The Statistician (2000)
49, *Part* 4, *pp.* 531–539

Focus on Sport

The Kelly criterion and bet comparisons in spread betting

John Haigh

University of Sussex, Brighton, UK

[Received February 2000. Revised June 2000]

Summary. The Kelly strategy, risking a fixed fraction of one's gambling capital each time when faced with a series of comparable favourable bets, is known to be optimal under several criteria. We review this work, interpret it in the context of spread betting and describe its operation with a performance index. Interlocking spread bets on the same sporting event are frequently offered. We suggest ways of investigating which of these bets may be most favourable, and how a gambler might make an overall comparison of the bets from different firms. Examples illustrate these notions.

Keywords: Gambling; Index bets; Kelly criterion; Optimality; Spread bets

1. Introduction

The operation of spread, or index, betting on sports events is described by Jackson (1994) and Haigh (1999). Bets are made on the outcome of a quantity X, such as the number of runs in cricket, the score difference in a rugby game or even the time until the first throw-in in a soccer match. The index betting firm offers an interval (a, b), known as the *spread*. A bettor may choose to *buy* X at unit stake β $(\beta > 0)$, in which case he receives $\beta(X - b)$, or to *sell* X at unit stake α $(\alpha > 0)$, in which case he receives $\alpha(a - X)$. Receiving negative amounts corresponds, of course, to making the appropriate payment. If X is not bounded below, then *any* buy bet could potentially require a payment of more than the bettor has, and similarly, if X is not bounded above, *any* sell bet could lead to an unsustainable loss. To avoid this, the gamblers or the firm may introduce a 'stop loss', by replacing an unbounded Y, such as a score difference between two teams, by the variable

$$X = \begin{cases} A & \text{if } Y \leqslant A, \\ Y & \text{if } A \leqslant Y \leqslant B, \\ B & \text{if } Y \geqslant B, \end{cases}$$

where A and B are finite constants. When stop losses are in place, the bettor can select the unit stake so that his maximum possible loss never exceeds his current fortune. Otherwise, since losses are legally enforceable, a bettor faces possible bankruptcy.

Address for correspondence: John Haigh, Centre for Statistics and Stochastic Modelling, School of Mathematical Sciences, University of Sussex, Falmer, Brighton, BN1 9QH, UK.
E-mail J.Haigh@sussex.ac.uk

 0039–0526/00/49531

Kelly (1956) considered how a gambler, faced with a series of favourable bets at given odds, should apportion his stakes. He showed that, to maximize the long-term growth rate of his capital by staking some proportion x of his current wealth, the optimal choice of x is the size of the advantage. If an outcome at odds of K: 1 has probability p, with $p > 1/(1+K)$, he should select $x = x_{\text{opt}} = \{p(K+1) - 1\}/K$. In practice, acting precisely in accordance with this criterion may be impossible, as the permitted stakes change by discrete amounts, and there are upper and lower bounds on the size of a bet. The strict use of this criterion may ask a gambler to bet on individual outcomes with negative expected gain, and to risk a substantial fraction of his capital, both of which may cause him discomfort. And, in gambles such as betting on horses, the values of the probabilities p are subject to great uncertainty. Despite these reasons why a gambler may not use x_{opt}, calculating its value gives a useful guide.

The optimality of the Kelly criterion has been studied under a range of conditions. We seek to interpret and apply these results in the context of spread betting, and we look to other ways in which a gambler might choose between competing favourable bets.

2. A review

Suppose that a bettor has initial capital V_0, and fortune V_n after n bets. These bets are on the outcomes of a sequence of independent random variables $\{X_k: k \geqslant 1\}$, all distributed as the generic variable X. Write $E(X) = \mu$ and $\text{var}(X) = \sigma^2$, and suppose that the spread (a, b) is such that $\mu > b$, so that a buy bet has positive expectation. (There is a parallel theory based on sell bets when $\mu < a$.) This favourable position may arise either because the spread betting firm has made an error of judgment or because the weight of other gamblers' money on sell bets has caused the firm to seek to attract buy contracts, to reduce its overall risk.

