Strategy with Forecast-ratio FR > 1.4:														
2008	-3.45	-2.01	-1.96	-1.08	-0.89	-2.82	-3.92	-1.54	-1.44	-3.70	**1.58**	-2.47	-3.15	-5.32	-2.30
2009	-1.80	-2.72	-3.11	-1.08	-1.12	-2.77	-2.98	-3.10	-3.17	-2.27	-1.65	-1.31	-2.48	-2.58	-2.30
2010	-3.08	-6.06	-2.89	-3.69	-6.98	-3.40	-2.40	-2.98	-5.80	-4.49	-5.11	-5.25	-4.61	-3.90	-4.33
2011	**0.31**	**0.98**	**1.66**	**0.04**	**0.34**	**2.39**	**0.50**	**2.32**	**2.19**	-0.18	**1.78**	**1.12**	-0.44	**2.94**	**1.14**
2012	-2.38	-1.13	-0.37	-1.17	-1.39	-0.05	-3.89	-0.75	**0.27**	-1.48	**0.33**	-0.76	-0.99	**0.72**	-0.93
2013	-2.68	-1.99	-1.69	-0.57	-1.55	-2.66	-2.24	-0.30	-2.69	-1.26	-1.05	-3.18	-1.90	-	-1.83
2014	-0.83	**1.94**	**0.03**	**2.27**	**0.14**	**1.84**	-0.75	-0.46	-0.23	-1.60	**0.53**	-1.39	-0.59	-	**0.07**
2015	-2.77	-1.53	-1.31	-0.98	-0.62	-3.88	-1.22	-2.46	-0.76	-3.62	-1.01	-3.04	-3.64	-1.12	-2.00
2016	-4.10	-2.63	-3.18	-0.60	-1.48	-1.14	-3.10	-1.83	-1.95	-2.45	**0.85**	-3.42	-	-0.57	-1.97
2017	-2.91	-0.77	-4.06	-1.26	-0.27	-4.98	-5.22	-0.61	-3.50	-2.73	-2.22	-2.64	-	-4.10	-2.71

Note: Indications are made in percent. The minimum number of bets per bookmaker for an individual season is 80. Positive ex-post returns are marked in bold.

Table 4: Seasonal Returns from Strategy FR > 1.4

the figures are lower for costumers in the German market subject to a 5% tax on gross winnings. Nevertheless, the substantial reductions in bookmakers' take-out indicates that the ordered probit model's information is also economically important and could be used to achieve superior, though still negative expected returns.

6 Conclusion

This study investigates the efficiency of the European football online betting market, a rising gambling form which provides a great example for an examination of market rationality in general. Overall, the results are mixed concerning the neutrality of bookmaker prices and the market functioning in accordance to the efficient market hypothesis postulated by Eugene Fama. However, in view of existing market distortions as taxes, switching costs of changing betting providers and limitation in competition, the results of the analysis are indicative of a rational market equilibrium surprisingly close to the efficiency benchmark.

Assessing the implied probability distribution of the bookmaker market and the actual distribution of match result outcomes, the prices appear to be markedly efficient and in line with the underlying probabilities. Furthermore, an analysis of the bookmaker margins over time reveals a clear trend of decreasing market take-outs, potentially affected by changes in regulation and an increase in competition. Simulating simple betting strategies, as selectively betting on match outcomes, indicates that there is some variation in expected betting returns conditional on the underlying outcome and match characteristics. However, the deviations were not in line with the often quoted favourite-longshot bias, but rather indicated anomalies in match observations with a strong visiting team or a general tendency of the market prices to underestimate draw outcomes.

In the weak-form efficiency regression analysis, the different approaches of detecting potential inefficiencies in the market also yielded mixed evidence. A clear home or favourite-longshot-bias, market anomalies often detected in the literature, cannot be found. Hence cognitive errors, risk-loving gambling behaviour or other sources for irrational behaviour of the betting audience seem not to be systematically exploited by the bookmakers in their pricing decision. A test for the effects of sentiment bias in form of a fan support proxy was also not able to reject the weak-form efficient market hypothesis. Even though there is some evidence that many punters avoid to bet on unfavourable outcomes for their supported team (Williams (2005)), this irrationality is not visible in the market prices.

In the second part of the empirical analysis, the forecasting performance of the bookmaker market was challenged by a statistical model incorporating public information which would have been available prior to each match day such as historical performance, distance between the opponents' locations and relative fan support. Though, the model was not able to outperform the bookmakers' forecasting accuracy, it was shown that by using only a limited information set and a relatively simple statistical model, a highly comparable forecasting performance can be achieved. Incorporating additional information about the teams' constitutions or measures for the relative importance of a match as well as improving the forecast estimation procedure by updating the model more frequently potentially improves the forecast accuracy even further. However, by examining the model's explanatory content besides the market prices, which by the semi-strong form efficiency hypothesis should include all relevant information, it was not possible to beat the explanatory power of the bookmaker prices by including the statistical model's forecasts.

The final part of the analysis was constituted by an examination of trading strategies based on the fundamental analysis embodied by the statistical benchmark model. By using different ratios for model vs. bookmaker forecast, a substantial improvement in ex-post returns was achieved for most of the bookmakers. Selecting bets only on outcomes for which the statistical model had forecasted a 40% higher likelihood of occurrence than the implied bookmaker probability, the overall ex-post market return was improved from -7.96% to -1.72%. Examining the difference in ex-post returns under bootstrapping, supported a significant influence of the statistical model on observed returns, even though it was not able to leave the negative domain. To sum up, interpreting the stricter version of market efficiency solely by its informational performance, the analysis revealed no doubts that prices incorporate all relevant information, though by using this fundamental analysis for trading it was possible to improve the ex post returns significantly.

Comparing the overall results of this analysis to the findings in the literature, the market appears to be more efficient. This may be attributable to the different source of historical price data and the use of a larger dataset, enabling a more rigorous examination. However, the extension of the dataset with trading volumes or dynamic price informations would be a good starting point for future research.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Ali, M. M. (1977). Probability and Utility Estimates for Racetrack Bettors. *Journal of Political Economy*, 85(4):803–815.

Avery, C. and Chevalier, J. (1999). Identifying Investor Sentiment from Price Paths: The Case of Football Betting. *The Journal of Business*, 72(4):493–521.

Brier, G. W. (1950). Verification of Forecasts expressed in Terms of Probability. *Monthly Weather Review*, 78(1):1–3.

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.

Christensen, R. H. B. (2015). Ordinal—Regression Models for Ordinal Data. R package version 2015.6-28. http://www.cran.r-project.org/package=ordinal/.

