Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2015)
178, *Part* 3, *pp.* 659–683

Joint modelling of goals and bookings in association football

A. C. Titman, D. A. Costain, P. G. Ridall and K. Gregory

Lancaster University, UK

[Received November 2012. Final revision April 2014]

Summary. A multivariate counting process formulation is developed for the quantification of association football event interdependences which permits dynamic prediction as events unfold. We model data from English Premier League and Championship games from the 2009–2010 and 2010–2011 football seasons and assess predictive capacity by using a model-based betting strategy, applied prospectively to available live spread betting prices. Both the scoreline and the bookings status were predictive of match outcome. In particular, the award of a red card led to increased goal rates for the non-penalized team and the home team scoring rate decreased once they were ahead. Overall the betting strategy profited with gains made in the bookings markets.

Keywords: Association football; Counting process; Dynamic prediction; Weibull distribution

1. Introduction

The course of a game of association football appears to change very quickly, often resulting in an apparently unpredictable course of influential events. Vast sums of money are spent by clubs to enhance the strength of their teams through the buying and selling of players, coaches and managers. However, even when these relative strengths and weaknesses have been allowed for, the game remains highly stochastic. The outcome is often determined by a chain of ‘freak’ events such as goals against the run of play, misconduct of a player and the marginal decisions of referees on the awarding of bookings.

Anecdotally, it is often claimed that a football match has been ‘ruined’ for a particular team because of a poor refereeing decision. The most direct case might be where the opposing team are erroneously awarded a penalty from which they subsequently score. Alternatively, the referee might issue a player a red card causing that player to be sent off the field for the remainder of the game and leaving his team a player short.

Such intramatch event dynamics are of intrinsic interest in this work. In particular, we aim to investigate the stochastic interplay between the goal scoring and bookings processes with a view to quantifying subsequent event-induced inhomogeneity in the home and away team process intensities (rates) and providing probabilistic statements regarding match outcomes in realtime.

Various researchers have considered factors which affect booking rates; in particular, Buraimo *et al.* (2010) and Dawson *et al.* (2007) investigated the potential for referee bias (a tendency to favour a particular team). Buraimo *et al.* (2010) developed a minute-by-minute probit analysis for the joint analysis of the awarding of red and yellow cards and on adjusting for factors such as

Address for correspondence: A. C. Titman, Department of Mathematics and Statistics, Fylde College, Lancaster University, Lancaster, LA1 4YN, UK.
E-mail: a.titman@lancaster.ac.uk

 0964–1998/15/178659

features of stadia, relative team ability, home team support and derby games found evidence of bias in referees. Similarly, Dawson *et al.* (2007) constructed bivariate models for the number of disciplinary sanctions (red and yellow cards) made against players in the English Premier League over the period 1996–2003. On adjusting for home advantage, relative team quality and match-specific factors (such as crowd size and whether the match is significant for the championship) they found that individual referee effects contribute significantly to the explanatory power of their conditional model.

In terms of investigating the effect of bookings on goal rates Ridder *et al.* (1994) considered the effect of red cards on scorelines: the so-called 'expulsion effect'. Assuming independent goal scoring intensities they developed a conditional model for the number of goals scored after the expulsion time, conditioning on the total goals scored. On the basis of games involving red cards in the two Dutch professional leagues (seasons 1989–1990, 1990–1991 and 1991–1992) and on adjusting for relative team strength and time since the kick-off they found that the scoring rate increases significantly for the team with 11 players (those not penalized). They also found that the earlier a red card is issued to an opposing team the more likely a team are to win the game. In a similar investigation into red card effects on goals, Vecer *et al.* (2009) found contrasting results. Using betting data on the Fédération Internationale de Football Association World Cup 2006 and Euro 2008 they showed that the scoring rate dropped significantly for the penalized team, whereas the rate change for the opposing, 11-man, team increased only slightly. The potential for differential home and away team effects was not considered in either case.

Models for interdependences in home and away team goal scoring rates *per se* also feature in the literature. Most closely related to the approach that is taken in this paper is the stochastic birth process modelling that was conducted by Dixon and Robinson (1998). They developed a bivariate Poisson model formulation which incorporates the attacking and defensive strengths of a team, home ground advantage, current scoreline and time left to play. Note that time of play was mapped to the interval [0,1] as opposed to actual time since kick-off. Non-stationarity in the baseline rate was modelled parametrically; model selection based on the Akaike information criterion was used to choose the functional form. They found that the goal rate was monotone increasing (linearly) for both the home and the away teams and, also, that the rates were dependent on the current scoreline. Most notably they found that when the home team have a narrow lead the home and away team scoring rates decreased and increased respectively. In a similar bivariate Poisson process framework for goals scored, Volf (2009) considered a semi-parametric model formulation consisting of a non-parametric baseline intensity and regressors reflecting rival teams' defensive strength and the state of the match. The model was used to analyse quarter-finalists data from the 2006 World Cup; model-based Monte Carlo predictions were then compared with the actual results. Although paucity of data (40 games) led to estimate uncertainty, Volf also found an increase in the goal rate once a team were losing.

A methodological aim when investigating interdependences between match events is to yield a flexible modelling framework which can be used for 'realtime' prediction of match outcomes. Notwithstanding limitations in the event types *a priori* considered, the models that have been used to investigate referee bias either focused on the overall match outcome (Dawson *et al.*, 2007) or conditioned out the goal events (Buraimo *et al.*, 2010) and thus cannot be used to predict dynamically. Similarly, Ridder *et al.* (1994) simplified their inferential procedure by conditioning on the total number of goals scored. In contrast, although the Poisson model formulation that was developed by Dixon and Robinson (1998) can be used dynamically, the methodology is limited in that they did not consider the effect of bookings on goal rates and vice versa. In addition, bookings constitute an increasingly popular betting market and their model does not permit bookings to be predicted or allow for investigation of other contributory factors such as referee decisions favouring home teams.

In this paper we seek to increase knowledge and understanding of the interplay between football event processes by not only modelling the interdependence between home and away team goals but also by quantifying the effect that match bookings (home and away team red and yellow cards) have on the course of and outcome of a game. We quantify the occurrence of influential events through use of a realtime eight-dimensional multivariate counting process, furnishing a flexible and elegant modelling framework. More specifically, we assume that the state of play at any instance during the course of a match can be summarized by the number of red cards (independent and red cards from a second yellow card), yellow cards and goals awarded to both the home and the away teams. In each period of play (the first and second half) distinct event-specific Weibull baselines are used to capture 'game time' non-stationarity and to allow for potential discontinuity at recommencement of play. Furthermore, in contrast with utilizing time-weighted scoring records to inform attacking and defensive strength (Dixon and Robinson, 1998) the influence of time-fixed, match-specific factors such as attacking and defensive strength, team form and home field advantage are implicitly handled by considering them as non-linear functions of the bookmakers' current prematch betting odds. Our primary interest is in the dynamic interdependence between the event processes, once match constant factors have been dealt with; information which may not easily be formalized. More generally, Andersen *et al.* (1993) have provided a thorough account of statistical models formulated on the basis of counting processes.

In the UK, association football constitutes a dominant share of sports betting markets. Aside from the fixed odds markets, which express potential gains (fractional) or returns (decimal) relative to a stake, speculators can use available data to trade in spread betting markets. Spread betting operates dynamically as events unfold and involves choosing to buy (a bet on the outcome being higher) or to sell (a bet on the outcome being lower) against the outcome predicted by the bookmaker: the so-called 'spread'. This can be, for instance, on the total number of goals or the total number of red cards awarded. A feature of this type of betting is that the gains or liabilities resulting from a modest bet may be almost unlimited. This is because you are rewarded or penalized according to the margin between your current bet and the final outcome.

To illustrate the predictive utility of our model, we consider its performance in live spread betting markets. The actions that are taken at any instance of time involve the placing of bets, either buying or selling, on four possible betting markets. These actions are taken to maximize an expected utility function. This is a function of the buying and selling live betting price and the current state of the game, and includes our risk propensity. The expectation of utility is carried out with respect to our model-based 'belief' (predictive distribution) in the final state of play. This belief is quantified by repeatedly simulating from the asymptotic distribution of our model parameters and in turn simulating the final outcome from the model conditionally on the current state of play. It is perhaps pertinent to note here that the model is formulated primarily for the quantification of goal and booking rate interdependences as opposed to providing a competitive betting strategy *per se*. In particular, we utilize bookmakers' prematch odds to adjust for time-fixed factors and hence some degree of dependence between the available spread and the expected match outcome will be induced.

The outline of the paper is as follows. A description of the data is given in Section 2 and the multivariate counting process model formulation is presented in Section 3. Section 4 presents the modelling results and in Section 5 predictive capacity is investigated via an automated, model-based, spread betting strategy applied prospectively to monitored prices. The paper concludes with a brief overview in Section 6.

