IMA Journal of Management Mathematics Advance Access published November 5, 2008

IMA Journal of Management Mathematics Page 1 of 18
doi:10.1093/imaman/dpn028

Markowitz portfolio theory for soccer spread betting

ALISTAIR D. FITT†
School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK

[Received on 16 August 2007; accepted on 30 July 2008]

Soccer spread betting is analysed using standard probabilistic methods assuming that goals are scored in a match according to Poisson distributions with constant means. A number of different possible forms of ‘edge’ (betting advantage) is identified. It is shown how the centre spreads of the more common bets in the ‘bet universe’ may be calculated. A more general question is then addressed, namely, how a punter should invest if they take a view that the online bookmakers have fixed the goal means incorrectly or some other edge is in their favour. It is shown that a Markowitz portfolio theory framework may be set up in such cases. This leads to the definitions of an ‘efficient betting frontier’ and an ‘optimal bet portfolio’. Examples are used throughout to illustrate the theory that is developed.

Keywords: spread betting; portfolio theory; soccer betting; arbitrage; sports betting.

1. Soccer spread betting

Sports spread betting has gained a huge following in the UK over the past 5–10 years. This popularity has been mirrored in other parts of the world, with the exception of the USA (which has essentially outlawed such betting for rather obscure ‘moral’ reasons and regularly arrests and imprisons those who run online betting sites) and countries where religious beliefs deem that gambling is not permitted. Soccer has proved particularly attractive to spread betters. Financially, a spread bet replicates a forward or a future contract and, in line with such contracts, is traded via an ‘exchange’ which is normally run by an online bookmaker (OB). A punter wishing to bet deposits a sum of money with their chosen OB, thereby opening an account and allowing them to bet using the internet. All further transactions are then carried out online. In all the examples given below, the OB used was one of the most popular internet spread betting companies (though for obvious reasons we shall not identify exactly which one).

1.1 *The mechanics of spread betting*

The mechanics of spread betting are simple: for each match, bets are placed on the outcome of a range of ‘indices’ offered by the OB. The index may consist of a simple quantity, such as the total number of goals scored in a match, or be much more involved and depend on esoteric combinations of goal times, corners and player bookings, etc. For each index, however, the betting mechanics are identical: the OB quotes a ‘spread’ $[B, T]$ and the punter has a choice (at a unit stake, say) of either ‘buying the spread at T’ or ‘selling the spread at B’. At the conclusion of the match, the final value S of the index is tallied, and the punter then receives an amount $S - T$ if they bought the spread and $B - S$ if they sold the spread. Receiving a negative amount (possible if the punter took an incorrect view) corresponds, of course, to paying a sum to the OB, thereby explaining why spread betting is normally only conducted via

†Email: a.d.fitt@maths.soton.ac.uk

an established OB account. Note that the size $(T_i - B_i)$ of the spread of the ith index simply determines the OBs profit margin: too small a spread and the OB will not cover its overheads, too large a spread and punters will no doubt seek better value elsewhere. For this reason, we henceforth concentrate our attention mainly on the centre of the spread, which will be denoted by $C = (T - B)/2$.

1.2 *Assumptions*

For the purposes of this study, we shall assume

(a) that the matches being studied are league matches (e.g. the English premiership) rather than international matches (which tend to have different properties since many are mismatches, most involve no ‘home’ team and the goals scored by the two sides tend to be negatively correlated);

(b) that goals for both sides are scored according to Poisson processes with constant means μ_{H} and μ_{A} for the home and away sides, respectively. The probability that the number of home goals N_{H} is equal to k is given by

$$\Pr(N_{\mathrm{H}} = k) = \frac{\mathrm{e}^{-\mu_{\mathrm{H}}} {\mu_{\mathrm{H}}}^k}{k!},$$

with a similar expression for the number of away goals. There has been much discussion in the literature (see, e.g. Dyte & Clarke, 2000; Greenhough *et al.*, 2002) concerning whether or not a Poisson model is appropriate: though this is a very interesting topic, space does not permit a further continuation of the debate in this study. Note, however, that if one’s sole motive for valuing such bets is to be able to predict the OB’s spreads, then to some extent it may be argued that it is more important to know what view the OB’s take rather than how goals are really scored. As we shall see below, a Poisson model replicates the OB’s spreads with a great deal of accuracy;

(c) that we will invest only in bets before the match begins (spread bets that are traded as the match progresses are also of interest; see Fitt *et al.*, 2006). Note also that some online sites also now offer spread betting exchanges where the odds are set by the punters themselves and this adds yet another potential complication;

(d) that, in line with the rules set by virtually all OBs, each half of the match lasts for 45 min only and both extra time and penalty shoot-outs are completely ignored. We will also consider only the ‘universe’ of spread bets (essentially all bets that involve only goals) that is detailed in the appendix. For reasons of brevity, we shall also ignore all issues of combination bets that involve ‘place bets’ (see, e.g. Haigh, 1999, 2000) and deal exclusively with spread bets.

2. Basic bet valuation

Our primary aims in this study are (a) to identify circumstances where a punter may enjoy an edge over the OB and (b) if such favourable circumstances do indeed exist, to indicate how the betting strategy employed may be optimized. We shall first distinguish between three different types of edge, which we shall label ‘arbitrage edge (AE)’, ‘online bookmaker edge (OBE)’ and ‘punter edge (PE)’.

The phenomenon of AE occurs only very rarely, and is sufficiently unlikely for us to discount completely in this study. We define AE as the circumstance where it is possible to make a guaranteed profit that is completely independent of the match result. AE is only likely to occur if an OB makes a catastrophic error in the quoted spreads or possibly if different OBs take widely different views of the outcome of a match: in any case, the market normally corrects any instances of AE extremely rapidly.

AE may also be possible if the extra complications of other OBs win and place fixed odds bets and/or the existence of spread betting exchanges are taken into consideration.

We define OBE as the circumstance where one or more of the calculated spreads (based on the OB's inferred Poisson goal means) lie outside the OB's quoted spread. When this occurs, the relevant bets can be considered to be 'positive expectation games'. Though such events are infrequent, they assuredly do happen (e.g. for Wigan versus Manchester United, UK premiership, 14 October 2006, the 'theoretical' bet 26 (Cs_{00}) centre spread value was 10.53, but the quoted spread was [11, 14]). One explanation for the existence of OBEs is the desire of OBs to 'balance their books'. If large amounts of money are bet on one side of the spread (as sometimes happens, e.g. when England are playing an important match), it may be in the OB's interest to quote spreads that are not 'theoretically' correct so as to hedge themselves against a large loss.

Finally, we define PE as the circumstance where one or more of the calculated spreads (based on the 'punter's' estimated Poisson goal means) lies outside the OB's quoted spread. This of course corresponds to the punter 'taking a view' which may prove to be correct or incorrect. The ubiquity of PE opportunities is, of course, largely a matter of opinion that can only be confirmed *post facto*, but it seems clear that possibilities for PE may exist if the OB uses a team's 'standard' goal performance to infer spreads when, in reality, special circumstance apply (e.g. when many of the team's star players are injured). For an example of this, see Section 5.

Our ultimate aim is, of course, to win money. Whether or not it is ultimately possible to make worthwhile profits from football spread betting ultimately depends to some extent on the efficiency of the football betting market. The efficiency of USA National Football League (NFL) sports betting has been previously considered in both Gray & Gray (1997) and Tassoni (1996). As far as soccer spread betting is concerned, however, the question of the general efficiency of markets seems to have received little attention.

3. The calculation of centre spreads

In order to be able to identify any of the favourable betting opportunities described above (and to eventually interpret portfolio theory for soccer spread bets), it is essential to be able to calculate the 'fair' value of each of the bets. In this section, we show how this may be done for each of the bets in our bet universe.

