Journal of Forecasting
J. Forecast. **28**, 426–444 (2009)
Published online 27 November 2008 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/for.1085

How Efficient is the European Football Betting Market? Evidence from Arbitrage and Trading Strategies

NIKOLAOS VLASTAKIS,[1] GEORGE DOTSIS[2] AND RAPHAEL N. MARKELLOS[1*†]
[1] *Department of Management Science and Technology, Athens University of Economics and Business, Greece*
[2] *School of Accounting, Finance and Management, Essex Business School*

ABSTRACT

This paper assesses the international efficiency of the European football betting market by examining the forecastability of match outcomes on the basis of the information contained in different sets of online and fixed odds quoted by six major bookmakers. The paper also investigates the profitability of strategies based on: combined betting, simple heuristic rules, regression models and prediction encompassing. The empirical results show that combined betting across different bookmakers can lead to limited but highly profitable arbitrage opportunities. Simple trading rules and betting strategies based on forecast encompassing are found capable of also producing significant positive returns. Despite the deregulation, globalization and increased competition in the betting industry over recent years, the predictabilities and profits reported in this paper are not fully consistent with weak-form market efficiency.

KEY WORDS betting markets; market efficiency; football; forecast encompassing

INTRODUCTION

Betting markets are experiencing an unprecedented growth over the past few years due to extensive deregulation, abolition of national monopolies and the advent of online gambling. Betting is evolving rapidly from a segmented state affair to a highly competitive global industry. The UK betting industry, with an annual turnover estimated at £39 billion for 2007, is leading the way internationally with the implementation of a series of significant changes, e.g., abolition of gambling tax for punters

*Correspondence to: Raphael N. Markellos, Department of Management Science and Technology, Athens University of Economics and Business, Office 915, 47A Evelpidon Str. 113 62, Athens, Greece. E-mail: rmarkel@aueb.gr
†Visiting Research Fellow, Centre for Research in International Economics and Finance (CIFER), Loughborough University, UK.

and legitimization of online betting.[1] In addition to the importance of this industry for the economy *per se*, betting markets have received much attention in the academic literature due to their similarities to financial markets. In a seminal paper, Thaler and Ziemba (1988) were among the first to argue that betting markets may be better suited than financial markets when testing for efficiency. The main advantage is that bets have a well-defined period of life at the end of which their value becomes certain and this makes the testing of market efficiency far less complicated.

In view of these developments, the present paper undertakes one of the first comprehensive empirical studies of international betting market efficiency. In particular, we examine the predictability of football match outcomes on the basis of information contained in the odds quoted by different bookmakers. Rather than concentrating on a single country or business model, we employ a broad dataset from six bookmakers which are based in different European countries and operate either online or through traditional coupons. We focus on European football matches since these constitute one of the most liquid and important betting markets. First, we examine whether arbitrage opportunities exist in the betting market. This is motivated by the impressive wealth of informal evidence and the vibrant tipster industry which has now emerged online with respect to the possibilities of arbitrage. To the best of our knowledge, the present paper is the first academic study that concentrates on investigating international arbitrage opportunities in European football betting using data from both online and traditional bookmakers. We then undertake an extensive descriptive analysis in order to examine the persistence of previously reported biases in odds and to see if any new biases have emerged in the online era. We implement a battery of betting strategies based on simple heuristic rules and on predictions from regression models. We also employ two model-encompassing techniques in order to combine the predictions from individual bookmakers and to subsequently formulate betting strategies. It is argued that if bookmakers employ different information and objectives when forming their odds, then encompassing may lead to superior forecasting performance. Aggregating information from models estimated over different odds data has not been previously explored in the market efficiency literature.

Despite the intense competition in the betting industry, the results from our empirical analysis suggest deviations from the standard weak-form market efficiency assumptions. In particular, we demonstrate that a limited number of highly profitable arbitrage opportunities, about one in every 200 matches, existed in the sample. These arbitrage opportunities could be exploited via combined betting across two or three bookmakers. However, we find that these opportunities become rarer in the more recent period studied. Moreover, they are further reduced if we account for the restrictions placed by the fixed-odds bookmaker in the sample on the number of matches in a coupon. If we consider only online bookmakers, then the number of arbitrage opportunities reduces to less than one in every 1000 matches. We analyse various rational explanations that may justify the existence of arbitrage opportunities. The descriptive analysis confirms that the well-known 'favourite-longshot' bias is prevalent in our dataset. It also provides new evidence of a consistent missestimation of the so-called 'home field advantage' in the bookmakers' odds in favour of the home team. The overestimation of the home field advantage combined with the favourite-longshot bias leads to a new reverse 'home-underdog' bias in the European football odds. We coin this as the 'away-favourite' bias. These patterns in the data motivate the estimation of Poisson count and multinomial logit regression models using odds from different bookmakers as regressors. Since

[1] For an overview of the recent developments in the betting industry see, for example, Mainelli and Dibb (2004) and the *Economist* Special Report on Gambling (30 September 2004). Detailed financial data on the UK industry can be found in the *Betting, Gaming and Lottery Duties* statistical factsheets (March 2007), prepared by HM Revenue and Customs.

DOI: 10.1002/for

differences do exist between the bookmakers under study, the use of forecast encompassing techniques is justified. The results confirm that a betting strategy which is based on encompassing Poisson model forecasts in order to forecast team scores is able to capture the away-favourite bias and to produce substantial out-of-sample profits.

The remainder of the paper is organized as follows. The next section reviews the relevant literature on betting markets and efficiency. The third section discusses the methodology with respect to the arbitrage and betting strategies considered. The fourth section presents the empirical results, while the final section concludes the paper.

LITERATURE REVIEW: BETTING MARKETS AND EFFICIENCY

Research on betting market efficiency and the underlying economics is certainly not new in the literature (for reviews see Sauer, 1998; Vaughan Williams, 1999, 2005; for recent relevant applications see the papers in the *Applied Economics* special issue on the Economics of Betting Markets, Volume 40, Issue 1, 2008). Direct similarities between specific financial assets and bets have also been discussed in the literature. For example, Ruhm (2003) demonstrated how positions in financial options can be viewed as simple bets, whereas Vecer *et al.* (2006) compared betting contracts with credit derivatives. Hodges *et al.* (2003) show that the return patterns from buying put options is analogous to the favourite–longshot bias in racing markets.

In general, informational efficiency requires that the market aggregates information from different sources, so that prices represent the best forecasts on the outcome of future events. This implies that no bettor or bookmaker can sustain returns that exceed transaction costs, i.e., the so-called margin or take. For bookmakers, this means that no bookmaker can operate at a greater margin than the others. For punters, it implies that no player can systematically achieve returns different from the margin. The use of betting markets in studying market efficiency and decision making under uncertainty can also be justified more broadly from the perspective of prediction markets. A recent strand of literature has shown that simple prediction markets, betting being the most common form, are able to efficiently aggregate information and to provide fairly accurate forecasts that outperform most moderately sophisticated benchmarks (see, for example, Wolfers and Zitzewitz, 2004).