The data that are analysed in the paper and the programs that were used to analyse them can be obtained from

http://wileyonlinelibrary.com/journal/rss-datasets

2. The data

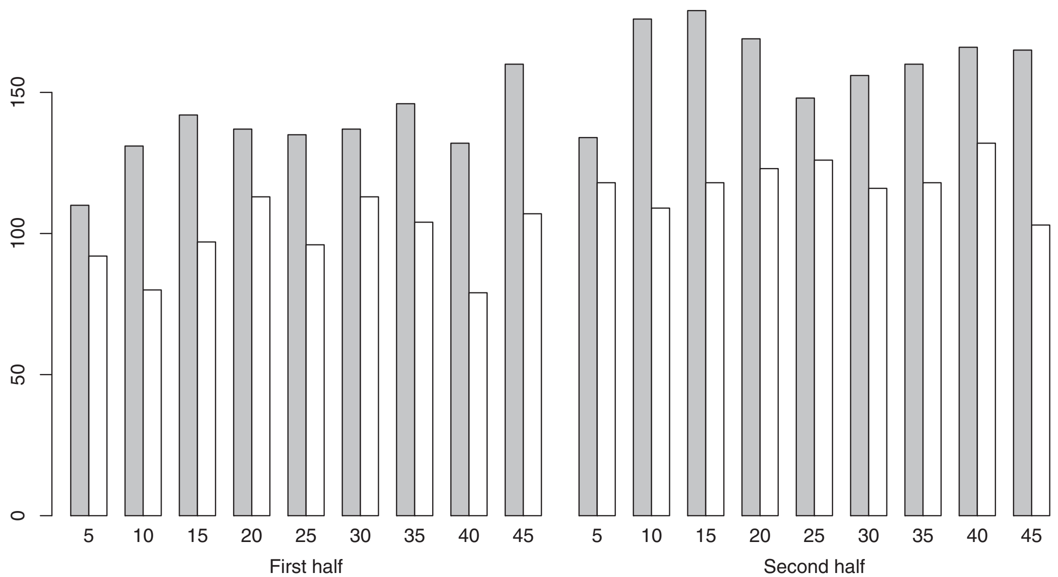
The data are taken from the live match reports on the BBC Sport Web site www.bbc.co.uk/sport and consist of the times of the match-specific events (home and away team goals and bookings) to the nearest second since the kick-off (for first events occurring in the first half) or since the start of the second half (for second-half events). All $M=1858$ Premier League and Championship games in the 2009–2010 and 2010–2011 football seasons for which a live match report was available were used. 758 were Premier League games (with two games missing) and 1100 Championship games (with four games missing). Additional data on prematch bookmakers' odds were obtained from www.football-data.co.uk.

The marginal frequencies for the home and away team goals and bookings are presented in Table 1. The overall counts indicated a possible home ground advantage in that the home team conceded fewer goals and received fewer disciplinary sanctions (bookings). More specifically, 2930 of 5038 goals (58%) were scored by the home team and, for example, 3263 away team yellow cards were issued compared with 2511 home team yellow cards. This finding is consistent with previous research outcomes; home team advantage is a well-documented phenomenon with hypothesized explanations including crowd support, familiarity, travel fatigue and referee bias (Pollard, 2008).

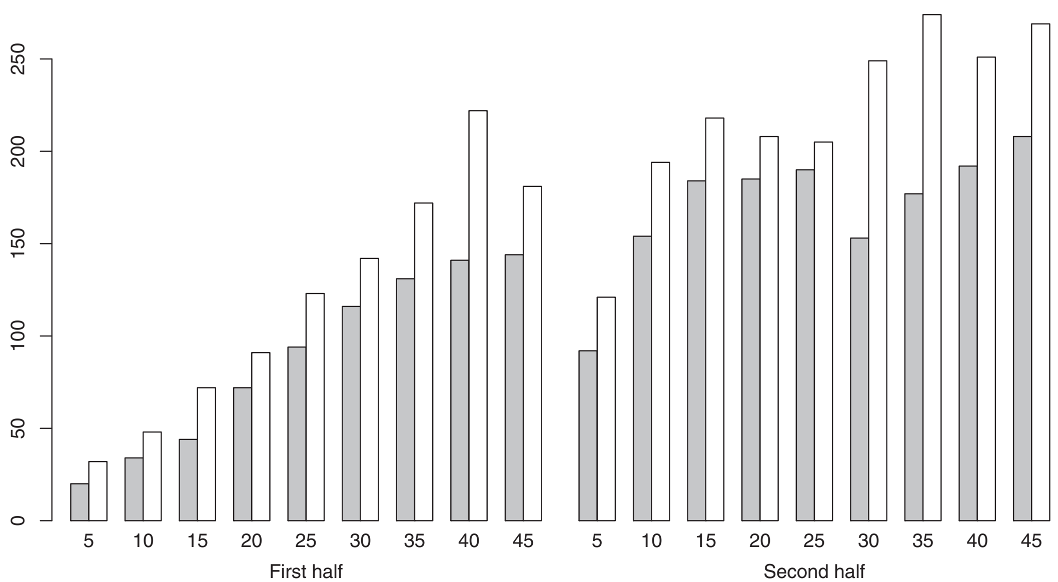
The marginal distribution of the number of goals and the number of bookings for the home and away team in each 5-min interval of normal time by period of play (half) are presented in Fig. 1. For comparability, events during stoppage time are not presented but are included in the statistical modelling. Throughout the 90 min of play the home and away team goal rates and bookings rates differed, indicating a 'sustained' home team advantage. Specifically, during all subintervals of the game the number of home team goals is higher, whereas the number of home team bookings is consistently lower. Although it may be tempting to speculate, the causal mechanisms underpinning these differences are unclear. For example, bookings rates could differ because the away team play more aggressively to compensate a home team advantage in

Table 1. Observed number of home and away team goals and yellow and red cards for the $M=1858$ English Premier League and Championship games in football seasons 2009–2010 and 2010–2011 for which a live match report was available (six games missing)

Event type	*Event count*
Home team goal	2930
Away team goal	2108
Home team yellow card	2511
Away team yellow card	3263
Home team independent red card	73
Away team independent red card	100
Home team red card from 2nd yellow card	56
Away team red card from 2nd yellow card	89



(a)



(b)

Fig. 1. Marginal distribution of (a) the number of goals and (b) the number of bookings in each 5-min interval of normal time for the first and second half of play: ■, home team; □ away team

attacking or defensive strength, or, alternatively, because referees' decisions favour the home team.

Temporal trends are evident as is some degree of home and away team variation in their development. Over the course of a game the home team goal rate generally increases over the first half of the game and dips at the start of the second half before peaking at around 15 min into the second half. The away team goal rate exhibits somewhat less structure, fluctuating in the first half of play, falling slightly at the beginning of the second half and then increasing slightly during the period 26–40 min into the second half.

For bookings (Fig. 1(b)) there is a more discernible pattern for both teams. In particular, during the first half of play the number of bookings is monotone increasing from kick-off until the last 5 min of the period when the away team rate falls. During the second period of play there is a degree of fluctuation and, moreover, at recommencement of play the rate is substantially lower for both teams than for the last 5 min of the first half. On first inspection one might be tempted to conclude that for each match the bookings rate increases over the first 40 min. However, as Dixon and Robinson (1998) noted, overinterpretation of trends in aggregated data can be problematic since a gradual increase could be due to a true increase in rate (e.g. due to tiredness), or, alternatively, could arise from variation due to dependence on the current score. For example, if the rate of bookings is constant until a goal has been scored or a booking has been made then when aggregated over matches the bookings pattern could be as observed in the first period of play. Analogous arguments hold for the apparent temporal trends in the goal rates; dynamics may simply reflect changes in scoreline and numbers of bookings.

3. Multivariate counting process model formulation for match events

Assume that in football games $j=1,2,\ldots,M$ each consisting of two periods $k=1,2$ of play (termed halves) events of interest, of types $l=1,\ldots,L$, occur at times T_{ijk}, $i=1,2,\ldots,n_{jk}$, where n_{jk} denotes the total number of events in the kth half of game j. The total number of events in a game is then given by $N_j=n_{j1}+n_{j2}$. We consider eight distinct types of event, $L=8$, the occurrences of which are denoted by an event indicator E_{ijk} given by

$$E_{ijk}=\begin{cases} 1, & \text{home team goal,} \\ 2, & \text{away team goal,} \\ 3, & \text{home team yellow card,} \\ 4, & \text{away team yellow card,} \\ 5, & \text{home team independent red card,} \\ 6, & \text{away team independent red card,} \\ 7, & \text{home team red card from second yellow card,} \\ 8, & \text{away team red card from second yellow card.} \end{cases}$$

In addition, the end of a half is treated as a censoring event taking event indicator $E_{ijk}=0$. Complete data for game j then consist of the sequence of event times for each period of play $T_{1jk}, T_{2jk}, \ldots, T_{n_{jk}jk}$ for $k=1,2$ and the corresponding event types E_{ijk} taking values in $l=\{1,\ldots,8\}$ except that $E_{n_{jk}jk}=0$ for $k=1,2$.

To describe the process of events we use a multivariate stochastic counting process formulation. Events in the first half are represented by a process

$$\mathbf{X}_{j1}(t)=\{X_{j1l}(t), j=(1,\ldots,M), l=(1,\ \ldots,8)\}$$

where each $X_{j1l}(t)$ is itself a counting process and for which at time t the process is orderly such that a unit increase in a single component $l\in\{1,\ldots,8\}$ is permitted. $X_{j1l}(0)=0$ for all l. Events in the second half are represented by a process

$$\mathbf{X}_{j2}(t)=\{X_{j2l}(t), j=(1,\ldots,M), l=(1,\ldots,8)\}$$

where $X_{j2l}(0)=n_{j1l}$ for all l and t here is the time since the game recommenced in the second half.

Each of the event processes has an associated intensity (hazard function)

$$h_{jkl}\{t;\mathcal{F}_j(t)\}=\lim_{\Delta t\to 0}\frac{P\{X_{jkl}(t+\Delta t)=i+1|X_{jkl}(t)=i\}}{\Delta t}$$

where $\mathcal{F}_j(t)$ represents the history, or so-called 'filtration', of the process $\mathbf{X}_j(t)$. The intensity of the process is taken to be a function of both time-fixed covariates and time varying covariates jointly denoted $z(t)$. The time varying components consist of the process states $X_{jkl}(t)$ (or functions thereof) and we make a Markov assumption that the hazard depends only on the current state:

$$h_{jkl}\{t;\mathcal{F}_j(t)\}=h_{jkl}\{t;\mathbf{X}_j(t)\}.$$

To facilitate prediction we assume a Weibull proportional hazards formulation and allow for distinct rate parameters in each half of a game so that the hazard function of event type l in period k takes the form

$$h_{jkl}\{t,\mathbf{X}_j(t)\}=h_{jkl}(t)\exp\{\beta^{\mathrm{T}}z_j(t)\} \tag{3.1}$$

where $h_{jkl}(t)$ denotes the event-specific baseline hazard function in game half k. The baseline hazard and cumulative baseline hazard functions are thus specified parametrically:

$$h_{jkl}(t)=\alpha_{kl}\lambda_{kl}(\lambda_{kl}t)^{\alpha_{kl}-1},$$
$$H_{jkl}(t)=\int_0^t h_{jkl}(u)\,\mathrm{d}u=(\lambda_{kl}t)^{\alpha_{kl}}$$

where α_{kl} and λ_{kl} respectively denote the event period-specific Weibull distribution shape and rate parameters and recall that time $t\geqslant 0$ is the time since start of play in the respective halves k. This model formulation permits the Weibull scale parameters to be non-stationary functions of the covariates $z_j(t)$ and allows for a different baseline hazard in each half of a game.

