Journal of Quantitative Analysis in Sports

Volume 5, *Issue* 4 2009 *Article* 5

Cruel and Unusual Punishment? An Analysis of Point Deduction in European Association Football Leagues

Nik J. Cunniffe* Alex R. Cook†

*University of Cambridge, njc1001@cam.ac.uk
†National University of Singapore, stacar@nus.edu.sg

Cruel and Unusual Punishment? An Analysis of Point Deduction in European Association Football Leagues*

Nik J. Cunniffe and Alex R. Cook

Abstract

We present a methodology to assess the effect of point deductions, which are widely applied as punishments to teams in European association football (soccer) leagues. In particular we outline a method to estimate the increased probability of relegation and the decreased probability of promotion that are associated with a certain penalty for a given team in a particular league. We illustrate our method in the context of two case studies involving points deductions: the relegation of Luton Town from English League Two in 2008–9, and the promotion of Juventus from Italian Serie B in 2006–7. By adapting the parameterisation of our model, we extend our analysis to the more challenging objective of designing standardised point deductions, and discuss how prior knowledge on performance influences the likely severity of any given punishment. We also investigate the relationship between the magnitude of any point deduction and the financial implications for the penalised team. Our techniques allow unambiguous quantification of the consequences of a form of punishment that has been widely, but not necessarily scientifically, implemented by a number of governing bodies in European association football.

KEYWORDS: European soccer, Poisson model, point deduction, league competitions, English League Two, Italian Serie B, Luton Town, Juventus

*We are grateful to Stephen Parnell and Sanjay Chaudhuri for their suggestions and comments on the work. Cook acknowledges research funding from the National University of Singapore.

1 Introduction

In European domestic association football (henceforth "football"), the best teams in each country are often arranged in an hierarchy of professional leagues. Within a particular league, pairs of teams play each other a number of times (generally once at home, i.e. at their own stadium, and once away, i.e. at their opponent's home stadium, although variations on this theme are possible), and receive a pre-specified number of points for each match they win, lose or draw. At the end of each season, and based on position in a final league table ordered by the number of points obtained, sometimes with further play-off matches, relegation and promotion shuffle teams between the different leagues. As well as gratifying supporters, promotion to a higher league can be very lucrative, leading to improved attendances, better deals with sponsors, the possibility of attracting better players, opportunities to play in more prestigious cup competitions (including Europe-wide competitions such as the Union of European Football Associations' Europa and Champions' Leagues), and potentially a larger share of any income that the controlling football association has negotiated with television companies. Relegation to a lower league is financially harmful for exactly the opposite reasons [Noll, 2002, Dobson and Goddard, 2001, Deloitte, 2009].

In many leagues, a punishment that may be applied by the governing body is point deduction, in which a number of points is removed from a team's total (either within a particular season or at the start of the next) in response to some perceived misdemeanour. In English football these penalties are routinely applied for financial irregularities, with a fixed ten point deduction for teams that enter Administration (similar to Chapter 11 Bankruptcy in the United States) having been widely and regularly applied since the 2004–5 season. The reasoning behind these penalties is that allowing teams to enter/exit Administration without punishment could be seen as a form of financial "doping" whereby a team could, theoretically, irresponsibly overspend in search of an on-pitch advantage, then, if that failed, clear its debts and emerge reborn and unscathed [Deloitte, 2009]. However point penalties are also applied in other European leagues, both for financial and other reasons. A particularly evocative example comes from Italian football: in 2006 team managers were found to have attempted to influence referee appointment in an extremely well-publicised match-fixing scandal, in which world-famous teams including Juventus, Milan and Lazio were implicated.

Point penalties, although arguably deserved, can appear to the casual observer to be decided on an ad-hoc basis. Even when there is consistency within a particular league or national hierarchy of leagues, the penalty for similar offences can vary wildly by country. It is obvious therefore to ask whether a particular penalty is fair. The most natural way to decide the severity of the penalty is in terms of its effect on league placing and in particular vis-à-vis relegation and promotion. Determining and imposing a penalty retrospectively once the season has finished, however, leaves sporting decision makers open to accusations of arbitrariness or favouratism. There is a clear need for a predetermined, standardised penalty. Since the effect of such a point penalty cannot be determined in advance, its effect must be quantified probabilistically.