3.1 *Detailed bet valuation*

We now proceed to value the centre spread for each of the bets that are described in the appendix. Before dealing with each bet in turn, it is helpful to establish a few subsidiary results. First, the probabilities HW, DR and AW of a home win, draw and away win, respectively, are given by

$$\mathrm{HW} = \sum_{k=1}^{\infty}\sum_{j=0}^{k-1} \frac{\phi^2 {\mu_{\mathrm{A}}}^j {\mu_{\mathrm{H}}}^k}{j!k!} = \sum_{k=1}^{\infty} \frac{\mathrm{e}^{-\mu_{\mathrm{H}}} {\mu_{\mathrm{H}}}^k \Gamma(k, \mu_{\mathrm{A}})}{k(\Gamma(k))^2}, \tag{3.1}$$

$$\mathrm{DR} = \sum_{k=0}^{\infty} \frac{\phi^2 (\mu_{\mathrm{H}}\mu_{\mathrm{A}})^k}{(k!)^2} = \phi^2 I_0(2\theta), \tag{3.2}$$

$$\mathrm{AW} = \sum_{k=1}^{\infty}\sum_{j=0}^{k-1} \frac{\phi^2 {\mu_{\mathrm{H}}}^j {\mu_{\mathrm{A}}}^k}{j!k!} = \sum_{k=1}^{\infty} \frac{\mathrm{e}^{-\mu_{\mathrm{A}}} {\mu_{\mathrm{A}}}^k \Gamma(k, \mu_{\mathrm{H}})}{k(\Gamma(k))^2}, \tag{3.3}$$

where $\Gamma(k, x)$ denotes the incomplete Gamma function, defined for real $k > 0$ by

$$\Gamma(k, x) = \int_x^\infty e^{-t} t^{k-1}\,dt,$$

I_0 is the modified Bessel function of the first kind, and we have defined, for convenience, $\theta = \sqrt{\mu_H \mu_A}$, $\alpha = \sqrt{\mu_H/\mu_A}$ and $\phi = \exp(-\mu_H/2 - \mu_A/2)$ (note that the three probabilities (3.1)–(3.3) are particularly useful for comparison with fixed odds bookmakers).

We also note that the probability that, in any given half of the match, the goal difference (home goals–away goals) is given by $N \geqslant 0$ is

$$P(H - A = N) = \sum_{k=0}^\infty \frac{\phi\left(\frac{\mu_H}{2}\right)^{k+N}\left(\frac{\mu_A}{2}\right)^k}{k!(k+N)!} = \phi\alpha^N I_N(\theta). \qquad (3.4)$$

The centre spreads C_i may now be determined. Many of the cases below are calculated by relatively simple Poisson enumeration: we give more detailed explanations for some of the more complicated centre spreads. We have

$$C_1 = \mu_H - \mu_A, \quad C_2 = \mu_H + \mu_A = \mu, \quad C_3 = 25\text{HW} + 10\text{DR}, \quad C_4 = 25\text{AW} + 10\text{DR},$$

$$C_5 = e^{-\mu}\left(10\mu + \frac{20\mu^2}{2!} + \frac{33\mu^3}{3!} + \frac{50\mu^4}{4!} + \frac{70\mu^5}{5!} + 100\sum_{k=6}^\infty \frac{\mu^k}{k!}\right)$$

$$= e^{-\mu}\left(100\,e^\mu - 100 - 90\mu - 40\mu^2 - \frac{67\mu^3}{6} - \frac{25\mu^4}{12} - \frac{\mu^5}{4}\right),$$

$$C_6 = 10\sum_{k=1}^\infty \frac{{\mu_H}^k {\mu_A}^{k-1}\phi^2}{k!(k-1)!} + 25\sum_{j=2}^\infty\sum_{k=j}^\infty \frac{{\mu_H}^k {\mu_A}^{k-j}\phi^2}{k!(k-j)!} = 10\alpha\phi^2 I_1(2\theta) + 25\sum_{j=2}^\infty \phi^2\alpha^j I_j(2\theta),$$

$$C_N = 10\alpha^{N-5}\phi^2 I_{N-5}(2\theta) + 25\sum_{j=N-4}^\infty \phi^2\alpha^j I_j(2\theta) \quad (N = 7, 8, 9, 10),$$

$$C_N = 10\alpha^{10-N}\phi^2 I_{N-10}(2\theta) + 25\sum_{j=N-9}^\infty \phi^2\alpha^{-j} I_j(2\theta) \quad (N = 11, 12, 13, 14, 15),$$

$$C_{16} = \mu_H, \quad C_{17} = \mu_A.$$

The bets where a given team scores a certain number of goals and wins (which are indices that are calculated at half and full time) are easiest to value by considering the state of the match at both half and full time. If we define

$$G_{j,k,p,q} = \frac{\phi^2\left(\frac{\mu_H}{2}\right)^j\left(\frac{\mu_A}{2}\right)^k\left(\frac{\mu_H}{2}\right)^p\left(\frac{\mu_A}{2}\right)^q}{j!k!p!q!},$$

then it is relatively easy to show that the expected value of the index bet ‘home team scores (exactly) N and wins’ is given by

$$75\sum_{k=0}^{N-1}\sum_{q=0}^{N-k-1} G_{N,k,0,q} + 50\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\sum_{q=0}^{N-k-1} G_{j,k,N-j,q}$$

$$+25\frac{\phi\left(\frac{\mu_{\mathrm{H}}}{2}\right)^N}{N!}\left(\sum_{k=0}^{N-1}\frac{\left(\frac{\mu_{\mathrm{H}}}{2}\right)^k}{k!}\left(1-\sum_{q=0}^{N-k-1}\frac{\phi}{q!}\left(\frac{\mu_{\mathrm{A}}}{2}\right)^q\right)\right).$$

Simplification occurs for the small values of N that are involved in bets 18, 19 and 20, giving

$$C_{18}=\frac{25\phi\mu_{\mathrm{H}}}{2}(1+4\phi),\quad C_{19}=\frac{25\phi{\mu_{\mathrm{H}}}^2}{16}(2+16\phi+16\phi\mu_{\mathrm{A}}+\mu_{\mathrm{A}}),$$

$$C_{20}=\frac{25\phi{\mu_{\mathrm{H}}}^3}{384}(8+128\phi+128\phi\mu_{\mathrm{A}}+4\mu_{\mathrm{A}}+64\phi{\mu_{\mathrm{A}}}^2+{\mu_{\mathrm{A}}}^2).$$

When the bet involves all values of N that exceed 3, matters are complicated somewhat, but enumeration eventually gives

$$C_{21}=75\sum_{j=4}^{\infty}\sum_{k=0}^{j-1}\sum_{p=0}^{\infty}\sum_{q=0}^{j+p-k-1} G_{jkpq}+50\sum_{j=4}^{\infty}\sum_{k=j}^{\infty}\sum_{p=0}^{\infty}\sum_{q=0}^{j+p-k-1} G_{jkpq}$$

$$+50\sum_{j=0}^{3}\sum_{k=0}^{\infty}\sum_{p=4-j}^{\infty}\sum_{q=0}^{j+p-k-1} G_{jkpq}+25\sum_{j=4}^{\infty}\sum_{k=0}^{j-1}\sum_{p=0}^{\infty}\sum_{q=j+p-k}^{\infty} G_{jkpq}.$$

Though some simplifications of these series are possible, the terms in the expression for C_{21} become very small for summation indices greater than about 7 and may essentially be ignored.