Assuming conditional independence, the likelihood for a given football game is simply the product over the game-specific events and hence the likelihood over the M games is

$$\prod_{j=1}^{M}\prod_{k=1}^{2}\prod_{i=1}^{n_{jk}}\prod_{l=1}^{8}h_{jkl}\{T_{ijk};X(T_{ijk}^{-})\}^{\delta(E_{ijk},l)}\exp\left[-\sum_{l=1}^{8}\int_{u=T_{(i-1)jk}}^{T_{ijk}}h_{jkl}\{u,X(T_{ijk}^{-})\}\,\mathrm{d}u\right] \tag{3.2}$$

where δ is the Dirac function and $X(T_{ijk}^{-})$ denotes the process state immediately before time t. Note that $\delta(E_{ijk},l)=0$ for all l at the end of the first period of play and also at full time.

To evaluate the likelihood we exploit the fact that $z_j(u)$ is constant over the interval $u\in(T_{(i-1)jk},T_{ijk})$ so that

$$\int_{T_{(i-1)jk}}^{T_{ijk}}h_{kl}\{u,z(u)\}\,\mathrm{d}u=\exp\{\beta^{\mathrm{T}}z_j(T_{ijk}^{-})\}\int_{T_{(i-1)jk}}^{T_{ijk}}h_{kl}(u)\,\mathrm{d}u$$

where $h_{kl}(u)$ is a deterministic function of time and

$$\int_{T_{(i-1)jk}}^{T_{ijk}}h_{kl}(u)\,\mathrm{d}u=H_{kl}(T_{ijk}^{-})-H_{kl}(T_{(i-1)kl}).$$

From expressions (3.1) and (3.2) the likelihood of parameter vectors $\theta=(\alpha,\beta,\lambda)$ can thus be expressed as

$$L(\alpha,\beta,\lambda)=\prod_{j=1}^{M}\prod_{k=1}^{2}\prod_{i=1}^{n_{jk}}\prod_{l=1}^{8}[h_{kl}(T_{ijk}^{-})\exp\{\beta^{\mathrm{T}}z_j(T_{ijk}^{-})\}]^{\delta(E_{ijk},l)}$$
$$\times\exp\left[-\sum_{j=1}^{M}\sum_{k=1}^{2}\sum_{i=1}^{n_{jk}}\sum_{l=1}^{8}\{H_{kl}(T_{ijk}^{-})-H_{kl}(T_{(i-1)jk})\}\exp\{\beta^{\mathrm{T}}z(T_{ijk}^{-})\}\right].$$

3.1. Modelling intragame event intensity functions

The hazard of each event type, $l\in 1,\ldots,8$, is assumed to be Weibull with respect to time since kick-off for events in the first half of play, $k=1$, and time since the game restarted for events in the second half, $k=2$. This means that the baseline event rates are assumed to be monotonically increasing or decreasing in time since the kick-off in the respective half, with increasing rates if the shape parameter is greater than 1. The exploratory analysis of marginal event rates identified a lower rate of events at the start of the second half compared with the end of the first half. Having separate Weibull intensities for the first and second halves enables such a feature to be accommodated. Sensitivity of the results to the Weibull model specification is considered in Section 5.

For goals distinct Weibull shape and scale parameters are assumed for the first and second halves. For all *bookings* we assume that the same Weibull shape parameter α_{kl} applies to a particular team (home or away) in a particular half, $k=1,2,$ but we allow the scale parameters λ_{kl} to vary. Effectively this means that we assume that the probability that a booking will be of a particular type remains constant, for a particular team, within a particular half, if the covariates remain the same. This assumption was adopted because of the small number of red card events, which made it difficult to estimate distinct shape parameters.

A general model could allow the covariate effects β to differ between halves of the match. However, because of small numbers of some events, particularly red cards in the first half, and lack of evidence of a difference when fitting models, in the analyses presented we assume that covariate effects are the same in each half. This means, for instance, that we assume that the effect of the award of a red card on a team's goal rate is the same regardless of the half. Specific details of the covariates $z(t)$ that are considered in the event intensities follow in Sections 3.1.1 and 3.1.2.

3.1.1. *Accounting for relative team ability, league and season*

The relative abilities of the opposing teams is likely to affect the respective goal rates and, perhaps to a lesser extent, the booking rates of a match. Previous work on modelling the dynamics of matches has built in team-specific covariates. Dawson *et al.* (2007) fitted a separate model to past match results to obtain estimates of the match outcome probabilities. They then used the estimated quantity $P(\text{home win})+0.5\,P(\text{draw})$ as a covariate in their bivariate Poisson count model for disciplinary points.

An alternative measure of relative team ability can be derived from the bookmaker's fixed price odds of match outcomes, which have been demonstrated to have high predictive ability (Forrest *et al.*, 2005). If the available decimal odds on an outcome f are O_f, then this implies that a bet of u units returns $O_f u$ if the outcome occurs, i.e. the total amount won is $(O_f-1)u$. Moreover, the implied probability of event f is $O_f^{-1}/\Sigma_F\,O_F^{-1}$ where the summation is over all possible events F. We propose to use

$$q\in A=\log(\bar{O}_{\mathrm{aw}}/\bar{O}_{\mathrm{hw}})$$

where $\bar{O}_{\mathrm{hw}}$ and $\bar{O}_{\mathrm{aw}}$ are the mean decimal odds from 10 leading UK bookmakers (including

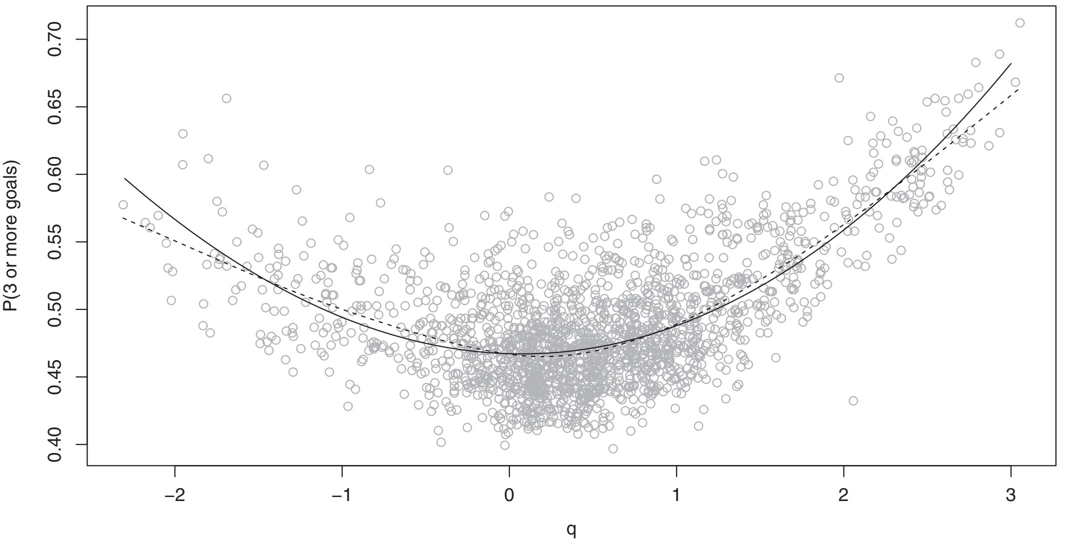


Fig. 2. Relationship between the bookmaker's implied estimate of q, a measure of relative team ability, and the bookmaker's implied estimate of the probability of at least 3 goals: ———, smoothing spline fit; - - - - - -, cubic regression model fit

Ladbrokes, William Hill and Bet365) of a home win and away win respectively. If the teams are evenly matched q will be approximately 0, whereas q will be positive in matches where the home team are expected to win and negative for matches where the away team are dominant. Rather than assume a direct linear relationship between q and goal scoring or booking rates, we use a B-spline function of q to allow a flexible range of possible relationships. On adjusting for q we also considered the effect of *league* and *season* on both scoring and booking rates.

A possible limitation of using q for the goal scoring rates is that the relative scoring abilities of the teams is not necessarily well reflected. A further match outcome that is often bet on is whether the total number of goals in a match is greater than or equal to 3. 3 goals are a popular number because the outcome has a roughly 50% chance of occurring in an average Premier League game. To inform the goal rates further, we take the implied probability $p_3 = 1/\bar{O}_{G\geqslant 3}/(1/\bar{O}_{G\geqslant 3} + 1/\bar{O}_{G<3})$ as an estimate of the probability of at least 3 goals in a particular match, where $\bar{O}_{G\geqslant 3}$ and $\bar{O}_{G<3}$ are the average prematch decimal odds of at least 3 and less than 3 goals from a set of bookmakers. Clearly, p_3 and q will be strongly related; Fig. 2 shows their relationship. Rather than use p_3 directly, we estimate $\hat{p}_3(q)$ by using the average value of p_3 for a particular value of q, where $\hat{p}_3(q)$ is estimated by using a regression model. A simple cubic regression model fits the data very similarly to a smoothing spline. The 'residual' value $r_3 \in A = p_3 - \hat{p}_3(q)$ is then used in the models for the home and away goal intensity.

3.1.2. *Modelling of event interdependences*

Red cards resulting from a second yellow card are directly related to the number of existing yellow cards that a team have accumulated. In particular, there is no possibility of receiving a red card through a second yellow card until at least one yellow card has been awarded. We therefore model the hazard that the home team receive a red from a yellow card as

$$h_{jk7}\{t, z_j(t), X_{jk3}(t)\} = h_{k7}(t)\, X_{jk3}(t)^{\nu} \exp\{\beta^{\mathrm{T}} z_j(t)\}$$

where $X_{jk3}(t)$ denotes the total number of yellow cards that the home team have obtained up to time t and $z_j(t)$ are other covariates, including states of, and functions of, the birth processes. An analogous model is used for the away team intensity $h_{jk8}\{t, z_j(t), X_{jk4}(t)\}$. If $\nu = 1$ then this would equate to the risk of red cards from yellow cards being directly proportional to the number of existing yellow cards. However, for more robust analysis we allow ν to be estimated.