Rebasing the variable, and using a stop loss if necessary, we may assume that 0 is indeed the greatest lower bound for X by requiring that the event $\{X < 0\}$ has zero probability, and that, for any $\delta > 0$, the event $\{X < \delta\}$ has positive probability. The bettor must select the unit stakes, $\{\beta_k: k \geqslant 1\}$, for a sequence of buy bets. The maximum possible loss on the kth bet is $\beta_k b$, so write $\beta_k = u_k V_k / b$, where $0 < u_k \leqslant 1$. Then plainly

$$\frac{V_n}{V_0} = \prod_{k=1}^{n} \left\{ 1 + \frac{(X_k - b)u_k}{b} \right\} \tag{1}$$

and the mean growth rate of the bettor's fortune after n bets is

$$G_n = \frac{1}{n} \ln\left(\frac{V_n}{V_0}\right).$$

The 'Kelly criterion' of maximizing this long-term expected growth rate is not the only option. Samuelson (1971) argued strongly against maximizing the geometric mean of the gambler's fortune, but Algoet and Cover (1988) claimed that decisions based on such a policy would be using a fundamentally important utility function. In the general context of economics, Hirshleifer (1977) asked 'According to what criterion does natural selection select when strategies have uncertain outcomes?'. Blume and Easley (1992, 1993) identified the expected growth rate of wealth share as the 'fitness' criterion that a market selects for. As noted below, the Kelly strategy has the added attraction of leading to the optimal bet under different criteria.

If $u_k = u$ is constant, and $G(u) = E[\ln\{1 + (X - b)u/b\}]$ is finite, the strong law of large numbers shows that $G_n \to G(u)$, almost surely. Finkelstein and Whitley (1981) showed that some essentially unique choice $u = u^*$ will maximize $G(u)$. Plainly, this is the solution of the equation

$$G'(u^*) = E\left\{\frac{X-b}{b+u^*(X-b)}\right\} = 0. \tag{2}$$

They also showed that, if V_n^* is the fortune after n plays using this fixed choice u^*, and V_n corresponds to any other permissible choice of strategy, including allowing $\{u_k\}$ to vary, then using u^* each time is superior in the following senses. (Recall, e.g. Loeve (1963), that $\{W_n\}$ is a *supermartingale* if the conditional expectation of W_{n+1}, given its past history, does not exceed W_n, i.e. the sequence tends to *decrease*. It is a *submartingale* if this conditional expectation is at least W_n, i.e. the tendency is to *increase*.)

(a) V_n/V_n^* is a supermartingale, with $E(V_n/V_n^*) \leqslant 1$, so that V_n/V_n^* converges almost surely to some finite value, and $E\{\lim(V_n/V_n^*)\} \leqslant 1$.
(b) V_n^*/V_n is a submartingale, with $E(V_n^*/V_n) \geqslant 1$, so that V_n^*/V_n converges almost surely to some real number (or to $+\infty$), and $E\{\lim(V_n^*/V_n)\} \geqslant 1$.

These results refine some of Breiman's (1961) work, which also showed that the use of a fixed fraction at each stage, for a sequence of favourable bets at given odds, will minimize the expected number of trials to reach a preassigned fortune—another possible criterion for the bettor.

Ethier and Tavare (1983) considered the behaviour of the value u^* from equation (2) when the difference $\mu - b = \epsilon$ is small. By an expansion of equation (2), they showed that the optimal bet uses

$$u^* \approx b\epsilon/\sigma^2. \tag{3}$$

The corresponding maximum safely achievable growth rate is

$$G(u^*) \approx \epsilon^2/\sigma^2, \tag{4}$$

i.e. the growth rate achieved by using the Kelly criterion is directly proportional to the *square* of the size of the punter's advantage, and inversely proportional to the variance of the quantity under consideration.

Browne (1998) considered betting in continuous time and listed several criteria under which a Kelly strategy of risking a fixed proportion of current wealth is found to be optimal.

Gottlieb (1985) noted that a diffusion approximation arises naturally, especially when σ^2 greatly exceeds ϵ. He used this to find a strategy to minimize the mean time for V_n to reach some preassigned goal with specified probability, under the condition that the intermediate values $\{V_k\colon 1 \leqslant k \leqslant n\}$ never fall below some lower bound. This is of direct interest in spread betting, as some minimum unit stake is always permitted. It turns out that here a strategy other than betting a fixed fraction of current wealth is recommended. Specifically, suppose that $c > 1 > d$, and we seek to minimize the mean time for $V_n/V_0 \geqslant c$, while ensuring that $V_k/V_0 \geqslant d$ for all $0 \leqslant k \leqslant n$, with the probability of reaching the desired goal at least p. When $V_k/V_0 = x$, Gottlieb (1985) showed that, provided that p exceeds some minimal value, the optimal unit stake in the next bet is $(b\epsilon/\sigma^2)(1 + \rho/x)$ where