The search for arbitrage opportunities is a natural starting point when examining market inefficiency. However, this issue has received somewhat little attention in the academic literature on betting, especially in an international context. This is probably due to the fact that betting markets were traditionally segmented across countries and the competition between bookmakers on the same sports event was not that intense. Data limitations and lack of odds from several bookmakers meant that arbitrage analysis was not practically easy to study. Hausch and Ziemba (1990) explored the potential for risk-free arbitrage profits in cross-track betting on US racetracks. They found that there are significant differences in prices on the same race from one track to the other due to the existence of different betting pools across tracks. They developed an arbitrage model to exploit such differences and by using historical data they showed that it can produce positive profits. Edelman and O'Brien (2004) developed their own model of arbitrage in racetrack betting and confirmed the existence of such opportunities in their sample of Australian thoroughbred races. In both these papers the identification of arbitrage opportunities is only possible *ex post*. This is due to the nature of pool betting, which does not allow for the payoff of bets to be certain until the last bet has been placed. This makes the exploitation of price discrepancies risky, thus rendering the use of the term arbitrage somewhat questionable. Pope and Peel (1989) mentioned a limited number of arbitrage opportunities

DOI: 10.1002/for

in their sample of four UK bookmakers' football odds from 1981 to 1982. More recently, Dixon and Pope (2004) examined the potential for arbitrage opportunities, using data from three bookmakers in the UK fixed-odds football betting market from 1993 to 1996. They concluded that there are no violations of no-arbitrage conditions. Lane and Ziemba (2004) developed strategies for 'perfect' and 'risk' arbitrage opportunities in the betting market for the exotic game of Team Jai Alai and examined the conditions that would allow the existence of such opportunities. Risky 'arbitrage' opportunities have also been identified by Paton and Vaughan Williams (2005) in the UK football spread betting market. They developed a so-called 'Quasi-Arbitrage' or 'Quarb' strategy, which exploits outliers—bookmakers whose spread differs substantially from the mean spread offered by all others—and found that it produced positive returns. Although arbitrage opportunities appear to be rare on the basis of the literature, especially in European football, casual online research reveals that a whole industry of specialized portals, information brokers and software packages has flourished in recent years on the alleged existence of vast arbitrage opportunities in sports betting. Since most of the academic research has examined odds in one market and with a limited number of traditional bookmakers, it would be instructive to examine the issue of betting arbitrage in the online era.

A second issue that has been examined extensively in the academic literature on betting is related to the existence of biases and their implications for market efficiency. Amongst the most prominent biases observed in sports betting markets is the favourite–longshot bias, first discovered in racetrack betting markets (see, for example Quandt, 1986, and Vaughan Williams and Paton, 1997). This bias implies that bets placed on favourites yield a higher return than bets placed on longshots. This phenomenon has been identified in a variety of sports betting markets, including the European football betting market (see, for example, Cain *et al.*, 2000, 2003). In some markets, such as the American football (NFL) and baseball spread betting markets, a reverse effect has been observed (see Golec and Tamarkin, 1991; Woodland and Woodland, 1994). Vaughan Williams and Paton (1998) provide a rationale that accounts for such disparities. Recently, Gil and Levitt (2007) verified the existence of the favourite–longshot bias in their study of the 2002 World Cup betting market. Notwithstanding, their overall conclusions were in favour of market efficiency as the market appeared to react quickly to the arrival of new information and there existed very few arbitrage opportunities.

Another bias that has been investigated in the literature is the home-field advantage which arises from the fact that a team plays better in its home ground.[2] Research on this issue in the NFL has shown that the betting market consistently underestimates the home-field advantage. This bias, in conjunction with the reverse favourite–longshot bias that has been observed in the NFL betting market, gives rise to what is termed in the literature the home-underdog bias. This justifies strategies whereby betting in home underdogs produces significantly higher returns than other simple strategies. However, the empirical evidence in support of such strategies is somewhat inconclusive (e.g., see Dare and MacDonald, 1996; Vergin and Sosik, 1999; Gandar *et al.*, 2001). In an effort to identify behavioural biases in sports betting markets, Avery and Chevalier (1999) found that the price paths of the NFL spread betting markets are affected by investors' sentiments, namely overreaction to new information. In the same direction, Durham *et al.* (2005) recently re-examined the same market, but found scant evidence in support of the regime-shifting model of investor sentiment of Barberis *et al.* (1998).

[2] The home-field advantage has proven difficult to assess and quantify in terms of probabilities and can be attributed to familiarity with terrain, social support from home crowds and referee biased behaviour (see, for example, Vergin and Sosik, 1999).

DOI: 10.1002/for

Another relevant direction in the literature on betting concentrates on the development of econometric models to predict the outcome of sports events. The study of the relationships in these models may reveal and take advantage of the biases in odds. Invariably, the economic value of the predictions resulting from such models is evaluated using betting strategies on historical data. For example, Pope and Peel (1989) developed a linear probability model to capture the relationship between objective probabilities of outcome occurrence and subjective probabilities implied in quoted odds. They used this model as a basis for several betting strategies in order to test weak-form efficiency. The authors undertook also a semi-strong form efficiency test by examining a betting strategy based on published predictions by specialists. They concluded that the market is efficient as no strategy yielded positive expected after tax returns. However, they were able to substantially decrease the expected losses, a fact that they considered as evidence that the odds do not meet the criteria of rational expectations. Cain *et al.* (2000) developed a model in which the goal-scoring processes of the home and away teams follow a Poisson and a negative binomial distribution, respectively, but found very few profitable opportunities. In contrast, Dixon and Coles (1997) and Goddard and Asimakopoulos (2004) reported substantial out-of-sample profits from strategies based on Poisson and ordered probit models, respectively. Both these papers test for semi-strong market inefficiency since they incorporate information on past match results and other publicly available explanatory variables.

METHODOLOGY

In order to understand how arbitrage and profitable betting strategies may be possible in football betting, it is useful to examine how odds are formed. Bookmakers are interested in securing a stable income through the margin, a percentage of the total stakes placed with them. The expected margin (gain) of a bookmaker on an event with n outcomes (a relevant model is discussed by Levitt, 2004) can be represented as

$$E(M) = 1 - \sum_{i=1}^{n} P_i \cdot w_i \cdot d_i \tag{1}$$

According to the above equation, the expected margin (M) on each match depends on the probability associated with each outcome (P_i), the percentage of bets on each outcome (w_i) and the quoted odds (d_i), respectively. As noted by Levitt (2004), this implies that there exist different ways by which bookmakers can set their prices. For example, bookmakers may try to forecast accurately game outcomes so that the odds reflect these expectations. Alternatively, they may try to forecast the distribution of bets on each outcome. Finally, it is possible that bookmakers use some combination of these two approaches.