More generally, for goal rates we consider the number of home and away team red and yellow cards and for bookings whether the home and away teams have accrued any bookings at all.

When considering the effect of scorelines on home and away *goals*, six current scores are considered distinct (0–0, 1–0, 0–1, equal with more than 1 goal, EE, home team ahead with 2 or more goals, H2A, and away team ahead with 2 or more goals, A2A). These classifications are the same as in Dixon and Robinson (1998) except that we do not make a distinction between a scoreline of 1–1 and higher drawn scores. For *bookings* in the model presented, we consider only whether the current score is drawn, the home team are ahead, HA, or the away team are ahead, AA.

4. Results

4.1. Main results

Statistical analyses were performed by using R version 2.10.1 (R Development Core Team, 2009). The R package `eha` (Broström, 2014) was used to fit the Weibull event process models. Table 2 gives the parameter estimates and standard errors for the Weibull proportional intensities model for home and away team goals. There is little evidence to suggest a difference between states 1–0 (Score_{10}) and H2A ($\text{Score}_{\text{H2A}}$) or between states 0–1 (Score_{01}) and A2A ($\text{Score}_{\text{A2A}}$). In general the home team's scoring rate decreases once they are leading, whereas the away team's rate increases once they are trailing.

Table 2. Estimated parameters and standard errors for home team goal and away team goal intensities

Parameter description	*Parameter*	*Results for home team goals:*		*Results for away team goals:*	
		Estimate	*Standard error*	*Estimate*	*Standard error*
Weibull rate parameter period 1	$\log(\lambda_{1l})$	−4.659	0.228	−3.538	0.174
Weibull shape parameter period 1	$\log(\alpha_{1l})$	0.107	0.029	0.088	0.034
Weibull rate parameter period 2	$\log(\lambda_{2l})$	−4.497	0.236	−3.307	0.185
Weibull shape parameter period 2	$\log(\alpha_{2l})$	0.075	0.025	0.050	0.030
Home team leads 1 goal to 0	Score_{10}	−0.139	0.059	−0.046	0.069
Away team leads 1 goal to 0	Score_{01}	−0.029	0.066	−0.097	0.075
Current score draw	Score_{EE}	−0.142	0.070	−0.285	0.085
Home team leads by 2 goals	$\text{Score}_{\text{H2A}}$	−0.172	0.064	0.009	0.076
Away team leads by 2 goals	$\text{Score}_{\text{A2A}}$	0.046	0.075	−0.046	0.085
Number of home team red cards	# Home Red	−0.194	0.135	0.481	0.112
Number of away team red cards	# Away Red	0.576	0.073	−0.816	0.176
Relative team ability measure: basis coefficient 1	q_{b1}	0.063	0.508	−1.259	0.466
Relative team ability measure: basis coefficient 2	q_{b2}	0.728	0.212	−0.648	0.266
Relative team ability measure: basis coefficient 3	q_{b3}	1.457	0.323	−2.246	0.391
Implied 'residual' measure of goal scoring intensity	r_3	0.993	0.568	3.040	0.647

Although the estimated hazard ratios differ in magnitude from the analogous (1–0) (0–1) effects that were found by Dixon and Robinson (1998) the general findings concur in that they conclude that 'if the home team is leading, the home and away rates generally decrease and increase respectively'.

The effect of red cards on scoring rates appear to differ for the home (# Home Red) and away (# Away Red) teams. The effect of a home red card on the home team's scoring rate is relatively modest (a reduction by 18%). In contrast, if the away team have at least one red card, the home team's scoring rate increases substantially (by 78%). Similarly, if the away team incur a red card their scoring rate drops by 56%. If the home team have a red card this increases the away team's scoring rate by 62%. Yellow cards appear to have a negligible direct effect on scoring rates for either team. Similarly, league and season did not significantly affect the goal rates and are not included in the final fitted model.

Our model assumes that the effect of red cards is linear on the logarithmic scale, i.e. the receipt of two red cards is twice as bad for a team as one, but it is not possible to verify this empirically because of the scarcity of matches where multiple red cards were incurred.

The effects of q are shown in Figs 3(a) and 3(b). As expected, the home goal rate increases with q whereas the away goal rate decreases. Similarly, having a positive value of r_3, implying a greater chance of at least 3 goals than expected by q, leads to higher goal scoring rates for the home team and especially for the away team.

Table 3 gives the parameters for red and yellow cards. A number of covariate effects apply to all types of bookings: hence the tabular repeats. Similarly, the same Weibull shape parameters apply across all bookings in a particular half. The baseline rate parameters change, however, with the rate of receipt of independent red cards being only 3% the rate for yellow cards. As expected, the rate of receipt of red cards from second yellow cards is closely related to the existing number of yellow cards awarded. The effect of $\log(X_{jk3})$ on the red card from yellow card rate was 1.15, suggesting that rates would be 122% higher if two players had already been booked compared with if only one had been booked. A slightly larger effect was observed for the away team. However, in each case the estimated effect is not significantly different from 1, suggesting that it might be adequate to assume direct proportionality with the number of yellow cards.

The score does not seem to have a strong effect on overall booking rates although there is a suggestion that booking rates for both teams drop if the home team are ahead. If the away team are leading this seems to lead to a slight increase in home booking rates and a slight decrease in away booking rates. Interestingly, a team's subsequent booking rate increases by around 25% if the opposing team are given a yellow or red card. It is not possible to establish whether this is because the opposing team change their playing style in retaliation or whether it reflects a referee's desire to try to appear even handed. Similarly, the rate of red cards awarded for a single offence increases once the team in question have at least one player with a yellow card. This seems to indicate that referees are very reluctant to award red cards for the first bookable offences in a match and also that to some extent they view a yellow card as a caution to all the players of a team.

The relative team ability covariate q had a smaller effect on booking rates than on goal rates but was statistically significant in both home and away cases (likelihood ratio statistic LR $=$ 38.8 on 3 degrees of freedom for the home team and LR $=$ 12.3 on 3 degrees of freedom for the away team). Figs 3(c) and 3(d) show the estimated effect of q on the hazards of each type of event. The respective booking rates are highest when the opposing team are slightly superior ($q \approx -0.8$ for the home team and $q \approx 1.2$ for the away team) and is lower both when the opposing team are greatly superior and when the opposing team are inferior. In both cases the B-spline fit was a substantial improvement compared with a linear fit. We also considered a model in

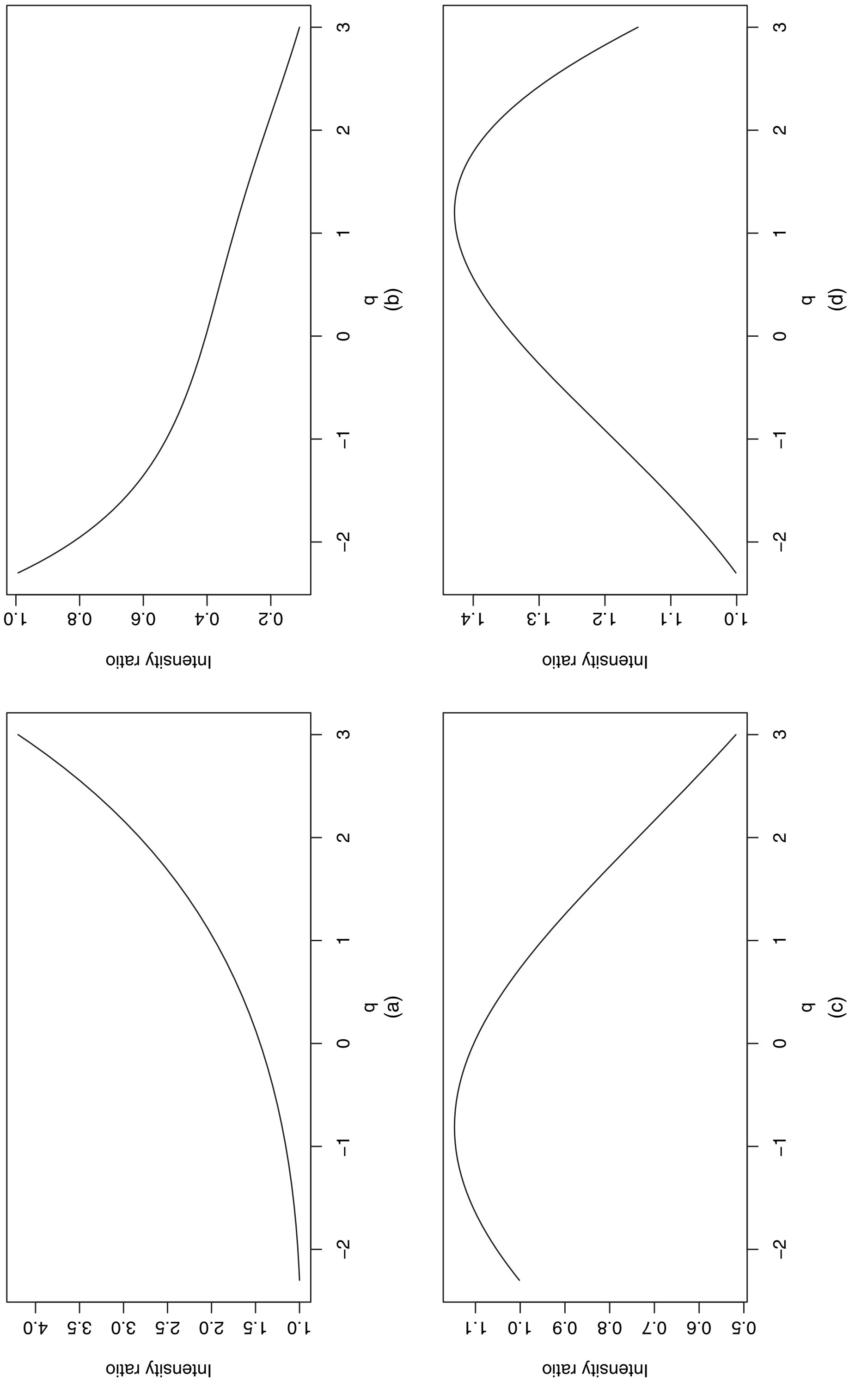


Fig. 3. Estimated effect of q, a measure of relative team ability, on scoring and booking rates (the functional form is based on a *B*-spline with two knots): (a) home team scoring intensity; (b) away team scoring intensity; (c) home team booking intensity; (d) away team booking intensity