$$\rho = \frac{c(d-1) - p(d-c)}{p(d-c) - (d-1)}.$$

3. The Kelly criterion in performance index betting

Suppose that rewards $R_1 \geqslant R_2 \geqslant \cdots \geqslant R_n$ are awarded to the n horses in a race, according to their finishing order, and that the spread for horse i is (a_i, b_i). We exclude the 'arbitrage' cases where the spreads have been set so carelessly that it is possible to place bets that have a guarantee of no loss, and the chance of a gain, regardless of the outcome of the race.

A bettor whose current fortune is F will buy horse i at unit stake $u_i F/b_i$ and sell horse i at unit stake $v_i F/(R_1 - a_i)$, for $i = 1, 2, \ldots, n$. For any i, at least one of u_i and v_i will be 0—it would be absurd both to buy and sell the same horse—but bets on several horses may be made. Write

$$w = 1 - \sum_i u_i - \sum_i v_i.$$

The outcomes are correlated, so if all the bets are made simultaneously the index firm need not insist that $w \geqslant 0$ to ensure that the bettor's fortune cannot become negative.

Provided that the mean value of the reward for some horse lies outside that horse's spread interval, a favourable bet will exist. Examples 1 and 2 illustrate what Kelly (1956) found for bets at odds: the bet combination that maximizes the long-term growth rate may use horses on which individual bets are unfavourable.

Consider the special case where $R_1 = R > 0$, $R_2 = R_3 = \ldots = R_n = 0$, i.e. only one horse gains a non-zero reward. The parallel with a set of conventional odds is that horse i is offered at odds of $(R - b_i){:}\, b_i$ to win the race, *and also* at odds of $a_i{:}\,(R - a_i)$ to fail to win. This second set of bets is not always offered by conventional bookmakers. The following algorithm (a justification for which is available on request from the author) identifies an optimal bet (in Kelly's sense). Write $p_i = \Pr(\text{horse } i \text{ wins})$.

Step 1: for any $s \geqslant 0$, write $X_s = \{i\colon Rp_i/b_i > s\}$ and $Y_s = \{i\colon Rp_i/a_i < s\}$.
Step 2: if both X_1 and Y_1 are empty, there is no favourable bet.
Step 3: otherwise, select $u > 0$ and $v > 0$ so that $X_u \cap Y_v = \varnothing$, and with $X_u \cup Y_v$ a proper subset of $\{1, 2, \ldots, n\}$. Calculate

$$t_{uv} = \left\{1 - \sum_{i\in X_u} p_i - \sum_{i\in Y_v} p_i\right\} \Big/ \left\{1 - \sum_{i\in X_u} b_i/R - \sum_{i\in Y_v} a_i/R\right\}.$$

Step 4: check the conditions

(i) if $i \in X_u$, then $Rp_i/b_i > t_{uv}$,
(ii) if $i \in Y_v$, then $Rp_i/a_i < t_{uv}$ and,
(iii) if $i \notin X_u \cup Y_v$, then $Rp_i/a_i \geqslant t_{uv} \geqslant Rp_i/b_i$.

If any of these fail, this choice of $\{X_u, Y_v\}$ does not lead to the optimal bet. But, if all of them hold, calculate

$$G_{uv} = \sum_{i\in X_u} p_i \ln(Rp_i/b_i) + \sum_{i\in Y_v} p_i \ln(Rp_i/a_i) + \ln(t_{uv})\left\{1 - \sum_{i\in X_u\cup Y_v} p_i\right\}.$$

Step 5: repeat step 3 for all possible choices of $\{X_u, Y_v\}$, and identify the choice that gives the maximum value of G_{uv}. For this choice, let $t = t_{uv}$. Then $u_i = p_i - tb_i/R$ and

$$v_i = (R - a_i)\frac{ta_i/R - p_i}{a_i}$$

determine the buy and sell stakes in an optimal bet.