Equation (1) implies that in order to calculate the actual margin that a bookmaker earns from a single match we need to know both the odds on every outcome and the distribution of bets across outcomes. Although information on quoted odds is publicly available, this is not the case for the distribution of bets. This means that we cannot calculate the actual margin of bookmakers but rather an *implied* margin, denoted hereafter as M'. Following the standard approach in the literature, this is estimated by assuming that the bets are equally distributed across outcomes and that the odds are set according to the true probabilities. In practice, bookmakers employ odds compilers who have special knowledge of specific sports in order to estimate the true probability of each possible

DOI: 10.1002/for

outcome. Assuming that odds are set on the basis of these true probabilities, then the *fair* odd on an outcome i is simply the reciprocal of the probability P_i of the occurrence of that outcome. However, if odds were priced exactly at their fair level according to the true probabilities then the expected bookmaker gain would be zero (this can be easily seen if one replaces d_i with P_i^{-1} in equation (1)). For this reason, actual odds are somewhat smaller than fair odds in order to allow a positive margin for bookmakers. Accordingly, actual odds do not correspond to true probabilities but to somewhat larger *implied* probabilities (denoted P'_i). The expected implied margin can be estimated as

$$E(M') = \left(\sum_{i=1}^{n} P'_i\right) - 1 = \left(\sum_{i=1}^{n} \frac{1}{d_i}\right) - 1 \tag{2}$$

Arbitrage opportunities in betting markets

From equation (2) it is obvious that if a punter was to place bets on all outcomes of an event with the same bookmaker, he or she would realize with certainty a loss equal to the bookmaker margin. In order to realize an arbitrage bet more than one bookmaker is needed. We denote as J the set of bookmakers quoting odds. The purpose of betting arbitrage is to take advantage of differences in the odds set by different bookmakers on the same event, so that the margin is reversed to the benefit of the punter, thus creating what is called an 'under-round' book. To achieve this, a bettor must take what is called a combined bet: select the maximum odds per outcome from the set J of available bookmakers and place a bet on each outcome with the bookmaker who offers the highest odd for that outcome. For an arbitrage opportunity to exist, the margin of the synthesized book ($\tilde{M}$) must be negative:

$$\tilde{M} = \left[\left(\sum_{i=1}^{n} \frac{1}{\max\limits_{j \in J} d_{ij}}\right) - 1\right] < 0 \tag{3}$$

where d_{ij} is the odd on outcome i reported by bookmaker j and $\max\limits_{j \in J} d_{ij}$ is the maximum odd on outcome i reported by all available bookmakers. If equation (3) holds, then the profits from arbitrage for punters are equal to the negative of the margin of the synthesized book. In order to exploit such an opportunity, a combined bet must be made by the punter. The total value of the combined bet must be divided between individual stakes in proportion to the odds of each outcome. It follows from equation (3) that the proportion ($\tilde{w}_i$) of the total amount staked in each outcome is

$$\tilde{w}_i = \frac{P'_i}{\sum_{i=1}^{n} P'_i} \tag{4}$$

A number of explanations can be put forward to explain the fact that arbitrage opportunities may arise, even within an efficient betting market (for a comprehensive study of the conditions that allow the existence of arbitrage in various settings, see Kallio and Ziemba, 2003; Ottaviani and Sørensen, 2005, discuss in the context of a theoretical model how arbitrage may arise in *pari-mutuel* betting due to the presence of a limited number of insiders). For example, it is possible that bookmakers are aware that they are posting odds that could be used in combined betting to achieve arbitrage profits. This does not necessarily imply a loss for them since they have to balance their own book rather than the synthesized book. Kuypers (2000) argued that inefficient odds could be set by

DOI: 10.1002/for

bookmakers as a result of their effort to maximize expected profits from bets placed by biased punters. Differences between books may arise between two bookmakers if, for example, one of them has a much stronger presence in a country with many supporters of a local team who, for sentimental reasons, may choose to bet 'against' the true odds. One can argue that differences in books arise also between bookmakers who operate purely online or non-online with fixed-odds through coupons (see Makropoulou and Markellos, 2007). Moreover, it is possible that difference between odds compilers may occur if one has superior knowledge of a local sport or market. This may also happen when the outcome of a particular sports event is difficult to predict or the expertise of odds compilers is hazy; for example, a match between two teams of equal merit or matches in lower divisions. It is also possible that for advertising reasons bookmakers post online non-profitable odds for a limited period of time on particular events. The rumours of arbitrage will boost traffic to their site and business will increase via cross-selling and banner-related revenues. Any short-term losses can be controlled since limits are placed on the monetary amount each punter can bet.

Betting strategies

Biases in quoted odds provide the motivation and basis for designing profitable betting strategies. A first step is to undertake a descriptive analysis in order to verify the existence of biases in a dataset and to evaluate their economic significance. This can be done easily by measuring the return produced by various simple betting rules that are designed to exploit biases in odds. In a second step, econometric forecasting models can be estimated in order to examine more carefully the predictability of match outcomes. In this paper we employ two alternative regression techniques: a Poisson count model to forecast match scores, and, a multinomial logit model to forecast match outcome probabilities.

It has been shown in the literature (e.g., Dixon and Coles, 1997; Cain *et al.*, 2000) that the goal-scoring processes of the home and away teams can be well approximated by independent Poisson processes. The density of the distribution of the number of occurrences of an event is given by

$$\Pr(Y=y)=\frac{e^{-\lambda}\lambda^{y}}{y!},\ y=0,1,\ldots, \tag{5}$$

where λ is the mean and variance parameter. The Poisson regression model is derived from the Poisson distribution through the parameterization of the relation between the mean parameter λ and the regressors x, using the exponential mean parameterization:

$$\lambda_i=e^{x_i'\beta},\ i=1,\ldots,n \tag{6}$$

The conditional density of y_i, the number of occurrences (goals) given x_i, the regressors, is

$$P(y_i|x_i)=\frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!},\ \ln(\lambda_i)=\beta' x_i \tag{7}$$

The Poisson regression model is estimated using the following maximum likelihood (ML) function:

$$\ln L(\beta)=\sum_{i=1}^{n}\{y_i x_i'\beta-e^{x_i'\beta}-\ln y_i!\} \tag{8}$$

DOI: 10.1002/for

The multinomial logit model assumes that the probability that the dependent variable y_i is observed in category s is given by

$$P(y_i = s) = \frac{e^{\beta_s' x_i}}{\sum_{k=1}^{S} e^{\beta_k' x_i}} \tag{9}$$

where y_i is a discrete variable that takes S different values and x_i is a vector of regressors. In our case, the dependent variable (y_i) will assume three values: home win (1), draw (X) and away win (2), so that $S = 3$. As previously, the independent variables (x_i) include the odds on all outcomes.