Table 3. Estimated parameters for home and away team booking event rates†

Parameter	Results for home team yellow		Results for away team yellow		Results for home team red		Results for away team red		Results for home team red from yellow		Results for away team red from yellow	
	Estimate	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*
$\log(\lambda_{1l})$	−4.363	0.125	−4.344	0.121	−6.445	0.169	−6.429	0.158	−6.372	0.191	−6.486	0.178
$\log(\alpha_{1l})$†	0.567	0.035	0.566	0.030	0.567	0.035	0.566	0.030	0.567	0.035	0.566	0.030
$\log(\lambda_{2l})$	−4.172	0.183	−4.170	0.174	−7.241	0.226	−7.174	0.213	−7.133	0.262	−7.256	0.243
$\log(\alpha_{2l})$†	0.179	0.025	0.201	0.022	0.179	0.025	0.201	0.022	0.179	0.025	0.201	0.022
Home Ahead†	−0.062	0.047	−0.102	0.041	−0.062	0.047	−0.102	0.041	−0.062	0.047	−0.102	0.041
Away Ahead†	−0.009	0.051	−0.019	0.045	−0.009	0.051	−0.019	0.045	−0.009	0.051	−0.019	0.045
log (# Home Yellow)	0		0		0		0		1.153	0.273	0	
log (# Away Yellow)	0		0		0		0		0		1.390	0.207
Any Home Booking	0		0.217	0.038	0.390	0.192	0		0		0	
Any Away Booking	0.247	0.045	0		0		0.227	0.179	0		0	
q_{b1}†	0.315	0.483	0.201	0.447	0.315	0.483	0.201	0.447	0.315	0.483	0.201	0.447
q_{b2}†	0.058	0.235	0.635	0.202	0.058	0.235	0.635	0.202	0.058	0.235	0.635	0.202
q_{b3}†	−0.689	0.358	0.159	0.308	−0.689	0.358	0.159	0.308	−0.689	0.358	0.159	0.308
PremLeague†	0.119	0.028	0.119	0.028	0.119	0.028	0.119	0.028	0.119	0.028	0.119	0.028

†The same effect is assumed for all types of booking.

which the effect of q for the away team's bookings was constrained to be the same as $-q$ for the home team. However, there was substantial evidence against this model compared with the model allowing unconstrained effects for q for each team (LR = 19.4 on 3 degrees of freedom). r_3 did not have any significant effect on booking rates and is not included in the final model for booking events.

When considering the effect of league and season on booking rates, on correcting for q the rate of bookings was higher in the Premier League than in the Championship, being around 13% higher (95% confidence interval (CI) 7–19%) in comparable match situations. Since in the vast majority of cases there is little league overlap in the pool of officiating referees it is not clear whether this is due to systematic differences in individual or general refereeing policies between leagues or due to more aggressive play in the Premier League.

4.2. Model extensions

4.2.1. *Team style of play or aggression*

The model that was presented in Section 4.1 uses the relative strengths of the teams to inform their expected booking rates. However, there are likely to be other factors, such as team playing style and the perceived importance of the game, which may also affect the booking rate. To follow the principle that was applied to scoring rates, we would seek to use available bookmakers' odds for the number of bookings as a proxy for team and match importance effects. However, historic data on the available odds were not available. Instead, as a proxy for playing style we considered the mean number of fouls that had been committed by a team in the previous 10 league games. Using fouls rather than directly using bookings gives a more stable estimate of team aggression because typically a team might commit 10–15 fouls in any match whereas they typically receive fewer than three cards. We found that this aggression index did have some effect in explaining booking rates, with teams with a recent history of matches with more fouls tending to be awarded a higher number of bookings, after adjusting for team quality. This equated to an increase of 3.5% (95% CI 0.9–6.3%) and 5.6% (95% CI 3.2–8.0%) in the rate of bookings per additional foul for the home and away team respectively. We considered whether the presence of two teams with aggressive playing styles further impacted on the number of bookings, but this interaction effect was non-significant. To some extent this may be because escalation-type effects (i.e. the increase in the opposition's booking rate once a team has been booked) are already a feature of the model.

We also found that matches involving two teams from the 'big 5' Premier League clubs (Manchester United, Chelsea, Liverpool, Arsenal and Manchester City, which were the teams with the greatest chance of winning the Premier League in the two seasons featured) had significantly higher booking rates after adjusting for relative team ability and aggression index (28% higher than other Premier League matches (95% CI 10–50%)). There may be scope to extend effects of this nature to other subgroups of matches, for instance relegation matches. Similarly, Buraimo *et al.* (2010) found that derby matches, involving teams that are local to one another, are associated with higher booking rates. A problem with such a variable is identifying which teams have a genuine local rivalry.

4.2.2. *Red card effects excluding penalties or free kicks*

One further issue with the model that was presented in Section 4.1 is that bookings, particularly red cards, are often accompanied by a penalty or free kick that allows the opposition team an immediate goal scoring opportunity. The current model assumes a change in the scoring intensities that persists for the remainder of the match. Some of the observed increase in the goal scoring rate of the opposing team may be a short-term effect that is directly associated with penalties or free kicks occurring as a direct consequence of the booking. Ideally, one might attempt to make a distinction between bookings that result in a goal scoring opportunity and those that do not. In the absence of this information, we instead assume that the award of a red card has a short-term effect (lasting for 2 min after the event) and a long-term effect (lasting from 2 min after the event until the end of the match). To fit this model, we create a time-dependent covariate that takes value 1 if the time since the last red card awarded is less than 2 min and takes value 0 otherwise. If half-time occurs before 2 min have elapsed we assume that the effect ends at half-time. Fitting this model suggests that the home team have a 226% higher goal scoring rate in the 2 min after an away team red card has been awarded. Nevertheless, the persistent increase in home scoring rate is 83.5% (95% CI 59–112%). Similarly, the away team have a 323% increase in goal scoring rate in the 2 min after a home team red card has been awarded but an increase of 62.2% (95% CI 30–102%) persists. Hence the estimates of the persistent effects of red cards on opposing team scoring rates are very similar to those in the simpler model that was presented in Table 2.

4.2.3. Referee effect

Previous researchers have identified apparent discrepancies in individual referees' propensities to issue yellow or red cards (Dawson *et al.*, 2007). As there is no significant overlap between the referees who officiated Premier League games and those who officiated in the Championship, to estimate individual referee effects in each league we fitted separate models to the bookings for each league. In each case, referees who officiated fewer than 15 games in the Premier League or 20 games in the Championship were pooled as a single category 'other' which we used as the reference category in each case. For the Premier League inclusion of referee effects led to a statistically significant improvement in the model fit (LR = 40.9 on 18 degrees of freedom). However, a substantial proportion of the heterogeneity was due to a single referee, Mark Halsey, who appeared anomalously lenient, issuing bookings at a rate that was 46% lower (95% CI 26–60%) than the other Premier League referees. In the Championship there is strong evidence of heterogeneity (LR = 89.4 on 27 degrees of freedom), but with no particular outliers among the individual referees.

5. Model validation

5.1. Assessment of modelling assumptions

A major assumption of the model is that the baseline intensities are of Weibull form. The appropriateness of this assumption can be assessed by comparing the Weibull baseline hazards with those obtained through an equivalent Cox proportional hazards model, which makes no assumptions regarding the shape of the baseline intensities. If the Weibull assumption is appropriate then we would expect the estimates of the baseline cumulative intensities from each model to be similar and the regression coefficients of the two models to be nearly identical. Fig. 4 shows the cumulative baseline intensities for home and away goals and bookings for the Weibull and Cox models in each half of the match.

In most cases the Weibull model gives a very close fit to the intensity that is estimated by the Cox model. The exceptions involve the home team in the second half, where there are slightly more than expected numbers of goals and bookings between the 55th and 65th min of the match (10–20 min through the second half). The estimated regression parameters for the Weibull and Cox models were virtually identical. The proportional intensities assumption was assessed by