Clarke, S. R. and Norman, J. M. (1995). Home Ground Advantage of Individual Clubs in English Soccer. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 44(4):509–521.

Constantinou, A. C. and Fenton, N. E. (2012). Solving the Problem of Inadequate Scoring Rules for Assessing Probabilistic Football Forecast Models. *Journal of Quantitative Analysis in Sports*, 8(1).

Dare, W. H. and MacDonald, S. S. (1996). A generalized model for testing the home and favorite team advantage in point spread markets. *Journal of Financial Economics*, 40(2):295–318.

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.

Dixon, M. J. and Pope, P. F. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, 20(4):697 – 711.

Dobson, S. and Goddard, J. (2000). Stochastic Modelling of Soccer Match Results. Working paper, University of Wales Swansea.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26.

EFS (2016). `http://www.european-football-statistics.co.uk/attn.htm`. Accessed: 2016-11-24.

EGBA (2016). Market reality - Factsheet. Technical report, European Gaming & Gambling Association.

Epstein, E. S. (1969). A Scoring System for Probability Forecasts of Ranked Categories. *Journal of Applied Meteorology*, 8(6):985–987.

EuropeanCommission (2012). Online gambling in the Internal Market. Commission Staff Working Document 596, European Commission.

Fama, E. F. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21(5):55–59.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417.

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5):1575–1617.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, pages 998–1004.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications*. Number 66 in Monographs on statistics and applied probability series. Chapman & Hall, London.

FootballData (2016). `http://www.football-data.co.uk/downloadm.php`. Accessed: 2016-11-18.

Forrest, D., Goddard, J., and Simmons, R. (2005). Odd-setters as forecasters: the case of English football. *International Journal of Forecasting*, 21(3):551–564.

Forrest, D. and Simmons, R. (2005). *Efficiency of the odds on English professional football matches*, pages 330–338. In Williams (2005).

Forrest, D. and Simmons, R. (2008). Sentiment in the betting market on Spanish football. *Applied Economics*, 40(1):119–126.

Franck, E., Verbeek, E., and Nüesch, S. (2010). Prediction accuracy of different market structures – bookmakers versus a betting exchange. *International Journal of Forecasting*, 26(3):448–459.

Franck, E., Verbeek, E., and Nüesch, S. (2011). Sentimental Preferences and the Organizational Regime of Betting Markets. *Southern Economic Journal*, 78(2):502–518.

Franck, E., Verbeek, E., and Nüesch, S. (2013). Inter-market Arbitrage in Betting. *Economica*, 80(318):300–325.

Gandar, J., Zuber, R., O'Brien, T., and Russo, B. (1988). Testing Rationality in the Point Spread Betting Market. *The Journal of Finance*, 43(4):995–1008.

Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2):331 – 340.

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

Golec, J. and Tamarkin, M. (1991). The degree of inefficiency in the football betting market. *Journal of Financial Economics*, 30(2):311 – 323.

GoogleMaps (2016). `https://www.google.de/maps/`. Accessed: 2016-11-23.

Gray, P. K. and Gray, S. F. (1997). Testing Market Efficiency: Evidence from the NFL Sports Betting Market. *The Journal of Finance*, 52(4):1725–1737.

Greene, W. H. (2012). *Econometric Analysis*. Pearson Education Limited, 7th edition.

Grossman, S. and Stiglitz, J. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*, 70(3):393–408.

Hill, I. D. (1974). Association Football and Statistical Inference. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 23(2):203–208.

Hvattum, L. M. and Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, 26(3):460–470.

Jensen, M. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6(2-3):95–101.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, 32(11):1353–1363.

Levitt, S. D. (2004). Why are gambling markets organised so differently from financial markets? *Economic Journal*, 114(495):223–246.

Long, J. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques. SAGE Publications.

Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36(3):109–118.

Mas-Colell, A., Whinston, M., and Green, J. (1995). *Microeconomic Theory*. Oxford University Press.

Nikkinen, J. (2014). The Global Regulation of Gambling: a General Overview. Working paper, University of Helsinki.

Peren, F. W. and Clement, R. (2016). *Der deutsche Glücks- und Gewinnspielmarkt - Eine quantitative Bemessung von regulierten und nicht-regulierten Glücks- und Gewinnspielangeboten in Deutschland*. MUR Verlag, München.

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):201–207.

Radner, R. (1979). Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices. *Econometrica*, 47(3):655–78.

Rebeggiani, L. (2009). The Liga Portuguesa Decision of the European Court of Justice - An Economist's View. *Rivista di Diritto ed Economia dello Sport*, 5(3):1353–1363.

Rebeggiani, L. (2015). Use and Misuse of Regulation in Fighting Betting Related Corruption in Sport - The German Example. *MPRA Paper No. 68610*.

Rebeggiani, L. and Breuer, M. (2017). Neue Ordnung, neues Glück? Ordnungs- und fiskalpolitische Aspekte des deutschen Sportwettenmarkts. *Wirtschaftsdienst*, 97(9):655–663.

Rebeggiani, L. and Rebeggiani, F. (2013). Which factors favor betting related cheating in sports? Some insights from political economy. In Haberfeld, M. and Sheehan, D., editors, *Match-Fixing in International Sports: Existing Processes, Law Enforcement, and Prevention Strategies*, pages 157–176. Springer.

Rubinstein, M. (1985). Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978. *The Journal of Finance*, 40(2):455–480.

Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36(4):2021–2064.

Sauer, R. D., Brajer, V., Ferris, S. P., and Marr, M. W. (1988). Hold Your Bets: Another Look at the Efficiency of the Gambling Market for National Football League Games. *Journal of Political Economy*, 96(1):206–213.

Shin, H. S. (1991). Optimal Betting Odds against Insider Traders. *Economic Journal*, 101(408):1179–85.

Shin, H. S. (1993). Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. *The Economic Journal*, 103(420):1141–1153.

Snyder, W. W. (1978). Horse Racing: Testing the efficient markets model. *The Journal of Finance*, 33(4):1109–1118.

Thaler, R. H. and Ziemba, W. T. (1988). Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries. *Journal of Economic Perspectives*, 2(2):161–174.

Vlastakis, N., Dotsis, G., and Markellos, R. N. (2006). Beating the Odds: Arbitrage and Winning Strategies in the Football Betting Market. Technical report, European Financial Management Annual Meeting Discussion Papers.

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.

Von Neumann, J. and Morgenstern, O. (1953). *Theory of games and economic behavior*. Science: Economics. Princeton University Press.

Williams, R. (2009). Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients Across Groups. *Sociological Methods & Research*, 37(4):531–559.

Williams, V. L. (2005). *Information Efficiency in Financial and Betting Markets*. Cambridge University Press.