The ‘away team scores N and wins’ bets may now be valued by interchanging μ_{H} and μ_{A} in bets 18–21, giving

$$C_{22}=\frac{25\phi\mu_{\mathrm{A}}}{2}(1+4\phi),\quad C_{23}=\frac{25\phi{\mu_{\mathrm{A}}}^2}{16}(2+16\phi+16\phi\mu_{\mathrm{H}}+\mu_{\mathrm{H}})$$

$$C_{24}=\frac{25\phi{\mu_{\mathrm{A}}}^3}{384}(8+128\phi+128\phi\mu_{\mathrm{H}}+4\mu_{\mathrm{H}}+64\phi{\mu_{\mathrm{H}}}^2+{\mu_{\mathrm{H}}}^2).$$

Finally, defining

$$H_{j,k,p,q}=\frac{\phi^2\left(\frac{\mu_{\mathrm{A}}}{2}\right)^j\left(\frac{\mu_{\mathrm{H}}}{2}\right)^k\left(\frac{\mu_{\mathrm{A}}}{2}\right)^p\left(\frac{\mu_{\mathrm{H}}}{2}\right)^q}{j!k!p!q!},$$

we have

$$C_{25} = 75\sum_{j=4}^{\infty}\sum_{k=0}^{j-1}\sum_{p=0}^{\infty}\sum_{q=0}^{j+p-k-1} H_{jkpq} + 50\sum_{j=4}^{\infty}\sum_{k=j}^{\infty}\sum_{p=0}^{\infty}\sum_{q=0}^{j+p-k-1} H_{jkpq} + 50\sum_{j=0}^{3}\sum_{k=0}^{\infty}\sum_{p=4-j}^{\infty}\sum_{q=0}^{j+p-k-1} H_{jkpq}$$
$$+25\sum_{j=4}^{\infty}\sum_{k=0}^{j-1}\sum_{p=0}^{\infty}\sum_{q=j+p-k}^{\infty} H_{jkpq}.$$

The Cs 'draw' bets (bets 26–28) are most easily valued using simple Poisson enumeration by considering the result in each half of the match. We find that for each 'individual' Cs draw bet, the expected value of the relevant index is given by

$$Cs_{NN} = 75\frac{\phi^2(\mu_{\mathrm{H}}\mu_{\mathrm{A}})^N}{4^N(N!)^2} + 25\frac{\phi(\mu_{\mathrm{H}}\mu_{\mathrm{A}})^N}{4^N(N!)^2}(1-\phi)$$
$$+50\left[-\frac{\phi^2(\mu_{\mathrm{H}}\mu_{\mathrm{A}})^N}{4^N(N!)^2} + \sum_{k=0}^{N}\sum_{j=0}^{N}\frac{\phi^2(\mu_{\mathrm{H}}\mu_{\mathrm{A}})^N}{4^N j!k!(N-j)!(N-k)!}\right]$$
$$= \frac{25\phi\theta^{2N}(4^{-N}+2\phi)}{(N!)^2},$$

thus giving

$$C_{26} = 25\phi(1+2\phi), \quad C_{27} = \frac{25}{4}\theta^2\phi(1+8\phi).$$

Matters are complicated for the Cs_{22} bet by the fact that this bet refers to *all* draws involving four or more goals. Since, e.g. a 2–2 half-time draw can become a 3–3, 4–4, etc., scoreline by full time, the bet cannot be valued simply by summing separate Cs_{NN} bets. Instead, it is simplest to proceed by enumerating each match that gives rise to index values of 75, 25 and 50. This gives

$$Cs_{22} = 75\phi^2\sum_{k=2}^{\infty}\sum_{j=0}^{\infty}\left(\frac{\theta^2}{4}\right)^{j+k}\frac{1}{(j!)^2(k!)^2} + 25\phi\left(\sum_{k=2}^{\infty}\left(\frac{\theta^2}{4}\right)^k\frac{1}{k!^2}\right)\left(1-\sum_{k=0}^{\infty}\left(\frac{\theta^2}{4}\right)^k\frac{\phi}{k!^2}\right)$$
$$+50\phi^2\sum_{k=2}^{\infty}\left(\frac{\theta^2}{4}\right)^k\frac{1}{k!^2} + 50\phi^2\frac{\theta^2}{4}\sum_{k=1}^{\infty}\left(\frac{\theta^2}{4}\right)^k\frac{1}{k!^2} + 50\phi^2\sum_{k=1}^{\infty}\left(\frac{\theta^2}{4}\right)^{k+1}\frac{1}{k!(k+1)!}$$
$$+50\sum_{k=2}^{\infty}\sum_{j=0}^{k-1}\sum_{i=0}^{\infty}G_{kjii+k-j} + 50\sum_{j=2}^{\infty}\sum_{k=0}^{j-1}\sum_{i=0}^{\infty}G_{kji+j-ki}$$

and so, after simplification, we find that

$$C_{28} = \frac{25\phi}{4}(8\phi I_0^2(\theta) + 4I_0(\theta) - 4 - \theta^2 - 8\phi - 8\phi\theta^2 + 8\phi\theta I_1(\theta)) + 100\phi^2\sum_{k=2}^{\infty}\sum_{j=0}^{k-1}\left(\frac{\theta^2}{4}\right)^{\frac{k+j}{2}}\frac{I_{k-j}(\theta)}{k!j!}.$$

To value the 'double' results, it is easiest to deal only with the goal difference in each half of the match, using (3.4). Enumeration proceeds by considering first-half/second-half goal differences that lead to each possible index outcome of 0, 10, 20, 25, 35 or 50 points. This gives

$$
\begin{aligned}
C_{29} = {} & 10\phi^2 \sum_{k=1}^{\infty} I_k(\theta)(\alpha^{-k} I_0(\theta) + I_k(\theta)) + 20\phi^2 I_0^2(\theta) \\
& + 25\phi^2 \sum_{k=1}^{\infty} \sum_{j=k+1}^{\infty} (\alpha^{k-j} + \alpha^{j-k}) I_k(\theta) I_j(\theta) + 35\phi^2 \sum_{k=1}^{\infty} (\alpha^k I_0(\theta) I_k(\theta) + I_k^2(\theta)) \\
& + 50\phi^2 \sum_{k=1}^{\infty} \left[\sum_{j=1}^{\infty} \alpha^{k+j} I_k(\theta) I_j(\theta) + \sum_{j=0}^{k-1} \alpha^{k-j} I_k(\theta) I_j(\theta) \right], \\
C_{30} = {} & 10\phi^2 \sum_{k=1}^{\infty} I_k(\theta)(\alpha^{k} I_0(\theta) + I_k(\theta)) + 20\phi^2 I_0^2(\theta) \\
& + 25\phi^2 \sum_{k=1}^{\infty} \sum_{j=k+1}^{\infty} (\alpha^{j-k} + \alpha^{k-j}) I_k(\theta) I_j(\theta) + 35\phi^2 \sum_{k=1}^{\infty} I_k(\theta)(\alpha^{-k} I_0(\theta) + I_k(\theta)) \\
& + 50\phi^2 \sum_{k=1}^{\infty} \left[\sum_{j=1}^{\infty} \alpha^{-k-j} I_k(\theta) I_j(\theta) + \sum_{j=0}^{k-1} \alpha^{j-k} I_k(\theta) I_j(\theta) \right].
\end{aligned}
$$

We now proceed to value the bets involving goal times. Note that, for simplicity, we shall assume in each of the valuations below that the match lasts for a total time of one unit (see Section 3.3 for details of how these prediction should be transformed into units of minutes).