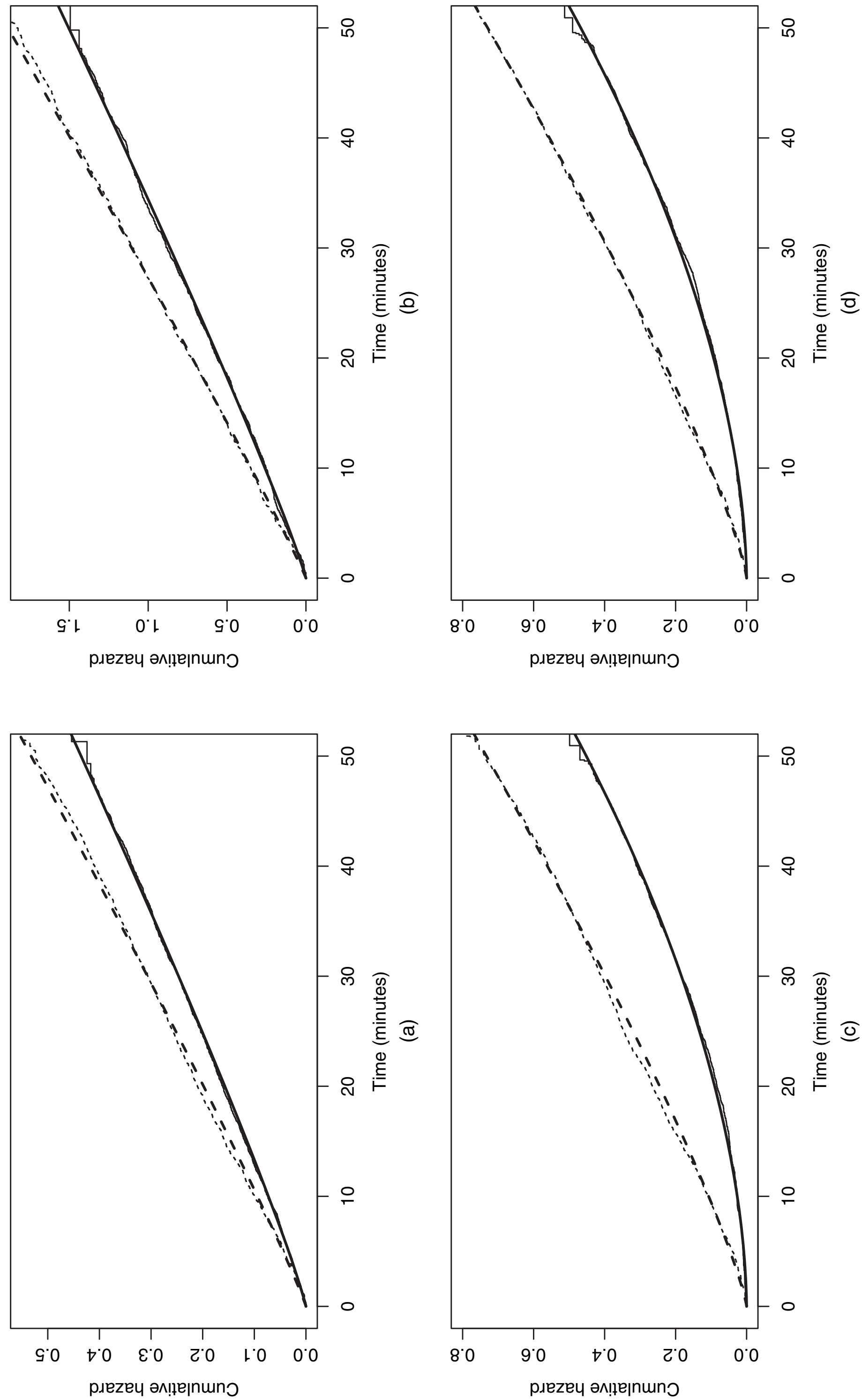


Fig. 4. Comparison of cumulative baseline intensities for home and away goals and bookings for Weibull (——, – – –) and Cox (——, - - - - -) models: (a) home team goal rates; (b) away team goal rates; (c) home team booking rates; (d) away team booking rates

considering the standardized Schoenfeld residuals (Grambsch and Therneau, 1994) of the Cox models. This suggested no evidence against the proportional intensities assumption.

5.2. Performance as an automated spread betting strategy

To illustrate the ability of the model to provide dynamic predictions of match outcomes in out-of-sample matches, we assessed the model's performance as the basis for an automated live spread betting strategy.

Spread betting allows participants to bet on a match outcome in two ways: buying or selling. Buying is a bet that the outcome will be higher than the buy price. Selling is a bet that the outcome will be lower than the sell price. Table 4, for example, gives the gain or loss that is incurred if someone buys or sells 1 unit of total goals at a price of 2.05. Note that selling at the same price is the exact converse of buying. In practice the buy price will always be higher than the sell price, with the difference being the so-called 'spread'. At a single instance in time, it is advantageous to bet only if the expected outcome lies outside the spread.

One method of allowing decisions to be made regarding the timing and magnitude of bets is to adopt a utility-based approach (Fishburn, 1970). If we were betting at a single point in time, e.g. before the start of the match, we would choose the action that maximizes the expected utility with respect to the estimated distribution of match outcomes. Let Y denote the outcome of a particular market in a match, and let $\rho(Y)$ denote the pay-out function given match outcome Y and bets placed. For instance, if we have bought 20 of $Y=$ total goals at a price of 2.05, our pay-out function is $\rho(Y)=20(Y-2.05)$. Our utility function is then some function $U\{\rho(Y)\}$. A common utility that is used in betting situations is a log-utility leading to 'Kelly betting' (Kelly, 1956). A practical issue with this utility function is that it assigns an unbounded negative utility to wealth of 0. It is possible to implement this to spread betting as in practice spread bets have an associated stop loss (a maximum loss that a bet can receive), so one could bet in such a way that the loss cannot exceed the current wealth. However, for simplicity we instead choose an exponential utility function:

$$U\{\rho(Y)\}=1-\exp\{-K\,\rho(Y)\}.$$

$K>0$ represents the degree of risk aversion, although for an exponential utility it serves only to alter the magnitude of bets rather than their timing. Owing to the constant absolute risk aversion property of exponential utility, the initial wealth is not relevant to the betting actions.

Table 4. Example to show the loss or gain from a spread bet on total goals scored

Total goals	*Gain*	
	Buy	*Sell*
0	−2.05	2.05
1	−1.05	1.05
2	−0.05	0.05
3	0.95	−0.95
4	1.95	−1.95
5	2.95	−2.95

Suppose that the current sell price is ϕ_S and the current buy price is $\phi_B > \phi_S$. Since $\rho(Y) = (Y - \phi_B)u$ if buying u units and $\rho(Y) = (\phi_S - Y)u$ if selling, we can express the action of buying or selling in terms of a single number u where $\rho(Y;u) = uY - \{u\}^+\phi_B - \{u\}^-\phi_S$ and $\{u\}^+ = \max(0,u)$ and $\{u\}^- = \min(0,u)$. Given that no previous bets have been made, the optimal bet at this time, given the utility function and the current estimated distribution of Y, is

$$u^* = \arg\max_u \mathbb{E}_Y[U\{\rho(Y;u)\}]. \tag{5.1}$$

However, the main potential benefit of a stochastic model for goals and bookings is in live betting. We may therefore wish to make multiple bets during a match. Suppose that we have already made I bets with stakes of u_i (where a negative u_i represents selling) at prices ϕ_i (representing a buying price if $u_i > 0$ and a selling price if $u_i < 0$); the current pay-out function is then $\rho(Y) = Y \Sigma u_i - \Sigma u_i \phi_i$. Given these past bets and current available prices ϕ_B and ϕ_S, the decision to buy or sell at the current time is based on maximizing the utility function with respect to u for a pay-out function

$$\rho(Y;u) = Y \sum u_i - \sum u_i\phi_i + uY - \{u\}^+\phi_B - \{u\}^-\phi_S,$$

over the current estimated distribution of the outcome Y. If $u^* \neq 0$ then we update the pay-out function with new stake $u_{I+1} = u^*$ and price

$$\phi_{I+1} = \begin{cases} \phi_B & \text{if } u^* > 0, \\ \phi_S & \text{if } u^* < 0. \end{cases}$$

It is also feasible to assume that we may wish to bet on several different outcomes from the same match at the same time, e.g. total goals and goal difference. We need to take into account that these outcomes may be correlated when deciding on bets. Let Y_j, $j = 1, \ldots, K$, represent the jth match outcome. Each outcome j has corresponding pay-out function $\rho_j(Y_j) = Y_j \Sigma u_{ij} - \Sigma u_{ij}\phi_{ij}$ where u_{ij} and ϕ_{ij} are the stakes and prices bet on outcome j at the ith time. The strategy then follows as in the single-market case except that we seek to find the $u_1^*, \ldots, u_K^*$ that maximize the expected utility function with respect to the pay-out $\rho(Y;\mathbf{u}) = \Sigma_j \rho_j(Y_j; u_j)$.

In practice, the expectation in equation (5.1) cannot be calculated in closed form. However, it is straightforward to simulate realizations of the outcome Y conditionally on the events that have occurred by the current time point. Given R simulated realizations $Y_1^*, \ldots, Y_R^*$, a Monte Carlo approximation of the expectation is given by

$$\frac{1}{R} \sum_{i=1}^{R} U\{\rho(Y_i^*;u)\}$$

and the optimal action is taken to be

$$u^* = \arg\max_u \frac{1}{R} \sum_{i=1}^{R} U\{\rho(Y_i^*;u)\}.$$

This can be optimized by using standard methods for numerical optimization, keeping the set of simulated Y^* fixed.

5.2.1. Application of betting strategy

The live spread betting prices listed at `www.sportingindex.com` for 93 Premier League or Championship matches that took place between October and December 2011 were recorded at 1-min intervals, which broadly corresponded to the frequency at which they were typically updated. We considered up to four markets: goal difference, total goals, total booking points

Table 5. Hypothetical gains from using the model to spread-bet on match outcomes goal difference, total goals, total bookings and booking difference based on an exponential utility with $K = 0.1$†

Method	*Overall*	*Goal difference*	*Total goals*	*Total bookings*	*Booking difference*
Fixed parameters	23.4	−12.9	−10.5	34.2	12.6
Approximate Bayes	26.2	−10.5	−12.7	35.8	13.6

†The overall hypothetical gain over all four markets is also tabulated

and booking points difference. Live betting on bookings was available in only 54 of the 93 matches. Given the available prices, we retrospectively applied the betting strategy, based only on models fitted to data before the 2011–2012 season. We used a model that incorporated the additional effects of the aggression index and 'big 5' effects and used the presence of Mark Halsey as referee as a covariate but no other referee effects (Section 4.2.3). The use of an exponential utility function allows the betting actions for different matches, some of which occurred simultaneously, to be calculated independently.

When choosing the optimal action we considered two approaches. Firstly we assumed that the estimated model parameters were known and fixed and hence simulated the Y^* via a model using the maximum likelihood estimates of the parameters. Secondly, we considered an approach which, for each realization $i = 1, \ldots, R$, first simulates a set of parameters θ^* from a normal distribution with mean equal to the maximum likelihood estimate and variance equal to the inverse of the Fisher information, and then simulated a Y_i^* from a model with parameter value θ^*. This approach can be considered to be an approximation to a fully Bayesian approach with flat prior distributions on the model parameters.