3.1. Example 1

Suppose that we have the position shown in Table 1. Then $X_1 = \{1, 2, 3\}$ and $Y_1 = \{6, 7\}$ are both non-empty, so that to buy any of $\{1, 2, 3\}$ or to sell either of $\{6, 7\}$ is a favourable bet. In step 3 above, $t_{11} = 25/23$, but condition (i) in step 4 fails for horse 3. Taking $X_u = \{1, 2\}$ and $Y_v = \{6, 7\}$ leads to $t_{uv} = 15/14$; here condition 4 holds, so this is a possible optimum and

Table 1. Artificial data, showing possible win probabilities and their respective spreads for a race with seven horses, with one reward $R = 100$

Horse	*Win probability*	*Spread*
1	0.25	(12, 14)
2	0.05	(2, 4)
3	0.20	(17, 19)
4	0.05	(3, 5)
5	0.20	(19, 21)
6	0.15	(25, 27)
7	0.10	(15, 17)

$$G_{uv} = 0.069988239.$$

With the same X_u, let $Y_v = \{5, 6, 7\}$; then $t = 25/23$ (again), and this time condition 4 holds, giving

$$G_{uv} = 0.070045508.$$

Several other choices for $\{X_u, Y_v\}$ satisfy condition 4, but none of them gives a higher value for G_{uv}. The (unique) optimal bet uses $u_1 = 0.097825$, $u_2 = 0.006521$, $v_6 = 0.365218$, $v_7 = 0.357247$ and $v_5 = 0.027803$. Notice that it does *not* include the favourable option to buy horse 3, but it *does* include the unfavourable option to sell horse 5!

Henery (1999) noted a common performance index, where the first four horses are awarded 50, 30, 20 and 10 points. There are several ways of using one parameter for each horse to specify the probabilities for the ordering of the first four horses. When more than two different rewards are given, an iterative process (details of which are available from the author) is needed to find the optimal bet combination. Example 2 uses this process to find the optimal bet when there are just three different rewards.

3.2. Example 2

Suppose that the winner scores 65, the second scores 35 and all others score 0. Let Table 1 give the win probabilities $\{p_i\}$, and the spreads, and use Harville's (1973) formula $p_{ij} = p_i p_j/(1 - p_i)$ for the probabilities for the first two horses. The optimal bet uses the same horses as in example 1, but with $u_1 = 0.1903$, $u_2 = 0.01659$, $v_6 = 0.4026$, $v_7 = 0.3562$, $v_5 = 0.04391$ and optimal growth rate $G = 0.1354$. Although $\Sigma(u_i + v_i) > 1$, even the least favourable outcome does not exhaust the bettor's capital.

4. Comparing spread bets

A gambler will wish to decide which of several apparently favourable spread bets is most attractive. When the advantage is small, approximation (4) could be a basis for choosing between competing bets.

4.1. Example 3

Given a homogeneous Poisson process of rate λ over the time period from 0 to F, spread bets may be offered on both X, the total number of events that occur, and on T, the time until the first event.

(This may be applicable to bets on the scoring of goals in soccer matches, which last for 90 minutes. Dixon and Robinson (1998) found that, although a *non-homogeneous* Poisson model fitted the data better (goals beget goals, and, in addition, more goals occur later in a match), a homogeneous Poisson model is not unreasonable.)

In this model, X has a Poisson distribution with parameter λF, and T has an exponential distribution, truncated at the value F. The mean and variance of T are then $(1-\alpha)/\lambda$ and $(1-\alpha^2)/\lambda^2 - 2F\alpha/\lambda$, where $\alpha = \exp(-\lambda F)$. If the spreads offered are consistent with the firm having underestimated λ, the buy price b for X will be too low, and the sell price s for T will be too high. On the basis of expression (4), buying X is preferable to selling T, if, and only if,

$$(\lambda F - b)\sqrt{(1-\alpha^2 - 2F\alpha\lambda)} > (\lambda s - 1 + \alpha)\sqrt{(\lambda F)}. \tag{5}$$

Spreads of (2.5, 2.8) for X and (35, 38) for T are common. Goal times are recorded as the whole minute in which they occur, so the interval for T is *4* minutes wide. If the mean number of goals expected in the match is $\lambda F = 3$, then the left-hand side of inequality (5) is 0.167 and the right-hand side is 0.317. Selling T is preferred to buying X here.