We address this using mathematical modelling. Our basic approach is to model the score in an individual game of football, using a Poisson model for the number of goals scored by each team. We then use this within-game model to quantify the probabilities of particular outcomes in a sequence of games and therefore in an entire league season. By artificially altering the equivalent of a league table that results from many iterations of our model, we can then investigate the effect of any penalty applied to any team. We use the model to ask the following questions.

- Can this modelling approach be used to determine the probability of any unfavourable outcome that would have resulted from a particular point deduction for a given team in a certain league, and so to assess reterospectively the "fairness" of historical punishments in particular cases?

- Can the model be used to design standardised punishments that are "fair" without reference to named teams or particular seasons, and so that could be applied in the future, where the standardisation is with respect to the expected severity over all teams?

We illustrate our questions in the context of two particular case studies.

- The relegation of Luton Town, the oldest professional football team in southern England, from League Two[1] at the end of the 2008–9 season, after a thirty point deduction at the beginning of the season. This point penalty was unusually large, with a deduction of ten points for entering Administration and twenty points for "improperly" exiting it, and was widely criticised for its perceived harshness by the UK sporting media. The deduction undoubtedly contributed to Luton Town being relegated

[1]There has been a confusing history of sponsorship-driven name changes in English football over the past fifteen years or so; to clarify, in 2008–9 "League Two" was actually the fourth highest league. In Italy "Serie B" was the second tier of Italian football in 2006–7, and has been since it came into existence in the 1930s.

DOI: 10.2202/1559-0410.1214

at the end of the football season, and so dropping entirely from the Football League.

- The promotion of Juventus from the Italian Serie B as champions in the 2006–07 season, having started that season after enforced relegation from Serie A and with nine points deducted. This penalty was a consequence of their involvement in the match-fixing scandal that rocked Italian football the year before, and was reduced twice on appeal from the original (far more severe) punishment. That the punishment was reduced from the originally proposed relegation to Serie C1 (i.e. the third league in Italian football) and that Juventus managed to return to the top tier of Italian football within a single season was the cause of much discourse.

2 Methods

2.1 Model formulation

We follow Maher [1982] and Lee [1997] and model the scores in an individual game by assuming that the numbers of goals scored by the home and away teams are independent Poisson variables. However we parameterise our model differently, and in what we feel is a more transparent manner. In particular we introduce two parameters, λ and μ, which are the mean numbers of goals scored by a "typical" team when playing at home or away respectively. Heterogeneity between teams is characterised by multiplicative "attack" (α_i) and "defence" (β_i) factors, which are related to the team's strength, and which scale the mean number of goals in any game involving team i. If team j plays against team k, with team j playing at home, we therefore model the number of goals J and K scored by teams j and k respectively as

$$J \sim \text{Poisson}(\alpha_j \beta_k \lambda), \tag{1}$$
$$K \sim \text{Poisson}(\alpha_k \beta_j \mu), \tag{2}$$

emphasising that higher values of α_i and/or lower values of β_i correspond to stronger teams. This framework is the generalised linear model with Poisson distributed data and exponential link function common in statistics [McCullagh and Nelder, 1989].

We note that this model is not novel, with different parameterisations of the same having been introduced by Maher [1982] and seemingly independently reinvented by Lee [1997]. In fact a number of more complex models elaborating this basic approach have more recently been proposed [Dixon and Coles, 1997, Dixon and Robinson, 1998, Karlis and Ntzoufras, 1999, 2003, McHale and Scarf, 2007, Everson and Goldsmith-Pinkham, 2008, Everson, 2008], the majority of which seek to provide a more accurate description of the joint distribution of home and away goals. However we contend that, in the interests of parsimony, this relatively simple model is sufficient for our purposes, and has the dual advantages of being well-tested and easy to describe and understand. We will return to these issues of model choice, presentation and selection in the Discussion.