The hypothetical market gains based on the maximum likelihood estimate and Bayes approximation methodology are presented in Table 5. Overall, the model gained 23.4 on the basis of $K = 0.1$, with the gain per game range being (−8.7, 13.9) having a heavy-tailed distribution. The model appeared to be more successful with bookings markets, gaining 34.2 on total booking points and 12.6 on booking points difference. In contrast, overall total goals lost 10.5 and goal difference lost 12.9. Note that these bets include a 'bid–ask' spread and the overall profit would be 46 higher if the same bets could be bought or sold at the midpoint of the spread. The better performance for bookings markets is consistent with the notion that the bookings markets may be less well developed than those for goals. The Bayes approximation gave broadly similar, but slightly higher, gains.

An alternative utility function, taking

$$U(Y) = \begin{cases} \log(1 + KY) & \text{if } Y \geqslant 0, \\ 1 - \exp(-KY) & \text{if } Y < 0, \end{cases}$$

representing increasing risk aversion once an expected profit has been secured in an *individual* match, gave broadly similar results, but allowed a greater proportion of matches to achieve positive gains.

5.2.2. *Case-study 1: Manchester City versus Wolves*

As a first casing example we show the model-based estimates compared with available live spread betting prices for the Premier League match between Manchester City and Wolves on October

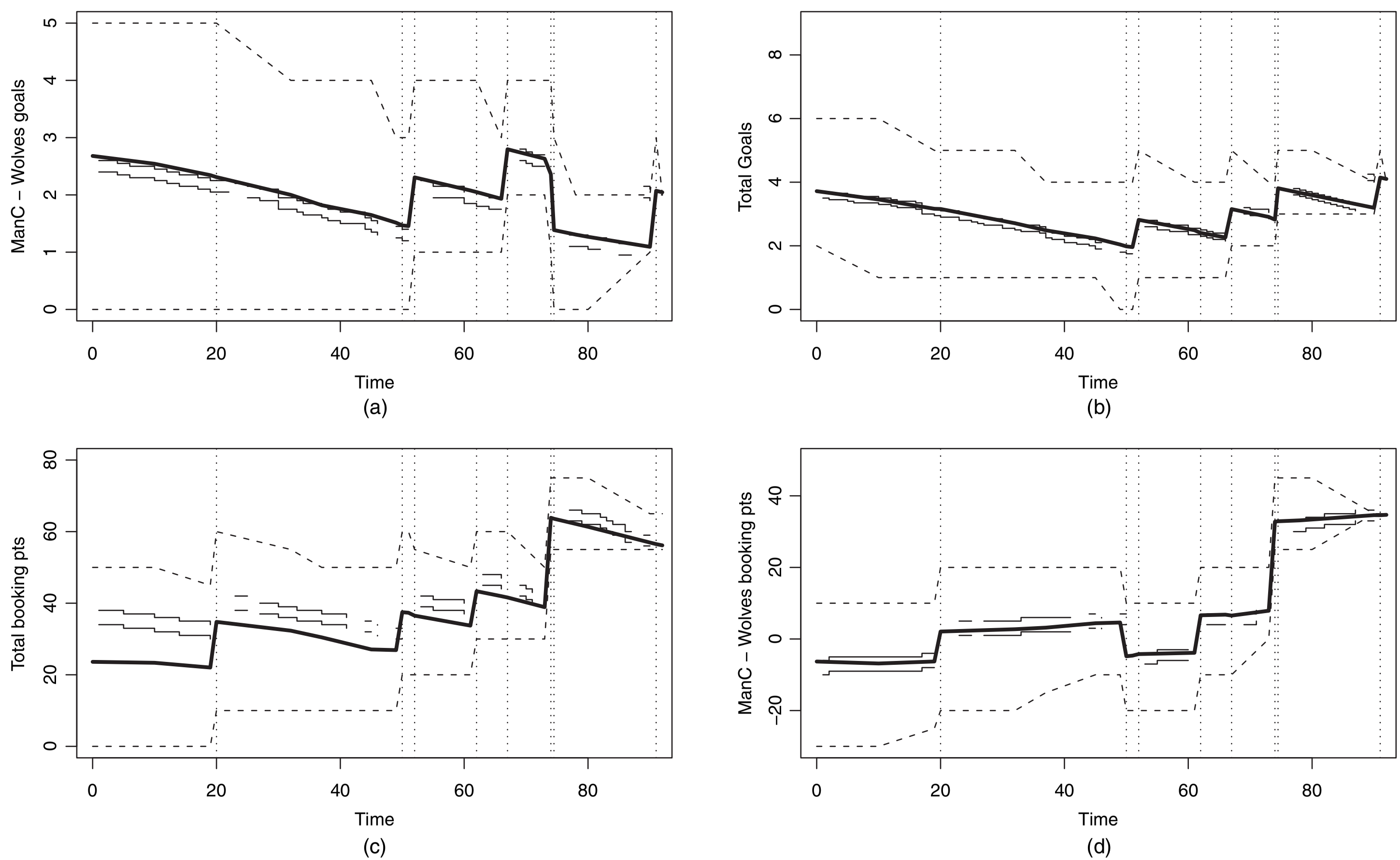


Fig. 5. Comparison of model-based live estimates and spread betting prices for Manchester City *versus* Wolves on October 29th, 2011 (——, buy and sell prices; ——, estimated mean from the model given the current value of the birth process; – – –, model-based 90% prediction intervals; ⁚, home team yellow card awarded at 20 min, away team yellow card awarded at 50 min, home team goal at 52 min, home team yellow card awarded at 62 min, home team goal at 67 min, home team independent red card awarded at 74 min, away team goal at 74 min (via a penalty) and home team goal at 91 min): (a) goal difference; (b) total goals; (c) booking points; (d) booking points difference

29th, 2011. Before the match, the available odds gave an estimate of $q=2.686$, implying that Manchester City are highly favoured and $r_3=0.04$ implying that the odds suggest a higher-than-expected chance of 3 or more goals. Prices for total goals and total bookings points were monitored through the match. Fig. 5 shows the dynamics of the prices and model estimates as the match progressed.

Initially there is reasonably broad agreement between the available spread and model-based estimates for all markets except total booking points where the available spread is higher than the model-based estimate. In periods where there are no events, both measures tend towards their current value with increasing gradient as the game progressed. The expected outcome jumps at each event. For instance the yellow card awarded at 20 min increases the expected number of booking points by around 15 points because of the estimated escalation effect but has a negligible effect on the expected goal difference. The betting strategy performs relatively poorly in this particular match because it initially estimated fewer bookings than the available spread in a match where a large number of bookings occurred.

5.2.3. Case-study 2: Queens Park Rangers versus Manchester City

As a second example we compare the model estimates with live spread betting prices for the Queens Park Rangers *versus* Manchester City match on November 5th, 2011. Before the match, the available odds gave an estimate of $q=-1.934$, implying that Manchester City were favoured, and $r_3=0.06$, implying that the odds suggest a higher-than-expected chance of 3 or more goals. Fig. 6 shows the dynamics of the prices and model estimates as the match progressed.

Initially, there is some disagreement between the model's predictions and the available spread betting prices. In particular, the model expects fewer bookings and more goals than the available spread. As the match progressed, the available spread and model predictions converged, with this occurring quickly for the goal difference market and much more slowly for the total booking points market. The betting strategy performs very well in this match because it correctly predicts a match with many goals but few bookings. Interestingly after the goal by Queens Park Rangers at 27 min, the spread betting prices reflect an expectation of higher total booking points in the match, whereas the statistical model suggests a small reduction in the expected booking points.

6. Conclusion

This paper presents a flexible approach for the investigation of association football event interdependences which allows for the dynamic prediction of match outcomes as events unfold. Specifically, influential events can be quantified and match predictions updated in realtime. The model can be readily extended to incorporate additional match events and can be further developed as a tool for participation in on-line spread betting markets.

The eight-dimensional counting process formulation for the joint modelling of home and away team goals and bookings provides a means of greater insight into event interplay. More specifically, existing approaches use bivariate model specifications, for example, to consider event subsets such as the effect of expulsion (a red card) on subsequent goal rates or the dependence of the goal rate on the current score. Also the generic model presentation in Section 3 provides added clarity in terms of possible modelling extensions and potential utility. In particular, previous models have been parameterized and presented in a manner that is specific to the modelling problem at hand. For example, the bivariate model formulation for goals in Dixon and Robinson (1998) included specific terms representing attacking and defensive strength, home team advantage, time of play and match state.