Woodland, L. M. and Woodland, B. M. (1994). Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market. *The Journal of Finance*, 49(1):269–279.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57(298):348–368.

7 Appendix

7.1 Seemingly Unrelated Regression Estimator by Zellner (1962)

Model (5) can be generalized in matrix form where X_e includes the respective covariates for equation e abbreviated by integer 1-3. The model is based on the following principles

- Strict exogeniety of X_e: $E[\epsilon|X_1, X_2, X_3]$
- Homoscedasticity of disturbances $\epsilon = [\epsilon_1^{'}, \epsilon_2^{'}, \epsilon_3^{'}] : E[\epsilon\epsilon^{'}]X_1, X_2, X_3] = \sigma^2 I_N$
- Correlation between disturbances across equations: $E[\epsilon_e \epsilon_{e'}^{'}]X_1, X_2, X_3] = \sigma_{ee'} I_N$

and is defined by:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} = X\beta + \epsilon$$

and the covariance matrix of the disturbances is given by:

$$V(\epsilon) = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I & \sigma_{13}I \\ \sigma_{21}I & \sigma_{22}I & \sigma_{23}I \\ \sigma_{31}I & \sigma_{32}I & \sigma_{33}I \end{bmatrix} = \Sigma \otimes I = \Omega$$

Since the exact covariance matrix is unknown, the procedure by Zellner (1962) estimates the equations first by usual single equation least-square estimation and uses the least square residuals to estimate $\hat{\Omega}$. Consequentially, the feasible generalized least square estimator $\hat{\hat{\beta}}$ (FGLS) is given by a two-stage estimation procedure with the following final estimation step:

$$\hat{\hat{\beta}} = \begin{bmatrix} \hat{\sigma}_{11}X_1'X_1 & \hat{\sigma}_{12}X_1'X_2 & \hat{\sigma}_{13}X_1'X_2 \\ \hat{\sigma}_{21}X_2'X_1 & \hat{\sigma}_{22}X_2'X_2 & \hat{\sigma}_{23}X_2'X_3 \\ \hat{\sigma}_{31}X_3'X_1 & \hat{\sigma}_{32}X_3'X_2 & \hat{\sigma}_{33}X_3'X_3 \end{bmatrix}^{-1} \times \begin{bmatrix} \sum_{j=1}^{3} \hat{\sigma}_{1j}X_1'y_j \\ \sum_{j=1}^{3} \hat{\sigma}_{2j}X_2'y_j \\ \sum_{j=1}^{3} \hat{\sigma}_{3j}X_3'y_j \end{bmatrix}$$