2.2 Fitting the model

A number of repositories make football results freely available, and in particular the website `http://www.football-data.co.uk/` makes extensive data for many years and for a range of countries available in machine-readable format. For each season corresponding to one of our case studies (see Introduction), we model the attack and defence parameters for each team, together with the two parameters controlling the performance of a typical team, as fixed properties that remain constant for the entire season. In a league of N teams, there are correspondingly $2N + 2$ parameters; N attack parameters $(\alpha_1, \alpha_2, \ldots, \alpha_N)$, N defence parameters $(\beta_1, \beta_2, \ldots, \beta_N)$, and the pair of parameters controlling the baseline number of home and away goals (λ and μ). As the model is otherwise over-determined, we follow Karlis and Ntzoufras [2003] in using sum constraints to ensure identifiability. In particular we constrain both α_i and β_i so that $\prod_i \alpha_i = \prod_i \beta_i = 1$; that is, on the log-scale the parameters sum to zero. This has the advantage of giving these parameters a clear interpretation as multiplicative ratios, controlling how the number of goals scored by a particular team's attack, and the number of goals conceded by a team's defence, compared to those scored/conceded by the typical team.

We estimate the $2N$ free parameters that remain within the Bayesian paradigm [Lee, 2009], using Markov Chain Monte Carlo integration to sample the joint posterior distribution of the parameters using random scan Metropolis Hastings proposals on the natural logarithm of a single parameter per jump [Gilks et al., 1995]. To ensure the constraints $\prod_i \alpha_i = \prod_i \beta_i = 1$ are met, no explicit proposals are made to α_1 or β_1; instead, when a proposal is made to $\alpha_{i>1}$ a deterministic proposal is made to α_1 that maintains the constraining equality, and similarly for β. In the interests of pseudo-objectivity we take independent, flat, improper priors on the log parameters.

DOI: 10.2202/1559-0410.1214

2.3 Using the model to assess point reductions for particular teams: assessing fairness in our case studies

Given the fitted model we can calculate the posterior probability of any particular result in any individual match, and thereby determine the probability of every possible sequence of results over the entire season for every team in a given league, in turn quantifying the probability of any league table at the end of the season with or without the application of point penalties. This is most easily done via Monte Carlo simulation, in which a large set of parameter values are drawn from their joint posterior distribution within teams and used to simulate a series of seasons, with each of the $2N-2$ games in a season (i.e. each team meeting each other team twice, once at home and once away) independently simulated. Each individual game is simulated by assigning a win, a loss or a draw according to the Poisson model in Equations (1) and (2). By simulating a large (10 000) number of seasons in this fashion, the probability of any outcome for any team can be obtained with great accuracy. By updating the final league tables that result[2] according to any point reduction for our focus team(s), these probabilities can be made conditional on the size of the punishment, allowing us to determine how the probability of any outcome for each of our focus teams depends on the size of any penalty. In both of the leagues we consider, promotion (and relegation) is automatic for teams that finish the season sufficiently high (or low) in the final table, but for the teams just below (or above) the cut-off point the outcome depends on a sequence of "play-off" matches between teams near the top (or bottom) of the league. These were simulated independently for each season with parameters according to the particular teams involved in the play-off matches during that iteration.

2.4 Designing fair penalties

The previously-described approach allows us systematically to estimate the probability of any particular outcome after the football season has happened, and therefore to assess reterospectively the impact of a range of historical penalties. A far more challenging objective is to determine suitable pun-

[2] In the context of our case studies we also assumed that any other point deductions apart from that we were interested in were applied unchanged, penalising Darlington by ten points and penalising Bournemouth and Rotherham United both by seventeen points in our simulations of the 2008–9 English League Two season, and similarly penalising Triestina and Pescara by one point and Arezzo by six points in our investigation of the 2006–7 Italian Serie B.

ishments in advance. The main difficulty is that the probability of any unfavourable event, and so what may be considered an unreasonable punishment, depends not only upon the size of the punishment, but also on the strength of the team that is penalised and the quality of its opposition.