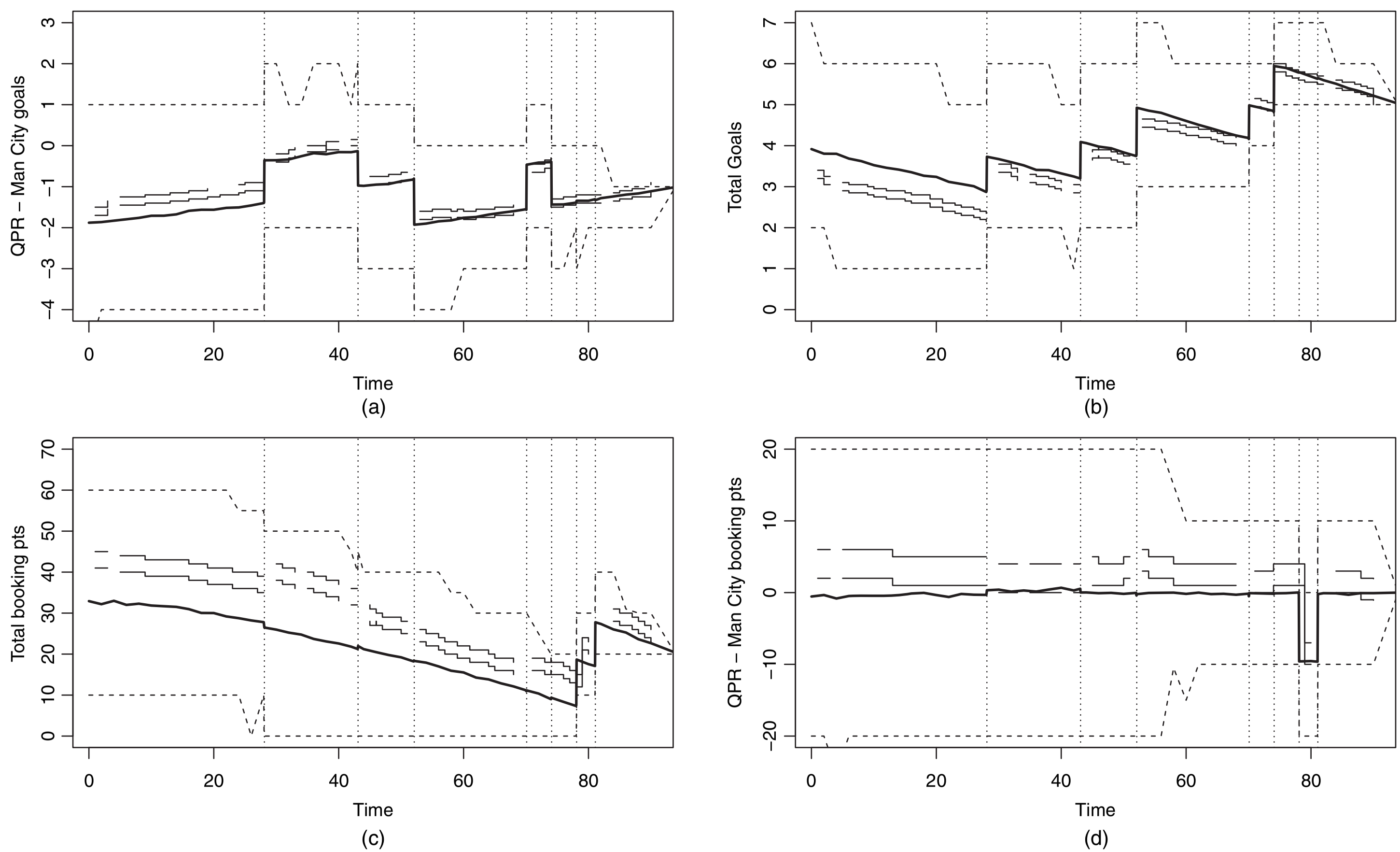


Fig. 6. Comparison of model-based live estimates and spread betting prices for Queens Park Rangers *versus* Manchester City on November 5th, 2011 (——, buy and sell prices; ——, estimated mean from the model given the current value of the birth process; – – –, model-based 90% prediction intervals; ⋮, home team goal at 27 min, away team goal at 42 min, away team goal at 51 min, home team goal at 68 min, away team goal at 73 min, away team yellow card awarded at 77 min and home team yellow card awarded at 80 min): (a) goal difference; (b) total goals; (c) booking points; (d) booking points difference

In terms of model findings, relative team ability, proxied via bookmakers' prior match odds, was found to have a significant effect on booking rates with a team's bookings peaking when the opposing team is slightly favoured. The award of yellow cards does not appear to have any direct effect on goal scoring rates, whereas red cards have a substantial detrimental effect, in particular, when the away team is reduced to 10 men. Once a team has been awarded a yellow card, the risk of a red card for a first bookable offence among fellow team members more than doubles. For the opposing team the booking rate also increases significantly, averaging 25%.

The application of the model to spread betting resulted in reasonable gains in markets relating to bookings but slight losses for goal markets. As we use bookmakers' prematch odds of match outcomes to inform goal rates, we would not expect our model to be able to outperform the bookmaker's predictions at the onset of a match. In most cases, the model prediction of goals will lie within the spread at the start of a match. When this does not occur it is likely to be because the prematch covariates q and r_3 are insufficient to capture all the bookmaker's prior information. In contrast, the model predictions for bookings do not rely on any direct bookmaker prediction of bookings but instead use our model of the relationship of relative team ability and number of fouls in previous matches on booking rates. Moreover, the market for betting on bookings seems to be a much less developed. As a consequence, our model has potential to outperform the bookmaker's predictions of bookings both at the onset of the match and subsequently. If the primary aim of the model were to spread-bet in practice, particularly for goal markets, it would be necessary to extend the model to use a greater range of prematch covariates indicative of form, independently of the bookmakers.

A large proportion of bookings in matches are as a direct consequence of fouls. It is likely that, in the same way that it is rare for the first booking of a match to be a red card, there will usually be some non-bookable fouls before the first yellow card is awarded. A more refined model, assuming that data were available, would model the fouls process, taking yellow and red cards as special types of foul event.

Weibull intensities were chosen for parsimony and to enable straightforward simulation of the model for prediction. Although the model appears to fit reasonably well there is some evidence of extra activity, in terms of both goals and bookings, for the home team during part of the second half. A straightforward model extension would allow the baseline hazards to be smooth non-parametric functions.

After the initial booking in a match, we find that subsequent bookings become more likely. To some extent, this could be explained by the existence of a shared match frailty affecting the booking event intensities. The most plausible model would perhaps be a model assuming both a frailty effect and event-specific events. As noted by Putter and van Houwelingen (2011), it is not necessarily possible to identify fully between two such models. Moreover, from the perspective of predicting outcomes the models will be similar—in the frailty model the award of a yellow card informs us about the frailty effect and hence the posterior estimate of the intensities at future times will increase (Aalen, 1987).

The model can quantify the effect of a particular refereeing decision on the outcome of a match but only with the strong assumption that the effect of a decision is independent of its perceived correctness. It seems quite plausible that the effect of a contentious red card decision would be different from cases where the decision to send a player off was clear cut. If contentious or incorrect decisions could be identified, perhaps by expert judgement, there could be scope to extend our model by increasing the number of match events to distinguish between such decisions. However, given the rarity of such events it is likely that data on a substantially greater number of matches would be required to estimate these effects adequately.

The betting strategy for on-line spread betting was based on optimizing the expectation of an exponential utility function. Exponential utility implies a constant level of risk aversion which might be realistic only for an individual with very large funds making a large number of bets. The methods applied can be easily adapted to include other utilities—although for functions such as log-utility a 'stop loss' (maximum amount that can be lost from a bet) needs to be factored in as the log-utility assigns infinite negative utility to becoming bankrupt. If betting on multiple games simultaneously, using utilities other than exponential would also require factoring in the positions taken on all the matches simultaneously.

The focus of this paper has been to investigate association football event interdependences, with a view to quantifying interactions between event processes and to provide a dynamic means of prediction. Notwithstanding this, the approach that is taken to modelling and decision making would also be broadly applicable to other football leagues and to some other sports. Consider, for example, rugby union. The distinct events of interest would include home and away team tries, converted tries, drop goals and disciplinary sanctions ('expulsions' and 'sin binnings'). The model and likelihood formulation would be broadly similar with appropriate changes to the static and time varying covariates.

More generally, an interacting counting process framework may be useful in financial or medical applications exhibiting similar dynamic properties. For instance, 'wait-listed' kidney transplant patients may be urgently called from an elective population years after diagnosis of chronic renal failure. During this period significant new medical issues may impact their operative and post-operative course. In particular, intercurrent events such as cardio-vascular disease, peritonitis, viral hepatitis, cancer, blood pressure and obesity may affect transplant candidacy, the perioperative and, or, post-operative course and implicate increased risk of comorbidity (Danovitch *et al.*, 2002; Wu *et al.*, 2005). Realtime monitoring of wait-listed candidates may thus lead to more informed selection procedures and also lead to reduced numbers of rescheduled surgeries and reduced cold ischaemia time.

References

Aalen, O. O. (1987) Mixing distributions on a Markov chain. *Scand. J. Statist.*, **14**, 281–289.
Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993) *Statistical Models based on Counting Processes*. New York: Springer.
Broström, G. (2014) eha: Event history analysis. *R Package Version 2.4-1.*
Buraimo, B., Forrest, D. and Simmons, R. (2010) The 12th man?: refereeing bias in English and German soccer. *J. R. Statist. Soc.* A, **173**, 431–449.
Danovitch, G. M., Hariharan, S., Pirsch, J. D., Rush, D., Roth, D., Ramos, E., Starling, R. C., Cangro, C. and Weir, M. R. (2002) Management of the waiting list for cadaveric kidney transplants: report of a survey and recommendations by the clinical practice guidelines committee of the American Society of Transplantation. *J. Am. Soc. Nephr.*, **13**, 528–535.
Dawson, P., Dobson, S., Goddard, J. and Wilson, J. (2007) Are football referees really biased and inconsistent?: evidence on the incidence of disciplinary sanction in the English Premier League. *J. R. Statist. Soc.* A, **170**, 231–250.
Dixon, M. J. and Robinson, M. E. (1998) A birth process model for association football matches. *Statistician*, **47**, 523–538.
Fishburn, P. C. (1970) *Utility Theory for Decision Making*. New York: Wiley.
Forrest, D., Goddard, J. and Simmons, R. (2005) Odds-setters as forecasters: the case of English football. *Int. J. Forecast.*, **21**, 551–564.
Grambsch, P. M. and Therneau, T. M. (1994) Proportional hazards tests and diagnostics based on weighted residuals. *Int. J. Forecast.*, **81**, 515–526.
Kelly, J. L. (1956) A new interpretation of the information rate. *Bell Syst. Tech. J.*, **35**, 917–926.
Pollard, R. (2008) Home advantage in football: a current review of an unsolved puzzle. *Open Sprts Sci. J.*, **1**, 12–14.
Putter, H. and van Houwelingen, H. C. (2011) Frailties in multi-state models: are they identifiable?; do we need them? *Statist. Meth. Med. Res.*, to be published.
R Development Core Team (2009) *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ridder, G., Cramer, J. S. and Hopstaken, P. (1994) Down to ten men: estimating the effect of a red card in soccer. *J. Am. Statist. Ass.*, **89**, 1124–1127.
Vecer, J., Kopriva, F. and Ichiba, I. (2009) Estimating the effect of the red card in soccer: when to commit an offense in exchange for preventing a goal opportunity. *J. Quant. Anal. Sprts*, **5**, article 8.
Volf, P. (2009) A random point process model for the score in sport matches. *IMA J. Mangmnt*, **205**, 121–131.
Wu, C., Evans, I., Joseph, R., Shapiro, R., Tan, H., Basu, A., Smetanka, C., Khan, A., McCauley, J. and Unruh, M. (2005) Comorbid conditions in kidney transplantation: association with graft and patient survival. *J. Am. Soc. Nephr.*, **16**, 3437–3444.