A gambler may seek to judge whether a collection of spread bets is likely to contain *some* favourable bet, according to his model. The firms construct the spreads by using both their own assessments of the chances of the different outcomes and their expectation of gamblers' behaviour. We suggest two possible considerations, based on different ways of judging how much risk the index firm has taken. Firstly, suppose that X is a random variable with a finite mean, whose distribution is governed by a parameter θ. The *firm's interval* for the spread (a, b) is defined as $B_X = \{\theta\colon a < E(X) < b\} = (A, B)$, say.

The interpretation is as follows. Both buy and sell bets will favour the firm when $a < E(X) < b$. Regarding θ as a random variable with distribution function $H(\theta)$, $H(B) - H(A)$ is then the probability that both buy and sell bets favour the firms, and so measures their unwillingness to be exposed to a risk of loss. If $H(\theta)$ is unknown, we can use the *widths* of the respective firm's intervals as proxies to compare spread bets offered on two different variables.

4.2. Example 4

In high scoring games such as rugby or American football, the number of points scored by a team might be modelled by some normal distribution. Spreads are offered on N, the total number of points scored in a match, and on S, the points superiority of one team. The widths of these two spreads are generally equal.

The values of N and S depend on the scoring rates of each team, which might be expected to have a positive correlation (e.g. both face the same weather and/or pitch conditions; if one team takes risks to score more points, that usually gives point scoring opportunities to the other, etc.). In these circumstances, N would have a larger variance than S, so the probability associated with an interval of given width around the mean value of N would be *smaller* than that for an interval of the same width for S. A punter is more likely to find a favourable bet on N. Conversely, if we expect a negative correlation between the teams' scores, S is more likely to be a source of a favourable bet.

4.3. Example 3 (continued)

Taking the homogeneous Poisson process as a model for goals in soccer, suppose that the spread for X is $(a-\delta, a+\delta)$ and that for T is $(b-\epsilon, b+\epsilon)$, and that the spread $(c-\phi, c+\phi)$ is offered for U, the time of the *second* goal. Note that $E(U) = 2(1-\alpha)/\lambda - \alpha F$.

For the parameter λ, the width of B_X is plainly $2\delta/F$. Define λ_0 by $b\lambda_0 = 1 - \alpha_0$, where $\alpha_0 = \exp(-F\lambda_0)$, so that the width of B_T is approximately $2\epsilon\lambda_0/(b - \alpha_0 F)$ when ϵ is small. For U, similarly define λ_1 by $(c + F\alpha_1)\lambda_1 = 2(1 - \alpha_1)$, where $\alpha_1 = \exp(-F\lambda_1)$. The width of B_U is approximately $2\phi\lambda_1/(c - \alpha_1 F - \alpha_1\lambda_1 F^2)$.

To compare the three bets, we use (2.5, 2.8) for X, (35, 38) for T and (61, 64) for U as typical of the spreads found for soccer matches. Then the widths of B_X, B_T and B_U are 0.00333, 0.00377 and 0.00320 respectively. The spread on B_T appears to be most conservative, whereas gamblers can expect best value from bets on U.

The natural estimate of λ, based on the spread (2.5, 2.8) for X, is $2.65/90 = 0.02944$, which exceeds the estimates $\lambda_0 = 0.02480$ and $\lambda_1 = 0.02507$. This is consistent with firms actually using a non-homogeneous model, where the overall scoring rate is expected to exceed the initial rate. However, we would expect the *comparison* of the widths of the firm's intervals to be fairly robust to the difference between a homogeneous and a non-homogeneous model.

4.4. Example 5: soccer, Barnsley versus Stockport County, September 10th, 1999 (spreads from William Hill)

We use the following spreads:

(a) goals superiority of Barnsley over Stockport, (0.7, 1.0);
(b) total goals in the game, (2.7, 3.0);
(c) total goal minutes, (135, 145) (i.e. sum the times at which all goals are scored);
(d) minutes that Barnsley are in the lead during the match, (33, 36);
(e) minutes that Stockport are in the lead during the match, (12, 15).

(Barnsley scored at 16 and 44 minutes, Stockport at 67, so the respective outcomes of the five betting opportunities were 1.0, 3.0, 127, 74 and 0.)

4.4.1. Analysis

The outcomes of (b) and (c) are strongly correlated. The ratio of the midpoints of the two spreads, $140/2.85 = 49.12 = m$, is an estimate of the average time, in minutes, at which a randomly chosen goal is scored. The firm's interval for the scoring rate in (c) corresponds to an expected number of goals in the range $(134.5/m, 145.5/m) = (2.74, 2.96)$, which is contained within the spread for (b). A punter, buy or sell, can expect better value from bets on (c) than on (b).