Poisson regression and multinomial logit regression are employed separately for each bookmaker in order to build forecasting models based on the information contained in the odds he has quoted. Subsequently, the individual forecasts are combined using two different encompassing techniques. The first, originally suggested by Bates and Granger (1969) and extended by Newbold and Granger (1974), is based on a weighted average of individual forecasts. These weights are calculated on the basis of the forecasting performance of each individual model. In the case of the Poisson models a natural weight can be obtained using the classic mean squared error (MSE) of each model. For the multinomial logit regressions the quadratic probability score (QPS) is used since this a popular metric of performance for probability forecasts (see Clements and Harvey, 2007). According to the second encompassing technique, originally proposed by Granger and Ramanathan (1984), individual forecasts are used as regressors for forecasting scores and outcomes for the Poisson and multinomial logit models, respectively.

In a final stage, the economic performance of individual and combined forecasts is assessed. For the Poisson count regression forecasts of goals the betting rule can be expressed as follows: '*If the forecast goal difference is positive and greater than some threshold T_1, bet on the home team; if it is negative and less than some threshold T_2, bet on away team; and, if its absolute value is less than some threshold T_x, bet on x*'. Accordingly, the rule for the multinomial logit forecast of outcome probabilities is the following: '*If the forecast probability for outcome i is greater than some threshold T_i bet on outcome i*'. Optimal threshold levels are determined using in-sample data by means of a genetic optimization algorithm as to maximize strategy return in a cross-validation sample. The strategy is finally tested using out-of-sample data.[3]

EMPIRICAL RESULTS

Our dataset contains odds on football matches reported on the Internet by five major online bookmakers: Bet365 (based in UK), Internet1×2 (Belize), Interwetten (Malta), Sportingbet (Malta) and William Hill (UK). These are coded hereafter as *A*, *B*, *C*, *D* and *E*, respectively. We also use odds from one fixed-odds bookmaker based in Greece, OPAP, coded hereafter as *F*, which did not allow online betting for the period under study. Online bookmakers are allowed to alter their odds at any time before a match takes place, whereas the odds in fixed-odds betting are fixed for a period before the match takes place. The odds from the online bookmakers are 'closing' odds, i.e., those that were quoted when bookmakers stopped accepting new bets. The dataset spans the period 2002–2004

[3] Standard hill-climbing optimization techniques were also used but were very difficult to converge and became trapped in local optima. The parameters used in the genetic algorithm optimization scheme were: population size 50, mutation rate 0.1 and crossover rate 0.5.

DOI: 10.1002/for

containing a total of 12,841 football matches and covering 26 different countries and events. It is a rather unique database, with respect to its size and the information it contains, and offers the opportunity to compare fixed and non-fixed odds. The available sets of odds are 55,977 and 28,092 for the online and fixed-odds bookmaker, respectively.

Arbitrage

As mentioned previously, in order for an arbitrage opportunity to exist, more than one bookmaker is needed. For this reason, the search for arbitrage opportunities was conducted in a subsample of 12,420 matches (96.7% of the total), for which at least two bookmakers quoted odds. We found a total of 63 cases of arbitrage in our sample representing a non-negligible 0.5% of all matches, or one in 200. More than half of the arbitrage opportunities returned more than 12%, with the maximum return exceeding 200%. The arbitrage returns had an average of 21.78% (SE 3.79%).

In practice, exploiting the 63 arbitrage opportunities would be limited by the fact that the fixed-odds bookmaker places restrictions on the minimum amount of matches that can be wagered upon in each coupon. Only for very few coupons is it allowed to place a single-match bet. It is interesting to note that if only online bookmakers are used in the analysis, then we detect just 10 arbitrage opportunities in 10,374 matches, i.e., implying a probability of 0.096%, or, less than one in every 1000 matches. This is to be expected, since fixed-odds bookmakers cannot alter their odds for a specific time period before the match takes place. This means that they cannot incorporate in their odds any new information that becomes available, making them more vulnerable to quoting odds that allow arbitrage (see Makropoulou and Markellos, 2007). As mentioned, an explanation of the remaining arbitrage opportunities could be based on price discrepancies that arise between bookmakers due to differences in pricing mechanisms, information, odd compiler opinions or clientele, respectively. Since the data used consist of closing odds and fixed odds, an explanation of arbitrage on the basis of advertising reasons cannot be justified.

A closer examination of the 63 matches for which arbitrage is possible provides insight into the conditions that allow the emergence of such opportunities. The majority of these matches (74.6%) involve international competitions, mostly EURO and UEFA Cup. Moreover, most of these matches involve teams of unequal merit since the average odds in the 'arbitrage sample' is almost three times greater than those for the whole sample. Home teams, which are the favourite 80% of the time in the whole sample, scored on average twice as many goals in the arbitrage sample as they did in the whole sample. Another interesting fact is that for roughly 13% of arbitrage matches there is a difference of opinion between bookmakers with respect to which team is the favourite. Such differences are much rarer (3%) over the complete sample. This analysis suggests that arbitrage opportunities arise in matches where there is increased ambiguity with respect to the comparative strength of each team. This tends to occur either for teams of comparable strength or for teams with highly unequal strength, respectively. Another explanation could be based on differences in the clientele between bookmakers, especially with respect to geographical dispersion. Although the fact that most of the bookmakers in the sample are online deprives us of any detailed knowledge in this respect, differences in clientele may imply, for example, variations in the support offered to particular teams and this may be reflected in the odds. This may also partially explain the high number of arbitrage opportunities that arise between the fixed-odds bookmaker, conducting business only in Greece, and the online bookmakers.

We must note that the number of arbitrage opportunities detected for the matches under study may well underestimate the true number. This is due to the fact that our sample of online bookmakers contains only closing odds, which are quoted just before the point bookmakers stop accepting further

DOI: 10.1002/for

bets. Any arbitrage opportunities that may have existed prior to that point may have been eliminated since the bookmakers can change the online odds. Finally, an inspection of the distribution of arbitrage across calendar time showed that arbitrage opportunities occur in clusters. This is to be expected since major sports events take place periodically.

Betting strategies

In order to evaluate the return from betting strategies we selected a subsample of the original data corresponding to four bookmakers, *B*, *D*, *E* and *F*, and 1486 matches, respectively. This subsample was selected so that it contains the maximum number of observations without missing values. Incorporating all six bookmakers was not possible since it would limit the sample size considerably.

As discussed previously, in an efficient market bookmakers should operate at comparable margin levels. The descriptive statistics of the implied margin (M') for the four bookmakers under study, given in Table I, suggest that the fixed-odds bookmaker charges a significantly higher margin. As argued by Makropoulou and Markellos (2007), this can be due to the fact that fixed-odds bookmakers must charge a premium on the margin in order to be compensated for the additional uncertainty that arises from the possibility of new public information emerging after the declaration of the odds.

In order to examine if the reported odds imply differences between bookmakers in terms of insider information we undertake the analysis proposed by Shin (1993). More specifically, we estimated the individual degree of so-called 'insider trading' for each set of quoted odds using the iterative method of Cain *et al.* (2001).[4] The relevant average values, also contained in Table I, are comparable with previous findings in the literature in terms of magnitude (e.g., see Cain *et al.*, 2003) and, as expected, they suggest that the fixed-odds bookmaker faces a significantly higher degree of insider trading. However, as argued by Makropoulou and Markellos (2007), the Shin model cannot differentiate between uncertainty arising due to variations in public information or to insider information, respectively. This means that the higher degree of insider trading estimated for bookmaker *F* may well reflect only the fact that he faces a higher degree of public information uncertainty. The corresponding differences between the online bookmakers are only marginal and can be considered consistent with market efficiency.