To value the total goal minutes bet, define X as the sum of the goal times during some interval (a, b) and consider $E(X \mid N(a, b) = n)$, the expected value of the total goal times given that exactly n goals were scored during (a, b). If n is zero, then X is zero. However, if $n = 1$, then by Poisson equipartition we would expect the goal to occur at the midpoint of the time interval, so that $E(X \mid N(a, b) = 1) = (a + b)/2$. By similar reasoning, we have

$$E(X \mid N(a, b) = n) = \sum_{k=1}^{n} \left(a + \frac{k(b-a)}{n+1} \right) = \frac{n(a+b)}{2}.$$

Thus, with $a = 0$ and $b = 1$,

$$C_{31} = \sum_{n=0}^{\infty} E(X \mid N(0, 1) = n) P(N(0, 1) = n) = \sum_{n=0}^{\infty} \frac{n\mu^n \,\mathrm{e}^{-\mu}}{2n!} = \frac{\mu}{2}.$$

For bets 32–40 below (involving the time of the nth goal), it is simplest to consider the Poisson probability density function (PDF). Denoting the time of the nth goal by T_n, we define $N(0, x)$ as the number of goals scored during the period $[0, x)$ and consider the probability $F_n(x)$ that at least n goals will occur during the time period $[0, x)$. This is evidently given by

$$F_n(x) = 1 - P(N(0, x) < n) = 1 - \sum_{k=0}^{n-1} \frac{(\mu x)^k \, \mathrm{e}^{-\mu x}}{k!}.$$

We also note that $P(T_n \leqslant x) = 1$ for $x = 1$ (since, according to the rules of the bets, if fewer than n goals are scored, then the time of the nth and all subsequent goals is tallied at 1). The PDF f_n of T_n may now be found in the normal way by differentiating the cumulative density function (CDF) to yield (for $0 \leqslant x < 1$) that

$$f_n(x) = -\sum_{k=0}^{n-1} \frac{\partial}{\partial x} \frac{(\mu x)^k \, \mathrm{e}^{-\mu x}}{k!} = \frac{\mu^n x^{n-1} \, \mathrm{e}^{-\mu x}}{(n-1)!}.$$

The bet will be tallied at $t = 1$ if fewer than n goals are scored, so that $P(T_n = 1) = 1 - F_n(1)$. The expected time τ_n of the nth goal is now seen to be given by

$$\tau_n = \frac{\mu^n}{(n-1)!} \int_0^1 x^n \, \mathrm{e}^{-\mu x} \, \mathrm{d}x + \mathrm{e}^{-\mu} \sum_{k=0}^{n-1} \frac{\mu^k}{k!}$$

so that (continuing to assume that a match lasts for a unit time)

$$C_{32} = \frac{1 - \mathrm{e}^{-\mu}}{\mu}, \quad C_{33} = \frac{1 - \mathrm{e}^{-\mu_{\mathrm{H}}}}{\mu_{\mathrm{H}}}, \quad C_{34} = \frac{1 - \mathrm{e}^{-\mu_{\mathrm{A}}}}{\mu_{\mathrm{A}}},$$

$$C_{35} = \frac{2 - \mathrm{e}^{-\mu}(2 + \mu)}{\mu}, \quad C_{36} = \frac{2 - \mathrm{e}^{-\mu_{\mathrm{H}}}(2 + \mu_{\mathrm{H}})}{\mu_{\mathrm{H}}}, \quad C_{37} = \frac{2 - \mathrm{e}^{-\mu_{\mathrm{A}}}(2 + \mu_{\mathrm{A}})}{\mu_{\mathrm{A}}},$$

$$C_{38} = \frac{6 - \mathrm{e}^{-\mu}(6 + 4\mu + \mu^2)}{2\mu}, \quad C_{39} = \frac{6 - \mathrm{e}^{-\mu_{\mathrm{H}}}(6 + 4\mu_{\mathrm{H}} + \mu_{\mathrm{H}}^2)}{2\mu_{\mathrm{H}}},$$

$$C_{40} = \frac{6 - \mathrm{e}^{-\mu_{\mathrm{A}}}(6 + 4\mu_{\mathrm{A}} + \mu_{\mathrm{A}}^2)}{2\mu_{\mathrm{A}}}.$$

To find the expected time $E(t_{\mathrm{f}})$ of the final match goal, we assume that no goal has been scored during the time period $[0, t)$ and also that the bet can be tallied at any time between 0 and 1 (since if no match goals are scored, the bet is tallied at 0). The CDF $F_{\mathrm{f}}(t, x)$ is then given by

$$F_{\mathrm{f}}(t, x) = P(t_{\mathrm{f}} = 0) + P(t \leqslant t_{\mathrm{f}} \leqslant x) = 1 - P(N(x, 1) \geqslant 1) = P(N(x, 1) = 0) = \mathrm{e}^{-\mu(1-x)}$$

so that

$$C_{41} = E(t_{\mathrm{f}}) = 0 \, \mathrm{e}^{-\mu} + \int_0^1 x \mu \, \mathrm{e}^{-\mu(1-x)} \, \mathrm{d}x = 1 - \frac{1}{\mu} + \frac{\mathrm{e}^{-\mu}}{\mu}.$$

To calculate the expected time of the winning goal, it is easiest to note that if the home team scores the winning goal in the time interval $(t, t + \mathrm{d}t)$, then three events must happen. First, the home team must score in the interval $(t, t + \mathrm{d}t)$, an event that has a leading-order probability of $\mu_{\mathrm{H}} \, \mathrm{d}t$. Second, the away team must score exactly m goals $(0 \leqslant m < \infty)$ in the match (an event with probability $\exp(-\mu_{\mathrm{A}}) {\mu_{\mathrm{A}}}^m / m!$) and finally, the home team must score m goals in the period $[0, t)$, which happens with probability $\exp(-t\mu_{\mathrm{H}})(t\mu_{\mathrm{H}})^m/m!$. Summing over all positive m and adding a similar result that applies when the winning goal is scored by the away team, we find that the density function $g_w(t)$ is given by

$$g_w(t) = \sum_{m=0}^{\infty} \frac{(\mu_{\mathrm{H}}\mu_{\mathrm{A}}t)^m}{(m!)^2}[\mu_{\mathrm{H}}\,\mathrm{e}^{-\mu_{\mathrm{A}}-t\mu_{\mathrm{H}}} + \mu_{\mathrm{A}}\,\mathrm{e}^{-\mu_{\mathrm{H}}-t\mu_{\mathrm{A}}}] = I_0(2\theta\sqrt{t})[\mu_{\mathrm{H}}\,\mathrm{e}^{-\mu_{\mathrm{A}}-t\mu_{\mathrm{H}}} + \mu_{\mathrm{A}}\,\mathrm{e}^{-}\,\mu_{\mathrm{H}} - t\mu_{\mathrm{A}}].$$

Integration of $g_w(t)$ between 0 and 1 may easily be shown to yield the result

$$\int_0^1 g_w(t)\mathrm{d}t = 1 - \phi^2 I_0(2\theta),$$

which is correct since the probability that no winning goal is scored is simply the probability that the match is a draw, which is given by (3.2). If we now recall that, in the event of a draw, the time of the winning goal is tallied at 0, we find that the centre spread for this bet is given by

$$C_{42} = \int_0^1 t g_w(t)\mathrm{d}t = \int_0^1 t I_0(2\theta\sqrt{t})[\mu_{\mathrm{H}}\,\mathrm{e}^{-\mu_{\mathrm{A}}-t\mu_{\mathrm{H}}} + \mu_{\mathrm{A}}\,\mathrm{e}^{-\mu_{\mathrm{H}}-t\mu_{\mathrm{A}}}]\mathrm{d}t$$

which must be calculated numerically as no closed-form expression appears to exist for the integral.