7.2 Information on Betting Markets

7.3 Analysis Results

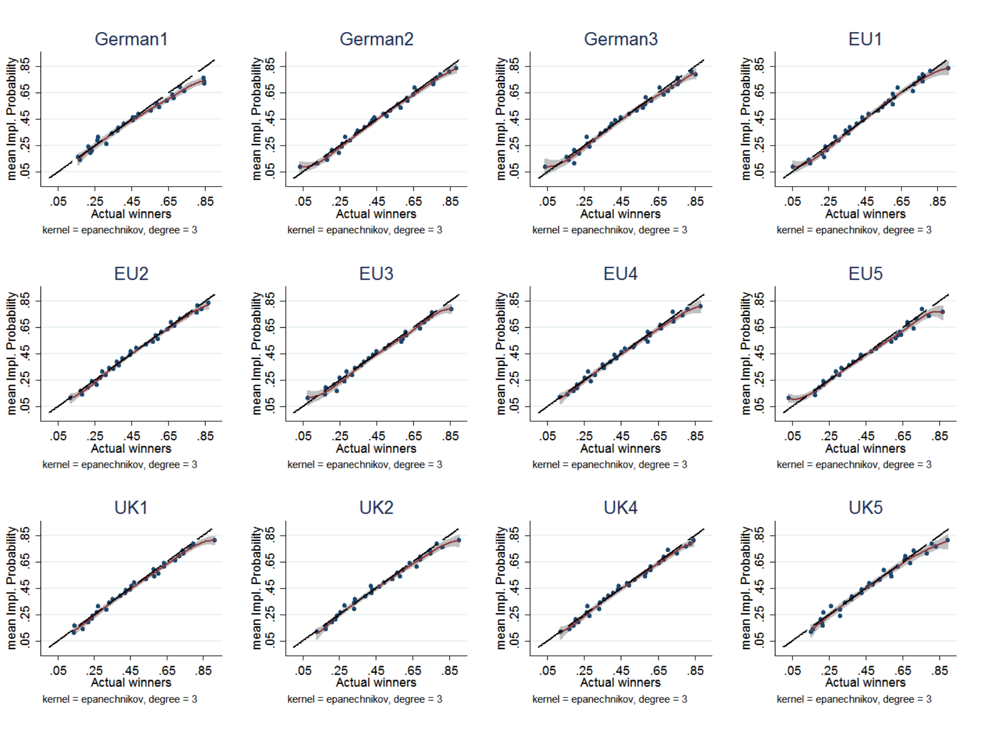

Figure 3: Implied probabilities versus outcome probabilities - Home victory

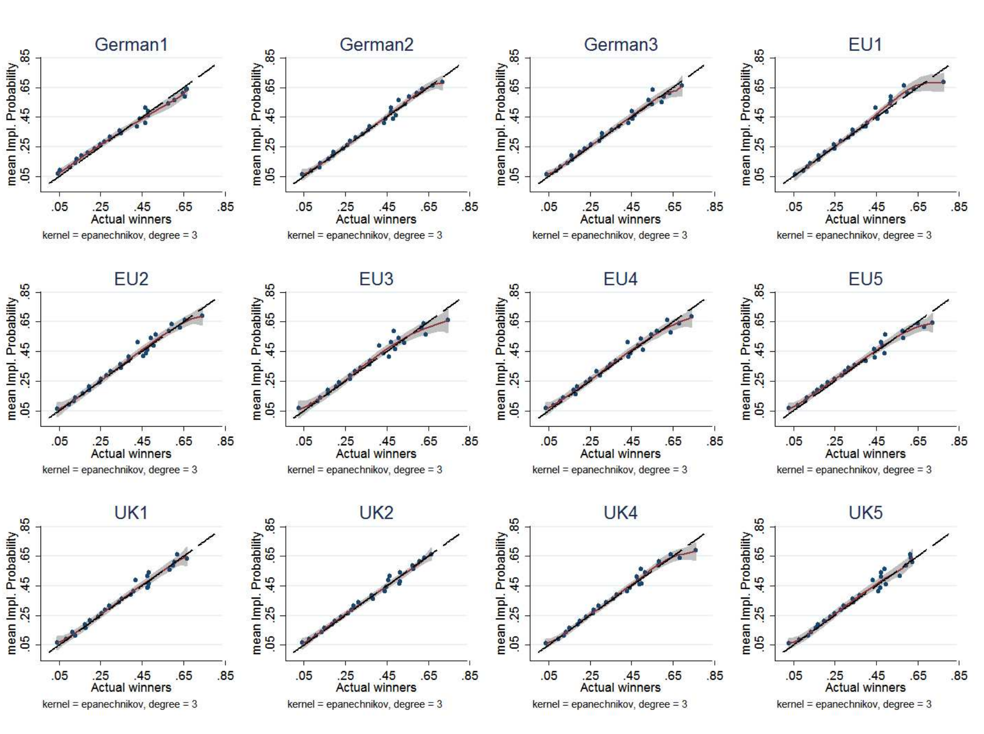

Figure 4: Implied probabilities versus outcome probabilities - Away victory

Table 5: Weak-form efficiency testing: SUR estimation results

	(1) GER 1	(2) GER 2	(3) GER 3	(4) EU 1	(5) EU 2	(6) UK 1	(7) UK 2	(8) UK 3	(9) EU 3	(10) EU 4	(11) UK 4	(12) EU 5	(13) UK 5	(14) UK 6
HomeV														
ImpProb_H	1.135*	1.032*	1.049*	0.996*	1.018*	1.046*	1.054*	1.017*	1.051*	1.042*	1.031*	1.122*	1.040*	1.030*
	(0.0298)	(0.0235)	(0.0246)	(0.0228)	(0.0231)	(0.0258)	(0.0243)	(0.0240)	(0.0243)	(0.0238)	(0.0238)	(0.0261)	(0.0267)	(0.0285)
_cons	-0.0424*	-0.00971	-0.0183	0.00133	-0.00579	-0.0118	-0.0179	-0.00439	-0.0157	-0.0126	-0.00921	-0.0424*	-0.0127	-0.00433
	(0.0137)	(0.0111)	(0.0116)	(0.0109)	(0.0110)	(0.0121)	(0.0114)	(0.0114)	(0.0114)	(0.0112)	(0.0112)	(0.0120)	(0.0126)	(0.0134)
AwayV														
ImpProb_A	1.135*	1.032*	1.049*	0.996*	1.018*	1.046*	1.054*	1.017*	1.051*	1.042*	1.031*	1.122*	1.040*	1.030*
	(0.0298)	(0.0235)	(0.0246)	(0.0228)	(0.0231)	(0.0258)	(0.0243)	(0.0240)	(0.0243)	(0.0238)	(0.0238)	(0.0261)	(0.0267)	(0.0285)
_cons	-0.0468*	-0.0140+	-0.0211*	-0.00323	-0.00946	-0.0194*	-0.0226*	-0.0129	-0.0215*	-0.0201*	-0.0150+	-0.0461*	-0.0201*	-0.0175+
	(0.00981)	(0.00772)	(0.00804)	(0.00752)	(0.00758)	(0.00842)	(0.00793)	(0.00790)	(0.00791)	(0.00783)	(0.00780)	(0.00849)	(0.00869)	(0.00929)
Draw														
ImpProb_D	-1.135*	-1.032*	-1.049*	-0.996*	-1.018*	-1.046*	-1.054*	-1.017*	-1.051*	-1.042*	-1.031*	-1.122*	-1.040*	-1.030*
	(0.0299)	(0.0236)	(0.0247)	(0.0228)	(0.0232)	(0.0259)	(0.0244)	(0.0241)	(0.0244)	(0.0239)	(0.0239)	(0.0262)	(0.0268)	(0.0286)
_cons	1.089*	1.024*	1.039*	1.002*	1.015*	1.031*	1.041*	1.017*	1.037*	1.033*	1.024*	1.089*	1.033*	1.022*
	(0.0222)	(0.0177)	(0.0186)	(0.0172)	(0.0174)	(0.0193)	(0.0182)	(0.0181)	(0.0182)	(0.0179)	(0.0179)	(0.0195)	(0.0202)	(0.0214)
N	10698	14927	14798	14571	15305	13360	14792	14272	15281	15348	15110	15372	12288	10921
F	5.425	0.651	1.256	0.404	0.358	1.255	1.501	0.891	1.445	1.371	0.779	5.406	1.169	1.088
P_value	0.00001	0.690	0.274	0.877	0.905	0.274	0.173	0.501	0.193	0.222	0.586	0.00001	0.320	0.367

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$

Table 6: Weak-form Efficiency Testing: OLS estimation with robust standard errors

	(1) GER 1	(2) GER 2	(3) GER 3	(4) EU1	(5) EU 2	(6) UK 1	(7) UK 2	(8) UK 3	(9) EU 3	(10) EU 4	(11) UK 4	(12) EU 5	(13) UK 5	(14) UK 6
HomeV														
P^{IP}_{Home}	1.140*	1.034*	1.052*	0.996*	1.017*	1.056*	1.058*	1.018*	1.054*	1.043*	1.034*	1.128*	1.046*	1.037*
	(0.0307)	(0.0242)	(0.0254)	(0.0235)	(0.0238)	(0.0267)	(0.0249)	(0.0248)	(0.0250)	(0.0246)	(0.0246)	(0.0267)	(0.0275)	(0.0297)
_cons	-0.0446*	-0.0105	-0.0196+	0.00145	-0.00567	-0.0166	-0.0194+	-0.00454	-0.0174	-0.0129	-0.0106	-0.0450*	-0.0157	-0.00730
	(0.0139)	(0.0112)	(0.0118)	(0.0111)	(0.0111)	(0.0122)	(0.0116)	(0.0116)	(0.0116)	(0.0114)	(0.0114)	(0.0121)	(0.0128)	(0.0137)
AwayV														
P^{IP}_{Away}	1.121*	1.014*	1.026*	0.981*	1.001*	1.021*	1.035*	1.004*	1.030*	1.028*	1.010*	1.099*	1.014*	1.008*
	(0.0338)	(0.0268)	(0.0281)	(0.0263)	(0.0265)	(0.0296)	(0.0276)	(0.0274)	(0.0276)	(0.0274)	(0.0271)	(0.0296)	(0.0305)	(0.0324)
_cons	-0.0427*	-0.00869	-0.0145+	0.00121	-0.00463	-0.0123	-0.0169*	-0.00901	-0.0154+	-0.0161*	-0.00896	-0.0391*	-0.0127	-0.0111
	(0.00993)	(0.00783)	(0.00820)	(0.00767)	(0.00772)	(0.00861)	(0.00803)	(0.00801)	(0.00804)	(0.00802)	(0.00792)	(0.00868)	(0.00884)	(0.00941)
Draw														
P^{IP}_{Draw}	-1.281*	-1.266*	-1.361*	-1.188*	-1.251*	-1.257*	-1.301*	-1.200*	-1.326*	-1.226*	-1.314*	-1.423*	-1.347*	-1.294*
	(0.116)	(0.0877)	(0.0951)	(0.0819)	(0.0865)	(0.100)	(0.0942)	(0.0894)	(0.0960)	(0.0891)	(0.0930)	(0.104)	(0.102)	(0.114)
_cons	1.196*	1.196*	1.269*	1.144*	1.187*	1.186*	1.221*	1.152*	1.240*	1.168*	1.232*	1.309*	1.259*	1.215*
	(0.0852)	(0.0651)	(0.0705)	(0.0611)	(0.0643)	(0.0737)	(0.0698)	(0.0665)	(0.0710)	(0.0661)	(0.0690)	(0.0769)	(0.0762)	(0.0841)
n	10698	14927	14798	14571	15305	13360	14792	14272	15281	15348	15110	15372	12288	10921
Chi2(6)	36.47	11.34	18.23	7.921	9.841	12.08	16.07	9.156	17.05	12.68	14.06	43.04	15.56	11.79
P_value	0.000002	0.0783	0.00569	0.244	0.132	0.0601	0.0134	0.165	0.00911	0.0484	0.0290	0.0000001	0.0163	0.0667