Since Table I reflects information only about the unconditional distributions, it is interesting to examine the correlation of the odds, margin levels and degrees of insider trading between different bookmakers. As expected, within each bookmaker, the odds on different outcomes appear to be significantly correlated, with the odds on 2 and X having the highest (positive) correlation in most instances. This implies that when the home team is the favourite, the probability of a draw is

Table I. Descriptive statistics of implied margin (M') and degree of insider trading (z)

	Bookmaker *B*		Bookmaker *D*		Bookmaker E		Bookmaker *F*	
	M'	z	M'	z	M'	z	M'	z
Mean	12.11%	6.16%	11.68%	5.95%	12.86%	6.55%	16.28%	8.30%
SD	0.70%	0.51%	1.04%	0.63%	0.72%	0.48%	0.74%	0.41%
Min.	9.03%	4.52%	7.39%	3.73%	6.71%	3.38%	14.72%	7.41%
Max.	15.19%	11.75%	36.11%	18.78%	25.61%	12.83%	20.00%	11.07%

[4] The relevant MATLAB routine is available upon request from the authors.

DOI: 10.1002/for

perceived to be significantly smaller than in the case when the away team is the favourite. Again, this is an indication of a potentially consistent missestimation of the home field advantage. In order to understand the behaviour of the margin, we regressed it against a number of 'synthetic' variables: the variance of odds, the range of odds and the expected gain or loss from attempting arbitrage in a particular match. The most interesting statistically significant result was that the range between the maximum and minimum odd on the same outcome reported by two different bookies is positively related to the implied margin. This could result from the fact that these two bookmakers perceive that risk is greater for matches where their competitors quote significantly different odds. Finally, the correlations between margin levels and degrees of insider trading were all significant at the 5% level and positive, except for two cases involving bookmaker *F*, but assume rather small values. More specifically, the average correlation between margin levels (*z*-values) is 19.5% (48.2%). This suggests that there may exist some pricing differences between the bookmakers under study.

To further examine the existence of biases in quoted odds, we evaluated the average return for a series of different simple betting rules and the results are shown in Table II. The existence of the favourite–longshot bias is confirmed since the rule that placed bets on favourites yielded a significantly higher average return than the rule that placed bets on longshots. Moreover, there is evidence that the home-field advantage is consistently overestimated. This is not apparent at first glance, since the rule that places bets on home teams has a significantly higher average return than the rule that places bets on away teams. However, one must keep in mind that the odds contain a favourite–longshot bias and the fact is that the home team is more often the favourite, as can be seen clearly in the table. This means that if one wishes to examine the effects of home-field missestimation he must first account for the favourite–longshot bias effects. Consequently, we have to examine the home ground effects separately for favourites and longshots. Indeed, the rules that take both home-field and favourite–longshot status into consideration reveal a totally different picture. Specifically, the rule that places bets on home favourites (longshots) has a significantly lower average return than the one that places bets on away favourites (longshots). This suggests that the home-field advantage is consistently, and perhaps purposefully, overestimated by bookmakers. Moreover, it is obvious that for all bookmakers the rule with the highest average return of all is the one that places bets on away favourites. Amazingly, this rule produces positive profits in the case of bookmaker B. The reason for the superior performance of this rule is that it exploits both the favourite–longshot bias and the home-field advantage missestimation. This combined effect is coined hereafter as the 'away-favourite' bias.

Table II. Average return (μ) for various simple betting rules

Rule	Bookmaker *B*		Bookmaker *D*		Bookmaker *E*		Bookmaker *F*	
	Bets	μ	Bets	μ	Bets	μ	Bets	μ
Bet on home team	1486	−9.30%	1486	−9.90%	1486	−10.61%	1486	−12.32%
Bet on draw	1486	−14.28%	1486	−12.14%	1486	−13.78%	1486	−17.43%
Bet on away team	1486	−17.17%	1486	−15.68%	1486	−17.30%	1486	−19.79%
Bet on favourites	1442	−3.98%	1447	−5.68%	1427	−5.88%	1442	−9.68%
Bet on longshots	1341	−22.31%	1423	−18.92%	1395	−20.61%	1413	−24.31%
Bet on home favourites	1128	−5.33%	1133	−7.00%	1113	−6.66%	1127	−9.83%
Bet on home longshots	107	−37.76%	117	−32.04%	121	−30.14%	131	−35.57%
Bet on away favourites	313	0.39%	314	−0.90%	312	−3.99%	313	−8.56%
Bet on away longshots	726	−24.94%	752	−22.13%	733	−24.50%	775	−26.63%

DOI: 10.1002/for

For the purposes of estimating the forecasting models and formulating betting strategies, the data were further divided into two subsamples: the first, containing 1186 observations, used for the estimation of models and betting strategy thresholds, and, the second, containing the last 300 observations, used for out-of-sample evaluation purposes. The results of the Poisson count regressions are presented in Table III. As suggested by the log-likelihood statistics, away team scores are easier to predict. As expected, the odds for the team whose score is the dependent variable in each regression are statistically significant and have a negative coefficient. This reflects the reverse relationship of odds with win probabilities: the greater the odds on a team, the less probable it is for that team to win and thus to score goals. The fact that the odds on draw have the same sign for both equations can be justified since draws are typically the outcome with the smallest score. More specifically, in our sample the average number of goals scored in matches that came to a draw is close to 2, whereas for home or away victory it is close to 3. Thus, the more probable it is for a match to end in a draw, the less probable it is for both teams to score many goals. Another interesting result is the fact that the implied margin enters the regression for the home team score in two of the four bookmakers with a positive coefficient. This suggests that bookmakers charge larger margins for home teams that score many goals and usually are favourites. It is possible that bookmakers exploit the home-field advantage missestimation bias by charging larger margins when there is a big difference in the relative strength of the two teams. As argued, in these cases odds may be more difficult to estimate. A similar explanation can be given for the fact that in the home team score regression results, for the two bookmakers for whom the margin is insignificant we find a statistically significant constant. For these cases, the implied margin may not accurately reflect the actual margin and the effect of the latter is captured by the constant. A somewhat puzzling result is the sign of the squared odd for the away team, in the away team score regression, which suggests an irrational relationship between odds and scores. If one examines the size of the relevant coefficients, it is easy to notice that this variable has a significant impact on the equation forecast only for very high odds, roughly over 30. As discussed previously, for matches where the relative strengths of teams are highly unequal, the