We have used simulation to explore the correspondence between bets on {(a), (b)} and bets on {(d), (e)}, taking independent homogeneous Poisson processes as models for both teams' goal scoring. For a given average number of goals, there was an excellent straight line fit between Time, the mean time that a team is in the lead, and Sup, the expected superiority of that team over the other. With an average of 2.85 goals per game, and Sup allowed to vary from -1.0 to 1.6, the simulated data led to

$$\text{Time} = 25.05 + 12.61\,\text{Sup}. \tag{6}$$

(Simulations for averages of 2.05 and 2.45 goals per game also gave excellent straight line fits, with only slightly different coefficients.)

Equation (6) translates the width 0.3 of the spread in (a) to a corresponding width of $12.61 \times 0.3 = 3.78$ for the expected spread in (d) or (e), just less than the actual width of 4 minutes. Further, equation (6) also enables us to translate the superiority interval (0.7, 1.0) for Barnsley's goals to the interval (33.88, 37.66) for the time that Barnsley are estimated to be in the lead, which corresponds well to the spread offered in (d). And, since Stockport's expected goal superiority is the negative of Barnsley's, equation (6) leads to (12.44, 16.22) as the suggested spread for the time that Stockport would be in the lead.

Not only are the locations of the spreads on the different bets offered seen to be consistent with each other, the index firm is also taking similar risks.

The notion of a firm's interval easily extends to more than one parameter.

4.5. Example 6: cricket, South Africa versus England, fifth test match, January 2000

In a drawn match, each team scores 10 points; otherwise the winner scores 25 points and the loser scores 0. IG Index offered initial spreads of (16, 17.5) for South Africa and (6, 7.5) for England.

Writing $p = \Pr(\text{South Africa win})$ and $q = \Pr(\text{England win})$, so that $1 - p - q$ is the chance of a draw, the mean scores to each team are $10 + 15p - 10q$ and $10 + 15q - 10p$. Let R be the region (a parallelogram) within the set $S = \{(p, q)\colon p \geqslant 0, q \geqslant 0, p + q \leqslant 1\}$ defined by $6 < 15p - 10q < 7.5$ and $-4 < 15q - 10p < -2.5$. So long as the point (p, q) is within R, all bets are unfavourable, so R would be the *firm's region*.

A second method of assessing the overall risk taken by the firm comes from an analogy with odds betting. Suppose that the odds offered against outcome k are α_k: 1, and

$$\sum_k \frac{1}{1+\alpha_k} = 1 + m,$$

where m ($m > 0$) is known as the *over-round*. Broadly, the smaller the over-round, the more likely it is that some horse is offered at favourable odds. Haigh (1999) considered spread bets on the two-horse race, where the winner scores R and the loser scores 0. Spreads (a, b) and (c, d) are on offer, with $d - c = b - a$; he noted that this bet is completely equivalent to an odds bet with an over-round of $(b - a)/R$.

In some two-player contests, there are rewards $R = R_1 > R_2 \geqslant \ldots \geqslant R_{n-1} > R_n = 0$, paired so that, if player A scores R_i, then B scores S_i. It can be shown (the details are available on request from the author) that if the spreads are (a, b) and (c, d) with $b - a = d - c$, and some (un-demanding) conditions hold, there is a set of odds, with over-round $(b - a)/R$, from which a bet can be constructed that is *identical* in its outcome with one of the four possible spread bets. All such sets of odds have this same over-round. This allows us to use the respective values of $(b - a)/R$ to compare the overall risks taken by firms with different reward structures for the same contest, and to judge where a favourable bet is most likely to be found. (We do not claim that the spread bet is *equivalent* to an odds bet with over-round $(b - a)/R$. There will be a set of odds bets that match each spread bet, but not normally a set of spread bets that match each odds bet, when $n > 2$.)

4.6. Example 6 (continued)

Here $n = 4$, $R_1 = 25$ and $R_2 = R_3 = 10$. Consider odds of 21:29 against a win for South Africa, 41:9 against a win for England, and 7:3 against a draw, which have an over-round of 6%, the same as $(b - a)/R = (17.5 - 16)/25$. The four spread bets at unit stake offered by IG Index give identical outcomes to the corresponding odds bets shown, whatever the match result:

buy South Africa, *or* bet 14.5 on South Africa to win, bet 3 on a draw;
sell South Africa, *or* bet 4.5 on each of England to win or a draw;
buy England, *or* bet 4.5 on England to win, 3 on a draw;
sell England, *or* bet 14.5 on South Africa to win, 4.5 on the draw.