The home- and away-goal minutes bets are valued in the same way as the total match goal minutes (bet 31 above), giving

$$C_{43} = \frac{\mu_{\mathrm{H}}}{2}, \quad C_{44} = \frac{\mu_{\mathrm{A}}}{2}.$$

Finally, to find the expected value of the total time that the home team will lead the match, we note that the probability that either side scores more than one goal in the time interval $I_t = [t, t + \mathrm{d}t)$ is $\mathrm{O}(\mathrm{d}t^2)$ and so may be ignored. The probability $P_{\mathrm{HL}t}$ that the home team leads in I_t is therefore given by

$$P_{\mathrm{HL}t} = P(H_t - A_t = 0)P(\text{H scores in } I_t) + P(H_t - A_t = 1)[1 - P(\text{A scores in } I_t)]$$
$$+P(H_t - A_t = k > 1),$$

where the quantity $H_t - A_t$ denotes the home–away goal difference at time t, so that $P(H_t - A_t = N \geqslant 0) = \exp(-\mu t)\alpha^N I_N(2t\theta)$. Calculating these expressions and proceeding to leading order in $\mathrm{d}t$, we find that

$$C_{45} = \sum_{k=1}^{\infty}\int_0^1 \mathrm{e}^{-\mu t}\alpha^k I_k(2t\theta)\mathrm{d}t, \quad C_{46} = \sum_{k=1}^{\infty}\int_0^1 \mathrm{e}^{-\mu t}\alpha^{-k} I_k(2t\theta)\mathrm{d}t.$$

Once again, no closed-form expression appears to exist for either the sum or the integral.

3.2 *Inferring values for the Poisson goal means*

When examining the numerical values of the centre spreads in relation to the OB's quoted prices, some thought must be given as to how the Poisson means μ_{H} and μ_{A} may be inferred. Probably, the easiest manner to proceed is to infer the two goal means directly from the quoted centre spreads for bets 1 and 2 (goal sum and goal difference). In most cases, the goal means thus calculated will tally exactly with bets

16 and 17 and calculation may now proceed. Note, however, that for reasons that are not entirely clear, the OB's quoted centre spreads of the goal total and difference are sometimes inconsistent with the total home and away goals (e.g. for Fulham versus Charlton, UK premiership, 16 October 2006, these four centre spreads were, respectively, 2.5, 0.4, 1.5 and 1.1). If this is the case, then one has to 'take a view' about exactly how to determine the Poisson goal means for subsequent calculations.

3.3 *Adjustments for added time at the end of each half*

For the calculations of goal times, etc. in Section 3.1, it was assumed for simplicity that a match lasted for a total of one time unit. The times thus calculated may then be converted to minutes simply by multiplying by 90. Note, however, that according to the quoted rules of the OBs, each half of the match consists of exactly 45 min, so that both the 45th and the 90th min of a typical match last for more than 1 min (e.g. with typical 'added time' periods $E_1 = 1$ min in the first half and $E_2 = 3$ min in the second half, so that the 45th min lasts for a total of 2 min and the 90th for 4 min). An adjustment should therefore be made to the goal times to reflect this fact. This may be done in a number of ways, but in this study we account exactly for each minute in the match, including periods of added time in each half. An event that would occur at time $t \in [0, 1]$ (were there is no added time) is therefore assumed to actually occur at minute $\bar{k}$ where $L_T = t(90 + E_1 + E_2)$,

$$\bar{k} = \begin{cases} \lfloor L_T + 1 \rfloor & (L_T < 45), \\ 45 & (45 \leqslant L_T < 45 + E_1), \\ \lfloor L_T + 1 - E_1 \rfloor & (45 + E_1 \leqslant L_T < 90 + E_1), \\ 90 & (90 + E_1 \leqslant L_T) \end{cases}$$

and $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.

4. General betting modus operandi: the spread bet as a 'risky asset'

In order to be able to identify any of the edges described above, it is necessary to examine the centre spreads for each of the bets listed in the appendix. This may easily be done using any one of a number of methods: for the studies reported in this paper, the OB's spreads were downloaded from the relevant web site (a task that may be automated using, e.g. a JAVASCRIPT programme) and rearranged to form a text data file. A MAPLE9© worksheet was then used to read the data from file and calculate each of the implied spread centres (as in the examples illustrated in the Appendix) so that they could be compared with the quoted spreads. If one then wishes to take a different view of what values the Poisson goal means should take, the implied spreads may then be recalculated with two revised goal means in order to determine which, if any, bets are favourable.

4.1 *Bet risk and correlation: an N-asset bet portfolio*

For many types of sporting bets, it is possible to propose a scheme for calculating the probability of success. For example, a horse that is quoted at win odds of 4–1 for a race plausibly has a probability of 0.2 of winning—suitable adjustments may also be made for 'over-roundness' of any book. For many such bets, however, it is much less clear how to define the risk of the bet or to quantify how multiple and related bets are correlated. For the scheme of bet valuation described above, however, risks and correlations may easily be calculated. Moreover, many pairs of bets have strong positive or negative correlations, suggesting that 'hedging' may be possible.

We may now construct an N-asset bet portfolio. Henceforth, we suppose that we 'know' the Poisson goal means μ_H and μ_A and that we have calculated the centre spreads C_n ($n = 1, 2, \ldots, 46$) for all the 46 bets listed in the appendix. We assume that, our calculations having been completed, we identify N cases (out of the total of 92 possible buy and sell bets), where C_n lies outside the interval (B_n, T_n). (For many bets, there will be no edge. These bets will not appear in the optimum portfolio.) As explained above, the values of the Poisson goal means may be those directly inferred from the OB's quoted odds or we may have 'taken a view'. In any case, if we assume that μ_H and μ_A are in some sense 'correct', essentially we have now identified N 'risky assets' S_i that each have a positive expectation $\bar{R}_i$, a known risk σ_i and known mutual covariances σ_{ij}. We may now argue in a similar (though not identical) manner to standard Markowitz portfolio theory (see, e.g. Elton *et al.*, 2003). We identify the 'efficient frontier' F in the normal way as the part of the curve $(\sigma_\pi(y), y)$ in risk–reward space with $\partial\sigma_\pi(y)/\partial y > 0$, where for a given portfolio reward and variance

$$\bar{R}_\pi = \sum_{i=1}^{N} \lambda_i \bar{R}_i = y, \quad \sigma_\pi^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \sigma_{ij},$$

we have solved the problem

$$\min_{\lambda_1,\lambda_2,\ldots,\lambda_N} [\sigma_\pi^2] \text{ subject to} \sum_{i=1}^{N} \lambda_i = 1, \;\; \bar{R}_\pi = \sum_{i=1}^{N} \lambda_i \bar{R}_i = y$$

by solving an $N \times N$ system of linear equations to give $\lambda_i = \lambda_i(y)$. The rational punter then optimizes their betting strategy by using a bet portfolio positioned on F, choosing their position on F according to their personal utility function (risk/reward preference).