Table III. Poisson count regression results

Variable	Dependent variable: home team score				Dependent variable: away team score			
	B	*D*	*E*	*F*	*B*	*D*	*E*	*F*
Constant	—	0.363	—	0.574	—	—	—	—
	—	(2.311)	—	(3.655)	—	—	—	—
Odd1	–0.262	–0.259	–0.264	–0.263	—	—	—	—
	(7.811)	(7.526)	(7.099)	(7.130)	—	—	—	—
Oddx	0.111	0.208	0.237	0.128	0.242	0.229	0.230	0.365
	(4.115)	(4.448)	(4.388)	(3.830)	(11.342)	(11.131)	(10.863)	(10.062)
Odd2	—	–0.0190	–0.023	—	–0.198	–0.188	–0.188	–0.327
	—	(2.370)	(1.982)	—	(9.540)	(9.491)	(9.133)	(8.092)
$(\text{Odd1})^2$	—	—	—	—	—	—	—	–0.005
	—	—	—	—	—	—	—	(4.519)
$(\text{Odd2})^2$	—	—	—	—	0.003	0.002	0.002	0.009
	—	—	—	—	(5.627)	(7.002)	(5.415)	(3.713)
Implied margin	5.206	—	2.353	—	—	—	—	—
	(4.723)	—	(2.126)	—	—	—	—	—
Adjusted R^2	—	0.085	—	0.080	—	—	—	—
Log-likelihood	–1,836.8	–1,839.4	–1,837.4	–1,843.6	–1,613.7	–1,615.0	–1,616.2	–1,607.9

Absolute values of z-statistics appear in parentheses below the estimated coefficients.

DOI: 10.1002/for

odds are somewhat less accurate, particularly for the team that is the longshot; e.g., a difference in odds between 30 and 31 means that the difference in implied probabilities is 0.001. Finally, the sign of the coefficients for the odds on the away team and for the squared odd for the home team, which are significant in some regressions, could reflect biases that have not been identified by the model. Comparable results are obtained through the use of a multinomial logit approach and are presented in Table IV. The predictability of outcome probabilities for each bookmaker is similar to that of team scores. However, the multinomial logit model has a disadvantage in that the estimated coefficients cannot be readily explained.

The estimated models are used to provide forecasts of the relevant dependent variables, goals or outcome probabilities, respectively, for each bookmaker. These forecasts are then combined by using the two encompassing techniques discussed earlier. Table V displays the estimation results for the encompassing regression of the Poisson model forecasts. Since the underlying forecasts represent count data, Poisson regression is used also for estimating the encompassing regression. We can note that the statistically significant bookmaker forecasts in the regression are those that had the best fit in terms of log likelihood in the original Poisson model regression. However, the bookmaker *B* forecast is statistically significant only at the 10% level. The application of logit multinomial regression in order to combine the probability forecasts (from the individual multinomial logit models for each bookmaker) did not lead to statistically significant results. Table VI summarizes the performance of the two encompassing approaches in forecasting the goals scored by the home and away team, respectively. It also reports the performance of the Poisson regressions for the individual bookmakers. Although the results suggest that encompassing in general leads to slightly more accurate predictions of goals, it is not clear which encompassing technique is superior from those considered. Table VII presents the performance of encompassing in forecasting match outcome

Table IV. Multinomial logit maximum-likelihood estimation results

Coefficient	*B*	*D*	*E*	*F*
β_{11}	0.6367	0.6988	0.1231	0.2826
	(0.8443)	(1.0158)	(0.1813)	(0.3970)
β_{12}	−0.4162	−0.3607	−0.4669	−0.4186
	(2.1543)	(2.2095)	(3.0434)	(2.3778)
β_{13}	0.0316	0.0428	−0.0334	0.0093
	(0.3387)	(0.5357)	(−0.4524)	(0.0986)
β_{14}	0.1877	0.1154	0.4488	0.3275
	(0.5855)	(0.4094)	(1.6438)	(1.1145)
β_{21}	−1.3210	−0.8463	−1.0985	−1.4618
	(1.2085)	(0.9242)	(1.1554)	(1.7059)
β_{22}	0.3562	0.3983	0.2819	0.3066
	(2.0130)	(2.4992)	(1.9707)	(1.8132)
β_{23}	−0.2380	−0.1619	−0.2498	−0.2833
	(2.0847)	(1.5305)	(2.5091)	(2.3092)
β_{24}	0.3952	0.1381	0.3930	0.5366
	(0.8986)	(0.3660)	(1.0361)	(1.3912)
Log-likelihood	−1165.5	−1169.1	−1167.4	−1170.6

Coefficient β_{ij} corresponds to the equation for choice *i* for variable *j*. The equation for choice 3 (draw) is not estimated, since the probabilities for draw are computed directly from the probabilities for home and away team victory. There are four variables: 1, 2, 3 and 4, denoting the constant, odds for home victory, odds for away victory and odds for draw, respectively. Absolute values of *z*-statistics appear in parentheses below the estimated coefficients.

DOI: 10.1002/for

Table V. Poisson count forecast encompassing regressions

Variable	Home team score	Away team score
Constant	–0.495	–1.062
	(5.781)	(8.259)
Bookmaker *B* forecast	0.250	0.453
	(1.786)	(1.932)
Bookmaker *D* forecast	—	—
	—	—
Bookmaker *E* forecast	0.334	—
	(2.492)	—
Bookmaker *F* forecast	—	0.646
	—	(4.155)
Bookmaker *D* squared forecast	—	–0.076
	—	(2.625)
Adjusted R^2	0.084	0.108
Log-likelihood	–1842.1	–1604.6

Absolute values of z-statistics appear in parentheses below the estimated coefficients.

Table VI. Performance of Poisson count forecasting models in predicting goals

	B	*D*	*E*	*F*	WCF	RCF
Forecast variable: home team score						
Weight in combination	0.2506	0.2499	0.2513	0.2482	—	—
MSE	1.6231	1.6272	1.6186	1.6385	1.6212	1.6310
Forecast variable: away team score						
Weight in combination	0.2493	0.2484	0.2481	0.2541	—	—
MSE	1.1775	1.1818	1.1835	1.1554	1.1668	1.1478

WCF and RCF stand for weighted and regression-based forecast encompassing, respectively. MSE is the mean squared error.

Table VII. Performance of multinomial logit forecasting models in predicting outcome probabilities

	B	*D*	*E*	*F*	WCF	RCF
Forecast variable: probability of home team victory						
Weight in combination	0.2503	0.2496	0.2503	0.2498	—	—
QPS	0.4479	0.4493	0.4480	0.4488	0.4479	0.4485
Forecast variable: probability of away team victory						
Weight in combination	0.2508	0.2502	0.2498	0.2491	—	—
QPS	0.3431	0.3440	0.3445	0.3456	0.3437	0.3441
Forecast variable: probability of draw						
Weight in combination	0.2499	0.2498	0.2502	0.2501	—	—
QPS	0.3872	0.3875	0.3869	0.3869	0.3869	0.3870

WCF and RCF stand for weighted and regression-based forecast encompassing, respectively. QPS is the quadratic probability score.