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$

Table 7: Robust OLS Estimation of Model (6) seasonally: Chi^2-statistic

	(1) GER 1	(2) GER 2	(3) GER 3	(4) EU 1	(5) EU 2	(6) UK 1	(7) UK 2	(8) UK 3	(9) EU 3	(10) EU 4	(11) UK 4	(12) EU 5	(13) UK 5	(14) UK 6
2007	12.57^{+}	14.71*	13.70*	-	12.62*	12.39^{+}	9.807	14.72*	12.86*	13.31*	12.89*	12.22^{+}	10.14	11.16^{+}
2008	15.93*	6.928	8.240	4.046	6.485	13.48*	10.86^{+}	7.579	8.747	8.739	7.989	14.89*	7.774	14.81*
2009	21.76*	2.964	2.896	1.262	2.170	2.412	2.088	2.528	3.935	3.405	2.351	8.229	3.080	2.140
2010	12.74*	4.406	6.139	4.157	4.438	3.926	5.304	4.761	6.323	6.257	6.134	8.736	5.943	4.879
2011	10.62	3.940	3.125	4.326	4.093	3.450	3.026	4.060	3.509	4.155	2.871	4.599	4.262	2.763
2012	6.399	1.225	1.470	2.081	1.128	0.803	2.899	0.871	2.326	0.660	1.210	3.828	2.324	4.300
2013	22.49*	8.590	9.639	8.814	9.767	7.426	9.944	7.300	10.49	6.363	9.910	14.03*	10.80^{+}	9.400
2014	14.22*	10.54	11.14^{+}	6.699	8.148	4.979	11.18^{+}	4.807	13.16^{+}	9.231	9.955	13.23*	10.83^{+}	-
2015	7.030	5.508	6.471	5.516	5.346	12.17^{+}	5.651	4.473	4.232	5.552	4.898	6.488	11.12^{+}	-
2016	6.555	5.736	6.423	5.608	4.383	6.053	2.532	4.767	4.104	3.708	5.810	7.101	-	2.539
2017	9.917	9.918	4.749	8.706	8.591	14.22*	10.43	9.542	11.01^{+}	10.30	9.780	13.68*	-	12.32^{+}

Note: The Table lists the test-statistic values for the robust OLS specification of model (6) estimated with seasonal subsamples for which the joint nullhypothesis of weak-form efficiency is tested. The resulting Wald-statistic values, which are following a Chi2(6)-distribution under H0, are listed for each bookmaker.

$^{+}$ indicates that the probability of falsely rejecting the null hypothesis (Type I error) is $p < 0.10$ and for * respectively $p < 0.05$.

Table 8: Clustered probit model

	(1) GER 1	(2) GER 2	(3) GER 3	(4) EU 1	(5) EU 2	(6) UK 1	(7) UK 2	(8) UK 3	(9) EU 3	(10) EU 4	(11) UK 4	(12) EU 5	(13) UK 5	(14) UK 6
$P^{ip}_{(home,away)}$	1.167*	1.048*	1.060*	0.989*	1.018*	1.029*	1.063*	1.029*	1.055*	1.049*	1.033*	1.136*	1.048*	1.031*
	(0.046)	(0.036)	(0.037)	(0.034)	(0.035)	(0.039)	(0.037)	(0.036)	(0.037)	(0.036)	(0.036)	(0.040)	(0.040)	(0.042)
$Home(d)$	0.010	0.011	0.010	0.014	0.012	0.018*	0.012	0.015+	0.013	0.015+	0.014	0.010	0.015	0.021*
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)
$\Delta Attend$	0.004	0.005+	0.005+	0.007*	0.006*	0.008*	0.006*	0.005+	0.006*	0.005*	0.006*	0.005*	0.005+	0.006*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
N	21396	29854	29596	29142	30610	26720.	29584	28544	30562	30696	30220	30744	24576	21842
N Cluster	10698	14927	14798	14571	15305	13360	14792	14272	15281	15348	15110	15372	12288	10921
pseudo R^2	0.092	0.087	0.085	0.089	0.087	0.086	0.088	0.089	0.086	0.087	0.087	0.085	0.088	0.086

The table lists marginal effects of clustered probit regressions; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1: $^{+}$ $p < 0.10$, * $p < 0.05$

Note: Each match observation is included twice for both winning outcomes (home win and away win) and standard errors are clustered for each match observation.