4.2 *The optimal portfolio*

In classical portfolio theory, it is usual to invest a proportion of one's total wealth in a risky portfolio and the remainder in a riskless asset with return $\bar{R}_0$. This allows us to identify an 'optimal portfolio' $\Pi = \lambda_1 S_1 + \lambda_2 S_2 + \cdots + \lambda_N S_N$, namely, the risky portfolio on the efficient frontier that lies at the point of tangency of the line $R = (R_\Pi - \bar{R}_0)(\sigma/\sigma_\Pi) + \bar{R}_0$ in risk–reward space (σ, R)—the capital market line (CML), which is calculated by solving

$$\max_{\lambda_1,\lambda_2,\ldots,\lambda_N} \left[\frac{\bar{R}_\Pi - \bar{R}_0}{\sigma_\Pi}\right] \text{subject to} \sum_{i=1}^{N} \lambda_i = 1. \tag{4.1}$$

Here, the situation is slightly different, for the timescale of spread betting means that no traditional riskless asset (bank account, building society) will yield worthwhile rewards over the 90 min of a match. The practical decision that a spread better has to make therefore concerns how much of their total available capital they should invest on each match. One may argue that this is tantamount to investing a certain proportion of one's capital in a risky portfolio and leaving the rest in a riskless asset with $\bar{R}_0 = 0$, and if this is one's point of view, then the standard CML problem may be solved to yield a (unique) optimum portfolio. The computations involved are routine, involve only the solution of linear equations and may easily be mechanized (using MAPLE9©, e.g.). All choice is therefore removed from the problem and there is only one combination of spread bets that the punter need ever consider.

Though the 'optimum portfolio' thus defined is clearly worth considering as a betting strategy, readers familiar with optimal betting strategies for 'binomial' bets will undoubtedly recognize that the problem of how much of one's total capital should be invested on a positive expectation bet is exactly that considered by Kelly (1956) and which led to the development of the 'Kelly criterion'. Spread bets are, of course, greatly different to binomial bets in their structure and the question of how to define the Kelly criterion for a single spread bet (see Chapman, 2007) or, more generally, for a portfolio of spread bets is not a trivial one. The sense in which the optimum portfolio identified above is really 'optimal' may therefore have to be revised if one wishes to maximize one's total expected return on a sequence of successive positive expectation portfolio spread bets while simultaneously minimizing the probability that one's balance may be wiped out by a single 'bad day'. The relationship between the zero-interest optimal portfolio and the Kelly stake is clearly worthy of further consideration.

5. Illustrative example

We end with an illustrative example. For the Wigan versus Manchester United premiership match of 14 October 2006, the OB's view was that the respective goal means should be $\mu_{\mathrm{H}} = 0.8$ and $\mu_{\mathrm{A}} = 1.8$. Our view was different, for it seemed as though historical goal averages (rather than current form) were being used to calculate the spreads. For this match, not only was Manchester United star Wayne Rooney (who had recently been out of form) clearly returning to his best but also he had scored four goals in three meetings with Wigan in the previous season. At the time Manchester United (current form WDLWWW) had a strong away record, having won seven of their last nine games away from Old Trafford, while the Wigan (current form LDDLWL) defence had recently been poor, having conceded at least one goal in all but one of their previous 14 premiership matches. The Manchester United strikers were in top form, and at this early point in the season they had had more goalscorers (seven), more shots on target (46) and had hit the woodwork more times (eight) than any other premiership side. Finally, the opposition coaches Sir Alex Ferguson and Paul Jewell had met five times in opposition, with United winning four of these games 4–0.

We therefore concluded that more realistic Poisson goal means were given by $\bar{\mu}_{\mathrm{H}} = 0.6$ and $\bar{\mu}_{\mathrm{A}} = 2.5$. Calculating the spreads implied by these values as detailed in Section 3.1 led to a number of bets with positive expected returns: for simplicity, we only carry out calculations for bets 2 (total goals), 16 (home goals) and 26 (Cs_{00}). The OB's respective spreads for these bets were (2.5, 2.7), (0.7, 0.9) and (11, 14); using $\bar{\mu}_{\mathrm{H}}$ and $\bar{\mu}_{\mathrm{A}}$ (with $\bar{\mu} = \bar{\mu}_{\mathrm{H}} + \bar{\mu}_{\mathrm{A}}$ and $\bar{\phi} = \exp(-\bar{\mu}_{\mathrm{H}}/2 - \bar{\mu}_{\mathrm{A}}/2)$) gave predicted centre spreads of 3.10, 0.60 and 7.56. The relevant 'risky assets' were therefore to buy total goals at $T_2 = 2.7$, sell home goals at $B_{16} = 0.7$ and sell C_{s00} at $B_{26} = 11.0$.

Evaluating the respective returns $\bar{R}_i$, variances σ_i^2 and correlations σ_{ij}, we have

$$\bar{R}_2 = \sum_{k=0}^{\infty} P(\mathrm{TG} = k)(k - T_2) = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-\bar{\mu}} \bar{\mu}^k}{k!}(k - T_2) = \bar{\mu} - T_2,$$

$$\bar{R}_{16} = \sum_{k=0}^{\infty} P(\mathrm{HG} = k)(B_{16} - k) = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-\bar{\mu}_{\mathrm{H}}} \bar{\mu}_{\mathrm{H}}{}^k}{k!}(B_{16} - k) = B_{16} - \bar{\mu}_{\mathrm{H}},$$

$$\begin{aligned}\bar{R}_{26} &= P(I = 0)[B_{26} - 0] + P(I = 25)[B_{26} - 25] + P(I = 75)[B_{26} - 75] \\ &= (1 - \bar{\phi})B_{26} + \bar{\phi}(1 - \bar{\phi})[B_{26} - 25] + \bar{\phi}^2[B_{26} - 75] = B_{26} - 25\bar{\phi} - 50\bar{\phi}^2,\end{aligned}$$

$$\sigma_2^2 = \sum_{k=0}^{\infty} P(\mathrm{TG}=k)[k - T_2 - \bar{R}_2]^2 = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-\bar{\mu}}\bar{\mu}^k}{k!}[k - T_2 - \bar{R}_2]^2 = \bar{\mu},$$

$$\sigma_{16}^2 = \sum_{k=0}^{\infty} P(\mathrm{HG}=k)[B_{16} - k - \bar{R}_{16}]^2 = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-\bar{\mu}_\mathrm{H}}\bar{\mu}_\mathrm{H}{}^k}{k!}[B_{16} - k - \bar{R}_{16}]^2 = \bar{\mu}_\mathrm{H},$$

$$\sigma_{26}^2 = B_{26}^2(1-\bar{\phi}) + (B_{26}-25)^2\bar{\phi}(1-\bar{\phi}) + (B_{26}-75)^2\bar{\phi}^2 - \bar{R}_{26}^2 = 625\bar{\phi}(1-\bar{\phi})(4\bar{\phi}^2 + 8\bar{\phi} + 1),$$

$$\sigma_{2,16} = E[(R_2 - \bar{R}_2)(R_{16} - \bar{R}_{16})] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \bar{\phi}^2 \frac{\bar{\mu}_\mathrm{H}{}^j \bar{\mu}_\mathrm{A}{}^k}{j!k!}(j + k - T_2 - \bar{R}_2)(B_{16} - j - \bar{R}_{16}) = -\bar{\mu}_\mathrm{H},$$

$$\begin{aligned}
\sigma_{2,26} = E[(R_2 - \bar{R}_2)(R_{26} - \bar{R}_{26})] &= P(0-0)(B_{26} - 75 - \bar{R}_{26})(0 - T_2 - \bar{R}_2) \\
&+ \bar{\phi}^2(B_{26} - 0 - \bar{R}_{26}) \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} \frac{(\bar{\mu}/2)^j(\bar{\mu}/2)^k}{j!k!}(j + k + T_2 - \bar{R}_2) \\
&+ \bar{\phi}^2(B_{26} - 25 - \bar{R}_{26}) \sum_{k=1}^{\infty} \frac{(\bar{\mu}/2)^k}{k!}(k - T_2 - \bar{R}_2) = \frac{25\bar{\phi}\bar{\mu}(1+4\bar{\phi})}{2},
\end{aligned}$$