DOI: 10.1002/for

probabilities. Encompassing in this case seems to lead to marginal, if any, improvements in forecasts accuracy.

The final step in the analysis is to use betting strategies in order to evaluate the economic significance of the forecasting approaches considered. For the purpose of estimating the betting strategy thresholds, the in-sample data were further divided into two subsamples. Roughly 80% of in-sample data was used in the estimation of thresholds, whereas the remaining 20% was set aside for cross-validation purposes. Once the thresholds were estimated, the profitability of each strategy was tested in the out-of-sample 300 matches and the results are given in Table VIII. In addition to the complete out-of-sample dataset, the strategies were also evaluated in a subsample consisting only of matches where the away team is the favourite. The motivation for this was provided by the previously discussed evidence on the profitability of the simple betting rule based on the away favourite bias. Overall, the results suggest that the betting strategies that are based on goal forecasts using the Poisson count regressions are more profitable. Betting strategy profitability can be enhanced if one concentrates only on the away-favourite matches. The most superior betting strategy corresponds to the regression-based encompassing of goal forecasts in the away-favorite subsample, yielding an impressive 13.3% average return per bet. This performance is the best even if we risk-adjust returns by dividing them by the standard deviation of betting returns. Moreover, this strategy was able to

Table VIII. Out-of-sample economic significance of forecasts

	B	*D*	*E*	*F*	WCF	RCF
Forecast variable: team scores						
Number of bets	24	25	20	8	31	56
Profitable bets/% of total bets	12/50%	14/48.3%	9/45%	6/75%	16/51.6%	29/51.8%
Total return	–0.930	–1.140	–0.450	4.750	0.600	3.180
Average return per bet	–0.039	–0.046	–0.023	0.594	0.019	0.061
Standard deviation	1.012	1.040	1.219	1.228	1.025	0.265
Return per standard deviation				0.484	0.019	0.231
Subsample of away favourites						
Number of bets	24	24	9	6	30	49
Profitable bets/% of total bets	12/50%	12/50%	6/66.7%	4/66.7%	16/53.3%	26/53.1%
Total return	–0.930	–1.300	1.000	0.300	1.600	6.520
Average return per bet	–0.039	–0.054	0.111	0.050	0.053	0.133
Standard deviation	1.012	0.988	0.852	0.831	1.025	0.444
Return per standard deviation			0.130	0.060	0.052	0.300
Forecast variable: outcome probabilities						
Number of bets	43	22	27	29	24	—
Profitable bets/% of total bets	19/44.2%	9/40.9%	10/37%	12/41.4%	11/45.8%	
Total return	–2.080	–6.000	–3.750	–5.400	–4.640	—
Average return per bet	–0.048	–0.273	–0.139	–0.186	–0.193	—
Standard deviation	1.123	0.908	1.212	1.026	0.922	—
Return per standard deviation						—
Subsample of away favourites						
Number of bets	29	20	7	20	21	—
Profitable bets/% of total bets	15/51.7%	9/45%	4/57.1%	11/55%	10/47.6%	
Total return	0.670	–4.000	–0.800	0.500	–2.800	—
Average return per bet	0.023	–0.200	–0.114	0.025	–0.133	—
Standard deviation	1.033	0.922	0.838	0.975	0.950	—
Return per standard deviation	0.022			0.026		—

WCF and RCF stand for weighted and regression-based forecast encompassing, respectively.

DOI: 10.1002/for

identify a significant number of profitable bets, placing a total of 49 bets in 300 matches, or 16.3% of all matches in the sample. This performance is not surprising, given the presence of the away-favourite bias in the odds and the fact that the strategy is based on the most accurate forecasting method. Overall, the results suggest that betting profits are possible using historical odds data from different bookmakers—something that is not consistent with market efficiency.

CONCLUSIONS

This paper examined the statistical and economic significance of forecasts of European football match outcomes. We extended the previous literature by analysing an extensive database of odds quoted by five major online bookmakers and one fixed-odds bookmaker. Despite the fact that efficiency should have increased over recent years due to regulatory changes and the emergence of the Internet, we found that the information contained in odds can be exploited to gain profits via arbitrage and active betting strategies. More specifically, although arbitrage opportunities appear to be rare, they are highly profitable, yielding returns between 12% and 200%. The implementation of simple betting strategies confirmed the existence of the favourite–longshot bias and revealed a new regularity in the odds which we named 'away-favourite' bias. In order to exploit more efficiently any information contained in the odds we estimated Poisson count and multinomial logit regressions, respectively. Moreover, the usefulness of encompassing these forecasts was also examined. Finally, the out-of-sample economic significance of the forecasts was also evaluated. Overall, the results suggested that formal econometric models lead to more accurate forecasts which can be employed successfully to form profitable betting strategies. Furthermore, encompassing goal forecasts seems to provide superior results.

The economic and statistical significance of the results in this paper suggest deviations from the weak-form international market efficiency paradigm for the European football betting market. As argued, this may well be the result of differences across bookmakers and players, variations in information and products, and behavioural biases of punters. The results are directly comparable with those reported for financial markets in terms of arbitrage opportunities (e.g., see Shleifer, 2000; Froot and Dabora, 1999) and international efficiency (e.g., see Kim *et al.*, 2006). Future research will concentrate on repeating the analysis undertaken in this paper using data from the increasingly popular betting exchanges (see Mainelli and Dibb, 2004, for an informal description). Betting exchanges operate similarly to a stock exchange by matching prices of buyers and sellers of bets through a double auction procedure. The profits of the exchange come from a small commission charged on the earnings, which is negligible compared to the traditional bookmaker take. Recent evidence by Smith *et al.* (2006) shows that betting exchanges are more efficient than other betting markets. It is possible that significant price discrepancies may emerge in comparison to bookmakers.

ACKNOWLEDGEMENTS

The authors are grateful for the comments and suggestions received from the participants of the 2006 European Financial Management (EFMA) Annual Meeting and from research seminar participants at the Essex Finance Centre, Essex University and ICMA, University of Reading, on a previous version of the paper entitled 'Beating the Odds: Arbitrage and Winning Strategies in the Football

Betting Market'. In particular, we wish to thank Jerry Coakley, Chris Brooks, Hiroyuki Nakata, Norvald Instefjord and George Papachristou. The usual disclaimer applies.

REFERENCES

Avery C, Chevalier J. 1999. Identifying investor sentiment from price paths: the case of football betting. *Journal of Business* **72**: 493–521.

Barberis N, Shleifer A, Vishny R. 1998. A model of investor sentiment. *Journal of Financial Economics* **49**: 307–343.

Bates JM, Granger CWJ. 1969. The combination of forecasts. *Operational Research Quarterly* **20**: 451–468.

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

Cain M, Law D, Peel D. 2001. The incidence of insider trading in betting markets and the Gabriel and Marsden anomaly. *Manchester School* **69**: 197–207.