[H]

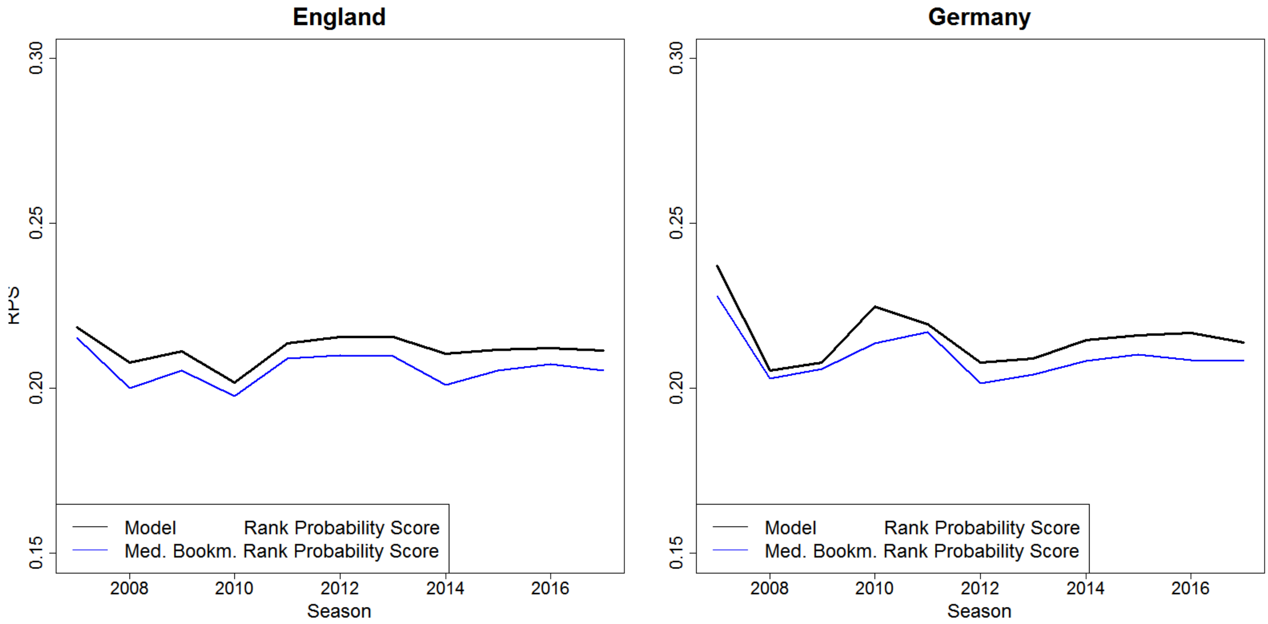

Figure 5: Forecast accuracy comparison

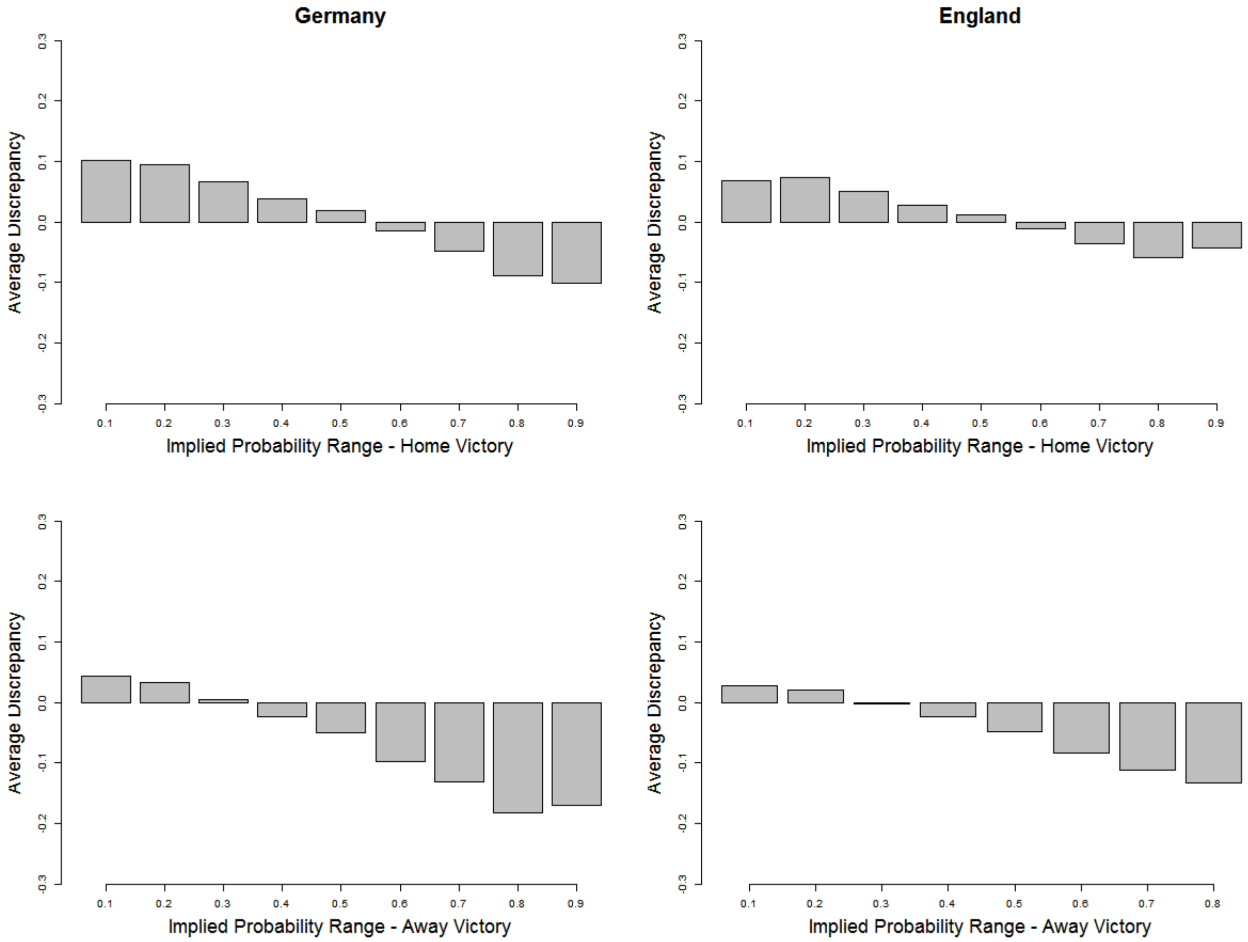

Figure 6: Discrepancy of model forecasts

Table 9: Ordered probit estimation results: Full Training sample: Seasons 2002-2006