$$\begin{aligned}
\sigma_{16,26} = E[(R_{16} - \bar{R}_{16})(R_{26} - \bar{R}_{26})] &= \bar{\phi}^2(B_{26} - 75 - \bar{R}_{26})(B_{16} - \bar{R}_{16}) \\
&+ \bar{\phi}^2(B_{26} - \bar{R}_{26}) \left[\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{(\bar{\mu}_\mathrm{H}/2)^j(\bar{\mu}_\mathrm{A}/2)^k(\bar{\mu}_\mathrm{H}/2)^p(\bar{\mu}_\mathrm{A}/2)^q}{j!k!p!q!}(B_{16} - (j+p) - \bar{R}_{16}) \right. \\
&\left. - \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{(\bar{\mu}_\mathrm{H}/2)^p(\bar{\mu}_\mathrm{A}/2)^q}{p!q!}(B_{16} - p - \bar{R}_{16}) \right] \\
&+ \bar{\phi}^2(B_{26} - 25 - \bar{R}_{26}) \left[-(B_{16} - \bar{R}_{16}) + \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{(\bar{\mu}_\mathrm{H}/2)^p(\bar{\mu}_\mathrm{A}/2)^q}{p!q!}(B_{16} - p - \bar{R}_{16}) \right] \\
&= -\frac{25\bar{\phi}\bar{\mu}_\mathrm{H}(1+4\bar{\phi})}{2}.
\end{aligned}$$

For illustrative purposes, it is easiest to scale bet 26 so that each of the three bets have risks, returns and covariances of similar orders of magnitude. Scaling B_{26} and the index values of 25 and 75 by a factor of 1/20 gives values

$$\bar{R}_2 = 0.4, \quad \bar{R}_{16} = 0.1, \quad \bar{R}_{26} = 0.172, \quad \sigma_2^2 = 3.1, \quad \sigma_{16}^2 = 0.6, \quad \sigma_{26}^2 = 0.752,$$
$$\sigma_{2,16} = -0.600, \quad \sigma_{2,26} = 0.760, \quad \sigma_{16,26} = -0.147.$$

The optimum portfolio weights calculated by solving (4.1) are given by

$$\lambda_1 = 0.252, \quad \lambda_2 = 0.551, \quad \lambda_3 = 0.197,$$

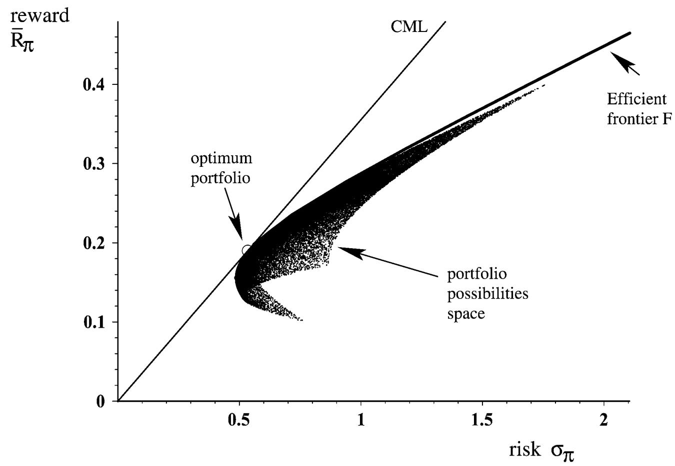

FIG. 1. Portfolio theory problem for three bets for Wigan versus Manchester United premiership match of 14 October 2006.

with associated reward, risk and CML given by

$$\bar{R}_\pi = 0.190, \quad \sigma_\pi = 0.534, \quad \bar{R} = 0.355\sigma.$$

The efficient frontier is given (for $\bar{R} > 0.156$) by

$$\sigma^2 = 1.299 - 13.684\bar{R} + 43.949\bar{R}^2.$$

Figure 1 shows the portfolio possibilities space (dots), the CML, the optimum portfolio and the efficient frontier F. For the record, Manchester United won the match 3–1. They fell behind in the fifth minute when Leighton Baines scored from a spectacular 30-yard free kick. Thereafter, however, Wayne Rooney did indeed reveal his top form, hitting the bar with a sublime piece of skill before Nemanja Vidic scored his first Manchester United goal after 62 min from a corner. Rooney then crossed for Louis Saha to score (66 min) before Ole Gunnar Solskjaer completed the scoring in the 90th min. From a spread betting point of view, the efficient frontier identified above was clearly optimal.

6. Concluding remarks

The two key matters that have been addressed in this study are (i) to value all the commonly traded soccer spread bets under the assumption that goals are scored as Poisson processes and (ii) to show how Markowitz portfolio theory may be used to bet ‘optimally’ when a sequence of positive expectation bets has been identified. We have not considered some of the other more offbeat bets that are sometimes traded such as the number of steps that the first penalty taker will take in his run up or the ‘rubbish decision ref’ bet (the number of times that the ball enters the net but no goal is scored), but it is possible to make plausible models for all of these more exotic products. Undoubtedly, it would be possible to pose many different models for the manner in which goals are scored. As far as we are concerned, however, our main aim is to understand how the OBs value such bets.

One possible problem in the portfolio theory approach that has been recommended is the sheer number of correlations that have to be calculated. A total of $46 \times 45 = 2070$ must, in principle, be calculated for the bet universe considered in this study. Though many of these calculations are related, some are quite hard to perform.

We have also not considered how the theory might be extended to cover (e.g.) international matches where there is more doubt about the underlying random nature of the goals that are scored. Note, however, that all the portfolio theory that has been presented still applies for different goal scoring models.

We also discussed various sorts of edges that may be exploited to give positive-expectation bets. The accuracy of a punter's view is all important here, and many approaches are possible. One of the most promising is to try to take advantage of the fact that there is considerable evidence that, in reality, the Poisson goal means are not constant, but increase slightly as a match progresses.

Whether profitable betting using the theory described above is a reality depends ultimately on the accuracy of one's betting model, for one has to identify circumstances when the OB's spreads are mispriced. There is evidence that many such mispricings exist, but betting is a business where creating a small edge requires hard work, and we trust that the reader will therefore understand why the details of such circumstances have not been explicitly addressed in this paper. Note also that we have chosen to estimate the goal means by using the quoted spreads. Many other methodologies are also possible, e.g. one might find the goal means by a best fit to all the spreads or use a moment method, testing one's results using a χ^2-test. Parameter risk is another issue that has not been mentioned, but could possibly be analysed with much benefit.

Finally, it is worth noting a fact that cannot be ignored: though the theory contained in this study undoubtedly allows one to bet 'optimally', as usual the OBs have the freedom to choose their customers. The inevitable result of this is that too much 'optimality' may eventually lead to one's account being wound up.

Acknowledgements

The author acknowledges helpful discussions with Dr Chris Howls (School of Mathematics, University of Southampton) and Mr Tasos Oikonomidis (School of Management, University of Southampton).

References

CHAPMAN, S. J. (2007) The Kelly criterion for spread bets. *IMA J. Appl. Math.*, **72**, 43–51.

DYTE, D. & CLARKE, S. R. (2000) A ratings based Poisson model for world cup soccer simulation. *J. Oper. Res. Soc.*, **51**, 993–998.

ELTON, E. J., GRUBER, M. J., BROWN, S. J. & GOETZMANN, W. (2003) *Modern Portfolio Theory and Investment Analysis*, 6th edn. New York: John Wiley.

FITT, A. D., HOWLS, C. J. & KABELKA, M. (2006) Valuation of soccer spread bets. *J. Oper. Res. Soc.*, **57**, 975–985.

GRAY, P. K. & GRAY, S. F. (1997) Evidence from the NFL sports betting market. *J. Financ.*, **52**, 1725–1737.