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

Clements MP, Harvey DI. 2007. Forecast encompassing tests and probability forecasts. Warwick Economics Research Paper Series (TWERPS) 774, Department of Economics, University of Warwick.

Dare WH, MacDonald SS. 1996. A generalized model for testing the home and favorite team advantage in point spread markets. *Journal of Financial Economics* **40**: 295–318.

Dixon MJ, Coles SG. 1997. Modelling Association Football scores and inefficiencies in the football betting market. *Applied Statistics* **46**: 265–280.

Dixon MJ, Pope PF. 2004. The value of statistical forecasts in the UK Association football betting market. *International Journal of Forecasting* **20**: 697–711.

Durham GR, Hertzel MG, Martin JS. 2005. The market impact of trends and sequences in performance: new evidence. *Journal of Finance* **60**: 2551–2569.

Edelman DC, O'Brian NR. 2004. Tote arbitrage and lock opportunities in racetrack betting. *European Journal of Finance* **10**: 370–378.

Froot KA, Dabora EM. 1999. How are stock prices affected by the location of trade? *Journal of Financial Economics* **53**: 189–216.

Gandar JM, Zuber RA, Lamb RP. 2001. The home field advantage revisited: a search for the bias in other sports betting markets. *Journal of Economics and Business* **53**: 439–453.

Gil R, Levitt SD. 2007. Testing the efficiency of markets in the 2002 World Cup. *Journal of Prediction Markets* **1**: 255–270.

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

Golec J, Tamarkin M. 1991. The degree of inefficiency in the football betting market: statistical tests. *Journal of Financial Economics* **30**: 311–313.

Granger CWJ, Ramanathan R. 1984. Improved methods of combining forecasts. *Journal of Forecasting* **3**: 197–204.

Hausch DB, Ziemba WT. 1990. Arbitrage strategies for cross-track betting on major horse races. *Journal of Business* **63**: 61–78.

Hodges SD, Tompkins RG, Ziemba WT. 2003. The favorite/long-shot bias in S&P 500 and FTSE 100 Index futures options: the return to bets and the cost of insurance. In *EFA 2003 Annual Conference*, paper no. 135.

Kallio M, Ziemba WT. 2003. Arbitrage pricing simplified. Sauder School of Business working paper.

Kim SJ, Moshirian F, Wu E. 2006. Evolution of international stock and bond. Market integration: influence of the European Monetary Union. *Journal of Banking and Finance* **30**: 1507–1534.

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

Lane D, Ziemba WT. 2004. Jai Alai arbitrage strategies. *European Journal of Finance* **10**: 353–369.

Levitt SD. 2004. Why are gambling markets organised so differently from financial markets? *Economic Journal* **114**: 223–246.

Mainelli M, Dibb S. 2004. *Betting on the Future: Online Gambling Goes Mainstream Financial*. Centre for the Study of Financial Innovation: New York.
Makropoulou V, Markellos RN. 2007. Optimal price setting in fixed-odds betting markets under information uncertainty. MSL working paper, Athens University of Economics and Business.
Newbold P, Granger CWJ. 1974. Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society, Series A (General)* **137**: 131–165.
Ottaviani M, Sørensen PN. 2005. Parimutuel versus fixed-odds markets. Working paper. Department of Economics, University of Copenhagen: Copenhagen, Denmark.
Paton D, Vaughan Williams L. 2005. Forecasting outcomes in spread betting markets: can bettors use 'quarbs' to beat the book? *Journal of Forecasting* **24**: 139–154.
Pope PF, Peel DA. 1989. Information, prices and efficiency in a fixed odds betting market. *Economica* **56**: 323–341.
Quandt RE. 1986. Betting and equilibrium. *Quarterly Journal of Economics* **101**: 201–208.
Ruhm DL. 2003. Distribution-based formulas are not arbitrage free. *Proceedings of the Casualty Actuarial Society* **90**: 97–129.
Sauer RD. 1998. The economics of wagering markets. *Journal of Economic Literature* **36**: 2021–2064.
Shin HS. 1993. Measuring the incidence of insider trading in a market for state-contingent claims. *Economic Journal* **103**: 1141–1153.
Shleifer A. 2000. *Clarendon Lectures: Inefficient Markets*. Oxford University Press: Oxford.
Smith MA, Paton D, Vaughan Williams L. 2006. Market efficiency in person-to-person betting. *Economica* **73**: 673–689.
Thaler RH, Ziemba WT. 1988. Anomalies: Parimutuel markets: racetracks and lotteries. *Journal of Economic Perspectives* **2**: 161–174.
Vaughan Williams L. 1999. Information efficiency in betting markets: a survey. *Bulletin of Economic Research* **51**: 1–39.
Vaughan Williams L (ed.). 2005. *Information Efficiency in Financial and Betting Markets*. Cambridge University Press: Cambridge, UK.
Vaughan Williams L, Paton D. 1997. Why is there a favourite–longshot bias in British racetrack betting markets? *Economic Journal* **107**: 150–158.
Vaughan Williams L, Paton D. 1998. Why are some favourite–longshot biases positive and others negative? *Applied Economics* **30**: 1505–1510.
Vecer J, Ichiba T, Laudanovic M. 2006. Parallels between betting contracts and credit derivatives. Working paper, Department of Statistics, Columbia University.
Vergin RC, Sosik JJ. 1999. No place like home: an examination of the home field advantage in gambling strategies in NFL football. *Journal of Economics and Business* **51**: 21–31.
Wolfers J, Zitzewitz E. 2004. Prediction markets. *Journal of Economic Perspectives* **18**: 107–126.
Woodland LM, Woodland BM. 1994. Market efficiency and the favorite–longshot bias: the baseball betting market. *Journal of Finance* **49**: 269–279.

Authors' biographies:
Nikolaos Vlastakis is a doctoral candidate in the Department of Management Science and Technology, Athens University of Economics and Business. His research interests include time series and panel econometrics, forecasting and prediction markets.

George Dotsis is a Lecturer at the School of Accounting, Finance and Management, Essex Business School, University of Essex. His research focuses on empirical asset pricing, derivatives valuation and portfolio theory. His work has been published in Applied Economics, Journal of Banking and Finance and Journal of Futures Markets.

Raphael N. Markellos is Senior Lecturer of Quantitative Finance in the Department of Management Science and Technology, Athens University of Economics and Business. His research interests and publications are in the area of applied time series econometrics and finance.

Authors' addresses:
Nikolaos Vlastakis, Department of Management Science and Technology, Athens University of Economics and Business, Office 915, 47A Evelpidon Str. 113 62, Athens, Greece.

DOI: 10.1002/for

George Dotsis, School of Accounting, Finance and Management, Essex Business School, University of Essex, Colchester CO4 3SQ, UK.

Raphael N. Markellos, Department of Management Science and Technology, Athens University of Economics and Business, Office 915, 47A Evelpidon Str. 113 62, Athens, Greece; Visiting Research Fellow, Centre for Research in International Economics and Finance (CIFER), Loughborough University, UK.

DOI: 10.1002/for