5. Discussion

There is no unique optimal way of taking advantage of favourable bets. A well-known case is when an outcome with probability $p > 0.5$ will pay out at even money: the policy of betting one's entire capital each time maximizes the expected fortune at every stage, but leads to certain ruin. The Kelly criterion of aiming to maximize the long-term capital growth rate has attractions, but it may turn out to be impractical or to indicate bets that are larger than a gambler is willing to make. We have described how this criterion applies in various forms of spread betting and noted that it has other optimality properties. MacLean *et al.* (1992) have looked at the benefits of using 'fractional Kelly' strategies, noting that one can trade off a small decrease in growth rate against, say, an increased chance that one's fortune is doubled before it is halved.

We have also looked at ways of inferring that favourable spread bets might be available, and of comparing correlated bets on the same event. Although the spread betting firms monitor each other's prices, the intersection of the different firms' intervals (or regions) may be sufficiently small for any gambler to locate some favourable bet. In some circumstances, an analogy with the over-round in odds betting enables us to compare the overall risks between different firms.

Jackson (1994) reported that bookmakers have been reluctant to use formal model building when setting their odds. When interlocking spread bets are offered, there is a clear danger that gamblers may have arbitrage opportunities. However, our examples indicate that, whether or not formal models are used to help to set spreads, the bets offered are broadly consistent with the formal models put forward in this paper.

Acknowledgements

I am grateful to Leighton Vaughan-Williams for the data of example 6, and for asking whether it might have an analogy with the conventional bookmakers' over-round. I thank David Firth for drawing my attention to the economics literature in this area, and two helpful referees whose comments have led to improvements in the presentation.

References

Algoet, P. H. and Cover, T. M. (1988) Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *Ann. Probab.*, **16**, 876–898.
Blume, L. E. and Easley, D. (1992) Evolution and market behaviour. *J. Econ. Theory*, **58**, 9–40.
———(1993) Economic natural selection. *Econ. Lett.*, **42**, 281–289.
Breiman, L. (1961) Optimal gambling systems for favourable games. In *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, vol. 1, pp. 65–78. Berkeley: University of California Press.
Browne, S. (1998) The return on investment from proportional portfolio strategies. *Adv. Appl. Probab.*, **30**, 216–238.
Dixon, M. J. and Robinson, M. E. (1998) A birth process model for association football matches. *Statistician*, **47**, 523–538.
Ethier, S. N. and Tavare, S. (1983) The proportional bettor's return on investment. *J. Appl. Probab.*, **20**, 563–573.
Finkelstein, M. and Whitley, R. (1981) Optimal strategies for repeated games. *Adv. Appl. Probab.*, **13**, 415–428.
Gottlieb, G. (1985) An optimal betting strategy for repeated games. *J. Appl. Probab.*, **22**, 787–795.
Haigh, J. (1999) (Performance) index betting and fixed odds. *Statistician*, **48**, 425–434.
Harville, D. A. (1973) Assigning probabilities to the outcomes of multiple entry competitions. *J. Am. Statist. Ass.*, **68**, 312–316.
Henery, R. J. (1999) Measures of over-round in performance index betting. *Statistician*, **48**, 435–439.
Hirshleifer, J. (1977) Economics from a biological viewpoint. *J. Law Econ.*, **20**, 1–52.
Jackson, D. A. (1994) Index betting on sports. *Statistician*, **43**, 309–315.
Kelly, Jr, J. L. (1956) A new interpretation of information rate. *Bell Syst. Tech. J.*, **35**, 917–926.
Loeve, M. (1963) *Probability Theory*, 3rd edn. Princeton: Van Nostrand.
MacLean, L. C., Ziemba, W. T. and Blazenko, G. (1992) Growth versus security in dynamic investment analysis. *Mangmnt Sci.*, **38**, 1562–1585.
Samuelson, P. A. (1971) The "fallacy" of maximising the geometric mean in long sequences of investing or gambling. *Proc. Natn. Acad. Sci. USA*, **68**, 2493–2496.