	(1)	(2)	(3)	(4)	(5)	(6)
X_i	1.011***	0.966***				
	(0.209)	(0.211)				
X_j	−1.204***	−1.084***				
	(0.208)	(0.210)				
Y_i	0.171**	0.164**				0.178***
	(0.072)	(0.073)				(0.065)
Y_j	−0.194***	−0.209***				−0.304***
	(0.071)	(0.072)				(0.064)
$X2_i$			0.905***	1.000***		0.972***
			(0.185)	(0.174)		(0.169)
$X2_j$			−0.879***	−1.026***		−0.744***
			(0.182)	(0.172)		(0.167)
$Y2_i$			0.156**	0.158**		
			(0.062)	(0.062)		
$Y2_j$			−0.169***	−0.169***		
			(0.062)	(0.062)		
$Xsep_i$					0.813***	
					(0.189)	
$Xsep_j$					−0.877***	
					(0.175)	
$Ysep_i$					0.095*	
					(0.057)	
$Ysep_j$					−0.166**	
					(0.068)	
$R1_i$	0.044	0.042	0.026		0.085**	0.022
	(0.036)	(0.036)	(0.037)		(0.035)	(0.037)
$R1_j$	−0.061*	−0.052	−0.048		−0.099***	−0.048
	(0.036)	(0.036)	(0.036)		(0.035)	(0.036)
$R2_i$	0.071**	0.060*	0.042		0.077**	0.041
	(0.035)	(0.036)	(0.036)		(0.035)	(0.036)
$R2_j$	−0.082**	−0.074**	−0.072**		−0.096***	−0.072**
	(0.036)	(0.036)	(0.036)		(0.035)	(0.036)
$Last_i$	0.048	0.030	0.045		0.105***	
	(0.036)	(0.037)	(0.036)		(0.035)	
$Prom_i$		−0.297***	−0.252***	−0.256***	−0.316***	−0.266***
		(0.074)	(0.074)	(0.074)	(0.074)	(0.075)
Rel_i		0.238***	0.219***	0.221***	0.263***	0.227***
		(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
$Prom_j$		0.292***	0.251***	0.260***	0.378***	0.276***
		(0.077)	(0.077)	(0.077)	(0.077)	(0.078)
Rel_j		−0.234***	−0.200***	−0.206***	−0.284***	−0.211***
		(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
$Attend_{ij}$		0.091***	0.106***	0.107***	0.102***	0.099***
		(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
$LnDist_{ij}$		0.027**	0.026**	0.026**	0.018	0.027**
		(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
AIC	12882	12832	12844	12842	12877	12828
N	6,221	6,221	6,221	6,221	6,221	6,221
Log-Lik.	-6,430.146	-6,398.851	-6,404.890	-6,409.044	-6,421.377	-6,398.220

Note: *p<0.1; **p<0.05; ***p<0.01

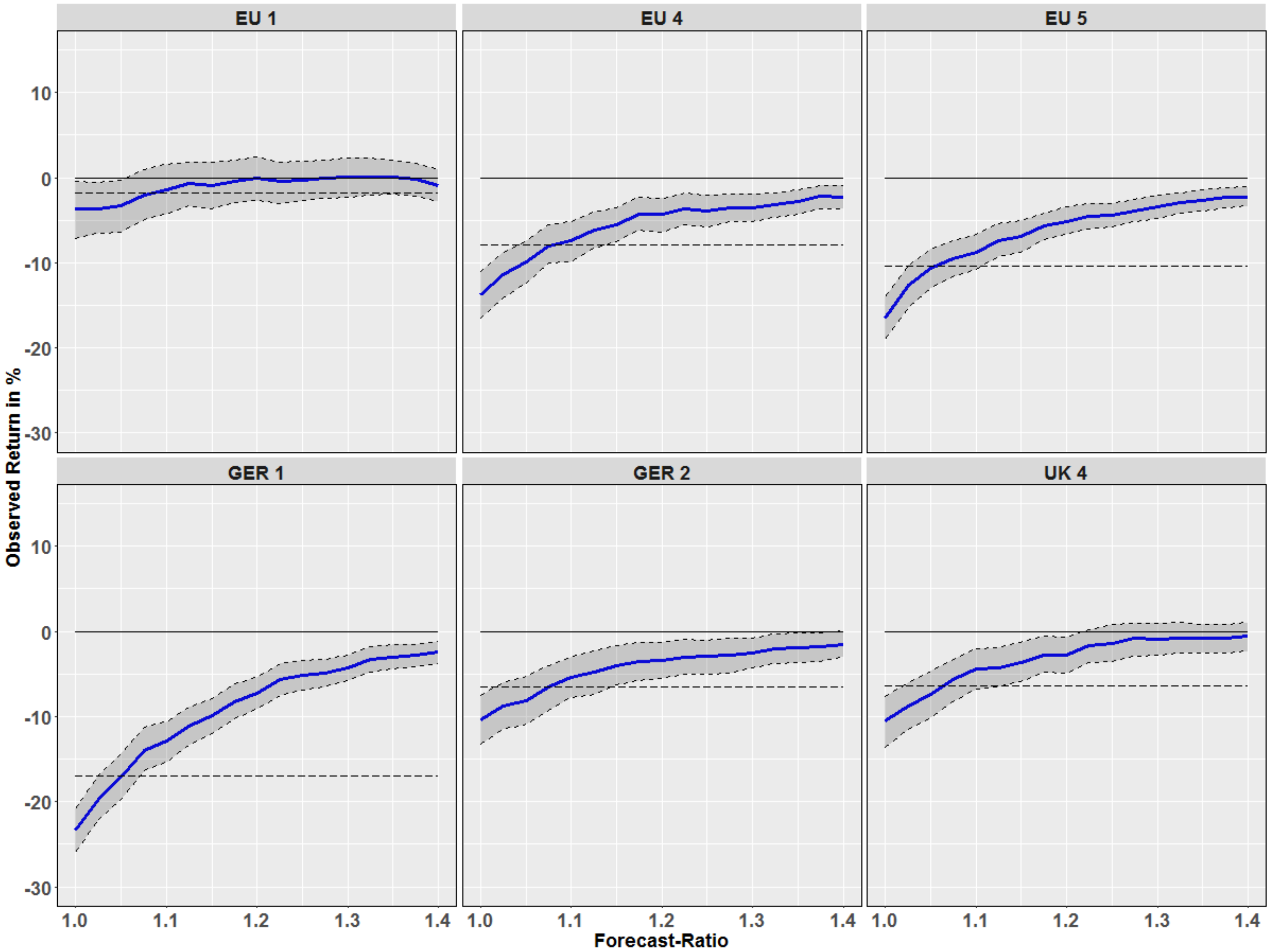

Figure 7: Observed returns conditional on different forecast-ratios

Note: Observed returns for the full odds set from 2007-2016 derived from forecast-ratios of model probabilities on bookmaker forecasts. The grey areas are 95% confidence Intervals obtained by bootstrapping with $R = 1000$. The horizontal lines are for zero return (solid line) and the expected return under random betting (dashed line).