GREENHOUGH, J., BIRCH, P. C., CHAPMAN, S. C. & ROWLANDS, G. (2002) Football goal distributions and extremal statistics. *Physica A*, **316**, 615–624.

HAIGH, J. (1999) (Performance) index betting and fixed odds. *The Statistician*, **48**, 425–434.

HAIGH, J. (2000) The Kelly criterion and bet comparisons in spread betting. *Statistician*, **49**, 531–539.

KELLY, J. L. (1956) A new interpretation of information rate. *Bell Syst. Tech. J.*, **35**, 917–926.

TASSONI, C. J. (1996) Representativeness in the market for bets on National Football League games. *J. Behav. Decis. Making*, **9**, 115–124.

Appendix

The universe of spread bets that is considered in this study consists of the indices shown below. A brief description of each bet has also been included, and the numbering conventions shown are adopted throughout this study. After each description an example spread is given—in this case, the quoted spread is from an OB offering bets on the Wigan versus Manchester United, UK premiership match that took place on 14 October 2006 (this match is also used in the example discussed in Section 5). Finally, the centre spread is calculated using the methodology outlined in Section 3.1 with goal means $\mu_{\mathrm{H}} = 0.8$, $\mu_{\mathrm{A}} = 1.8$ and adjusted times calculated according to Section 3.3 with $E_1 = 1$ min and $E_2 = 3$ min.

1. Home − away spread (total home goals − total away goals) (Wigan versus Manchester United spread (−1.1, −0.9), calculated centre spread −1.0).
2. Total goals (total home goals + total away goals) (Wigan versus Manchester United spread (2.5, 2.7), calculated centre spread 2.6).
3. Home-win index (25 points for a win, 10 for a draw, 0 for a loss) (Wigan versus Manchester United spread (5.0, 6.5), calculated centre spread 6.32).
4. Away-win index (25 points for a win, 10 for a draw, 0 for a loss) (Wigan versus Manchester United spread (17.0, 18.5), calculated centre spread 17.54).
5. 'Goal rush' (total goals index: 0 = 0 points, 1 = 10 points, 2 = 20 points, 3 = 33 points, 4 = 50 points, 5 = 70 points, >5 = 100 points) (Wigan versus Manchester United spread (30, 33), calculated centre spread 31.25).
6. Home − 1 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (1.5, 3.0), calculated centre spread 2.35).
7. Home − 2 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (0.0, 1.5), calculated centre spread 0.65).
8. Home − 3 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (0.0, 0.0), calculated centre spread 0.14).
9. Home − 4 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (0.0, 0.0), calculated centre spread 0.02).
10. Home − 5 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (0.0, 0.0), calculated centre spread 0.00).
11. Away − 1 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (11.0, 12.0), calculated centre spread 11.44).
12. Away − 2 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (5.5, 7.0), calculated centre spread 6.04).
13. Away − 3 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (2.0, 3.5), calculated centre spread 2.61).
14. Away − 4 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all

losses = 0 points) (Wigan versus Manchester United spread (0.0, 0.0), calculated centre spread 0.94).

15. Away − 5 handicap (win by >handicap = 25 points, win by exact handicap = 10 points, all losses = 0 points) (Wigan versus Manchester United spread (0.0, 0.0), calculated centre spread 0.29).
16. Total number of home goals (Wigan versus Manchester United spread (0.7, 0.9), calculated centre spread 0.8).
17. Total number of away goals (Wigan versus Manchester United spread (1.7, 1.9), calculated centre spread 1.8).
18. Cs: home team score 1 and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (4, 6), calculated centre spread 5.70).
19. Cs: home team score 2 and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (3, 5), calculated centre spread 4.36).
20. Cs: home team score 3 and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (0.5, 2), calculated centre spread 1.57).
21. Cs: home team score 4+ and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (0, 1), calculated centre spread 0.43).
22. Cs: away team score 1 and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (12, 15), calculated centre spread 12.82).
23. Cs: away team score 2 and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (13, 16), calculated centre spread 14.69).
24. Cs: away team score 3 and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (7, 10), calculated centre spread 8.87).
25. Cs: away team score 4+ and win index (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (5, 7), calculated centre spread 5.74).
26. Cs: 0–0 (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (11, 14), calculated centre spread 10.53).
27. Cs: 1–1 (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (6, 9), calculated centre spread 7.8).
28. Cs: all other draws (25 if correct at half time, 50 if correct at full time: max 75) (Wigan versus Manchester United spread (1, 3), calculated centre spread 2.49).
29. Home team double result (25 for win, 10 for draw, 0 for loss, half time and full time added) (Wigan versus Manchester United spread (13, 15), calculated centre spread 14.06).
30. Away team double result (25 for win, 10 for draw, 0 for loss, half time and full time added) (Wigan versus Manchester United spread (32, 34), calculated centre spread 32.91).
31. Total goal minutes (sum of all minutes in which goals are scored) (Wigan versus Manchester United spread (125–135), calculated centre spread 117).
32. Time (minutes) of first match goal (tallied at 90 if no goals are scored in the match) (Wigan versus Manchester United spread (35, 38), calculated centre spread 32.04, adjusted centre spread 34).
33. Time (minutes) of first home goal (tallied at 90 if no home goals are scored in the match) (Wigan versus Manchester United spread (64, 67), calculated centre spread 61.95, adjusted centre spread 64).
34. Time (minutes) of first away goal (tallied at 90 if no away goals are scored in the match) (Wigan

versus Manchester United spread (42, 45), calculated centre spread 41.73, adjusted centre spread 44).

35. Time (minutes) of second match goal (tallied at 90 if no second goal is scored in the match) (Wigan versus Manchester United spread (60, 63), calculated centre spread 57.40, adjusted centre spread 59).
36. Time (minutes) of second home goal (tallied at 90 if no second home goal is scored in the match) (Wigan versus Manchester United spread (82, 85), calculated centre spread 83.46, adjusted centre spread 87).
37. Time (minutes) of second away goal (tallied at 90 if no second away goal is scored in the match) (Wigan versus Manchester United spread (68, 71), calculated centre spread 68.59, adjusted centre spread 71).
38. Time (minutes) of third match goal (tallied at 90 if no third goal is scored in the match) (Wigan versus Manchester United spread (76, 79), calculated centre spread 74.07, adjusted centre spread 77).
39. Time (minutes) of third home goal (tallied at 90 if no third home goal is scored in the match) (Wigan versus Manchester United spread (88, 90), calculated centre spread 88.79, adjusted centre spread 90).
40. Time (minutes) of third away goal (tallied at 90 if no third away goal is scored in the match) (Wigan versus Manchester United spread (82, 84), calculated centre spread 82.06, adjusted centre spread 85).
41. Time (minutes) of last match goal (tallied at 0 if the match is goalless) (Wigan versus Manchester United spread (61, 64), calculated centre spread 57.96, adjusted centre spread 60).
42. Time (minutes) of winning goal (tallied at 0 if the match ends in a draw) (Wigan versus Manchester United spread (34, 37), calculated centre spread 32.06, adjusted centre spread 34).
43. Home team total goal minutes (sum of all minutes in which home goals are scored) (Wigan versus Manchester United spread (35, 37), calculated centre spread 36).
44. Away team total goal minutes (sum of all minutes in which away goals are scored) (Wigan versus Manchester United spread (80, 82), calculated centre spread 81).
45. Home team lead minutes (time (minutes) for which home team is in the lead) (Wigan versus Manchester United spread (8, 11), calculated centre spread 12.20).
46. Away team lead minutes (time (minutes) for which away team is in the lead) (Wigan versus Manchester United spread (32, 35), calculated centre spread 37.38).