Table 10: Brier Scores and RPS per Season: Full Sample

Season	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Brier Score: Home Win											
GER 1	0.235	0.224	0.224	0.227	0.244	0.215	0.215	0.219	0.224	0.224	0.230
GER 2	0.234	0.222	0.222	0.223	0.244	0.212	0.213	0.217	0.223	0.223	0.229
EU 1	-	0.223	0.225	0.244	0.211	0.214	0.216	0.223	0.223	0.229	0.222
EU 3	0.231	0.223	0.223	0.225	0.244	0.212	0.213	0.217	0.224	0.223	0.229
EU 5	0.232	0.224	0.222	0.225	0.243	0.213	0.213	0.218	0.224	0.224	0.230
UK 1	0.233	0.221	0.221	0.225	0.243	0.212	0.213	0.218	0.224	0.223	0.228
UK 2	0.228	0.222	0.222	0.225	0.244	0.211	0.213	0.217	0.224	0.223	0.229
Model	0.243	0.224	0.230	0.234	0.250	0.222	0.220	0.224	0.234	0.228	0.238
Brier Score: Draw											
GER 1	0.173	0.203	0.187	0.188	0.182	0.194	0.195	0.189	0.192	0.198	0.188
GER 2	0.175	0.203	0.187	0.187	0.181	0.193	0.195	0.189	0.191	0.198	0.187
EU 1	-	0.186	0.187	0.180	0.193	0.195	0.189	0.191	0.197	0.187	0.201
EU 3	0.172	0.203	0.186	0.188	0.180	0.193	0.195	0.189	0.191	0.198	0.187
EU 5	0.172	0.203	0.187	0.187	0.180	0.194	0.196	0.189	0.192	0.198	0.188
UK 1	0.168	0.204	0.186	0.188	0.180	0.193	0.195	0.190	0.191	0.197	0.188
UK 2	0.170	0.203	0.187	0.187	0.181	0.194	0.195	0.189	0.192	0.198	0.187
Model	0.171	0.204	0.187	0.187	0.182	0.195	0.197	0.190	0.193	0.198	0.189
Brier Score: Away Win											
GER 1	0.218	0.178	0.174	0.192	0.196	0.185	0.186	0.197	0.190	0.190	0.194
GER 2	0.218	0.174	0.173	0.189	0.197	0.184	0.182	0.195	0.190	0.189	0.194
EU 1	-	0.174	0.191	0.197	0.184	0.182	0.194	0.190	0.189	0.195	0.173
EU 3	0.216	0.175	0.173	0.190	0.197	0.184	0.182	0.196	0.190	0.189	0.195
EU 5	0.216	0.176	0.173	0.191	0.196	0.184	0.183	0.197	0.189	0.190	0.194
UK 1	0.214	0.176	0.172	0.191	0.197	0.184	0.183	0.196	0.189	0.189	0.194
UK 2	0.213	0.175	0.173	0.191	0.197	0.182	0.183	0.195	0.190	0.189	0.194
Model	0.223	0.177	0.179	0.197	0.196	0.190	0.191	0.201	0.194	0.194	0.198
Rank Probability Score											
GER 1	0.226	0.201	0.199	0.209	0.220	0.200	0.200	0.208	0.207	0.207	0.212
GER 2	0.226	0.198	0.197	0.206	0.220	0.198	0.198	0.206	0.207	0.206	0.211
EU 1	-	0.198	0.198	0.208	0.221	0.198	0.198	0.205	0.207	0.206	0.212
EU 3	0.224	0.199	0.198	0.208	0.220	0.198	0.198	0.206	0.207	0.206	0.212
EU 5	0.224	0.200	0.198	0.208	0.220	0.199	0.198	0.208	0.207	0.207	0.212
UK 1	0.224	0.198	0.196	0.208	0.220	0.198	0.198	0.207	0.206	0.206	0.211
UK 2	0.221	0.199	0.198	0.208	0.221	0.196	0.198	0.206	0.207	0.206	0.211
Model	0.233	0.201	0.204	0.216	0.223	0.206	0.205	0.212	0.214	0.211	0.218

Note: All Scoring measures are computed on same set of match observations, hence observations with missing bookmaker forecasts are excluded. The measures are individual seasonal averages whereby the Brier Scores are defined by: $BRS^e = \frac{1}{N}\sum_{i=1}^{N}\left(Y_{ei} - P_e^{IP}\right)^2$ and the Rank probability Score by: $RPS = \frac{1}{r-1}\sum_{i=1}^{r}\left(\sum_{j=1}^{i} Y_{ej} - \sum_{j=1}^{i} P_{ej}^{IP}\right)^2$.

Season	2007-2009	2010-2011	2012-2013	2014-2015	2016-2017
Number of matches	2628	2581	3752	1895	2367
(1) Log-Likelihood: ordered probit regression of match results on bookmaker covariate					
EU 2	-2652.4	-2630.1	-2809.6	-1915.7	-2390.6
EU 4	-2653.2	-2632.1	-2807.9	-1917.0	-2389.9
EU 5	-2654.7	-2635.8	-2809.1	-1920.2	-2395.2
UK 1	-2643.2	-2630.4	-2808.9	-1915.9	-2392.0
UK 4	-2652.4	-2632.4	-2807.7	-1916.3	-2390.8
GER 2	-2645.0	-2623.9	-2808.6	-1916.1	-2392.5
(2) Log-Likelihood: ordered probit regression of match results on model covariate					
Model	-2688.7	-2668.0	-2851.8	-1958.5	-2435.8
Best Performance	UK 1	GER 2	UK 4	EU 2	EU 4
(3) Log-Likelihood: ordered probit regression of match results on bookmaker and model covariates					
EU 2	-2651.7	-2630.0	-2809.0	-1915.6	-2389.9
EU 4	-2652.6	-2632.0	-2807.6	-1916.8	-2389.5
EU 5	-2653.9	-2635.8	-2808.8	-1920.0	-2394.6
UK 1	-2643.1	-2630.4	-2808.4	-1915.8	-2391.3
UK 4	-2651.7	-2632.4	-2807.3	-1916.2	-2390.2
GER 2	-2644.6	-2623.7	-2808.2	-1915.9	-2391.9
(4) Chi2(1) statistic of LR Test for significance of Model forecast in model (3)					
EU 2	1.406	0.039	1.321	0.182	1.482
EU 4	1.047	0.047	0.651	0.326	0.853
EU 5	1.652	0.051	0.597	0.347	1.153
UK 1	0.259	0.073	0.991	0.307	1.271
UK 4	1.274	0.001	0.772	0.194	1.250
GER 2	0.834	0.398	0.787	0.433	1.271

Note: $^{+}$ indicates that the probability of falsely rejecting the null hypothesis (Type I error) is $p < 0.10$ and for * respectively $p < 0.05$

Table 11: Informational contribution of model forecasts