One way of normalising for this is to devise a notion of a "typical" season, and to investigate penalties for teams of varying strength relative to that. In particular, instead of sampling once from the posterior distribution of parameters for each team to create our league, on each iteration we choose a random sample of teams from the N available with replacement (i.e. probably duplicating certain teams), and sample the α_i and β_i parameters from the joint posterior densities of the teams in question. On each iteration we then choose our focus team at random, and consider the effect of any point reduction relative to the performance of that team. In this fashion we determine the impact of any penalty on a "typical" team regardless of its strength. Note that we do not impose any penalties to other teams in this case, although this could be effected easily if joint penalties were being considered.

A particularly controversial aspect of the penalty applied to Juventus was its size compared to the historical strength of the team. Historically the most successful team in Italian football, Juventus would naturally be expected to dominate the Serie B season. Therefore we also consider how the impact of any punishment varies with the strength of the team involved. Specifically we follow the above approach on each iteration to populate a typical league, but instead of randomly choosing the focus team, we choose from the subset of teams which were within the top six strongest or weakest (quantified by position in the final real-world league table). Our personal standpoint is that a fair penalty should be the same regardless of the strength of the transgressing team; however, it is clear that if teams vary in quality then the implications of the penalty will differ accordingly. In this way one can consider the potential effects of a single proposed penalty on different strata of teams.

In addition to investigating the effect of any point penalty on the probabilities of promotion and relegation, we also demonstrate how it is possible within this framework to quantify the expected financial cost of a particular point penalty. We assign to promotion and relegation the multiplicative payoffs a and b, where $a > 1$ is a measure of the proportionate expected gain for a promoted team (relative to the baseline of staying in the same league for the next season) and $b < 1$ is similarly a measure of the loss associated with relegation. Calculating appropriate values of a and b is a highly challenging econometrical issue (see e.g. Dobson and Goddard [2001], Noll [2002], Dobson and Goddard [2004], Szymanski and Smith [1997]), so here we take arbitrary but plausible values of a and b.

DOI: 10.2202/1559-0410.1214

We then define the expected cost of an N point penalty, $\phi(N)$, as

$$\phi(N) \quad = \quad a[p(0) - p(N)] + b[r(N) - r(0)] + 1[q(0) - q(N)], \qquad (3)$$

where $p(N)$ is the probability of promotion conditional on an N point deduction, $r(N)$ is defined similarly with respect to relegation, and $q(N) = 1 - p(N) - r(N)$ is the probability of neither, i.e. of maintaining the status quo. The expected change in revenue is made relative to the baseline performance with no penalty applied (i.e. $N = 0$).

3 Results

3.1 Relegation of Luton Town in 2008–9

The posterior distribution of the parameters of the fitted model for the 2008–9 English League Two season, and the expected impact of differently-sized point penalties upon Luton Town, are shown in Figure 1. With the exception of the false outliers caused by imposed penalties (e.g. Luton Town and Bournemouth), the defence ratio (β) increases with final position in the real-life league table (Figure 1b), although the relationship between league position and the attack ratio (α, Figure 1a) is less marked, particularly at the top end of the table, suggesting that in this league a relatively impervious defence is more important in terms of promotion than a prolific attack. The negative correlation between the point estimates of α and β for individual teams (Figure 1c) indicates that strong teams generally tend to do well in both attack and defence, whereas weaker teams are successful in neither.

To assess the goodness-of-fit of the fitted model, we calculated the posterior predictive distribution of the number of points each team obtained from a Monte Carlo sample of 10 000 realisations of the 2008–9 English League Two season, and for each team compared this distribution to the actual number of points obtained. The results are presented in Figure 2a. For all teams, the actual number of points fell well within the predicted range. To standardise the comparison across teams, the posterior predictive cumulative distribution function evaluated at the actual points is calculated and plotted in Figure 2b: if the data come from the model these should be uniformly distributed on $[0, 1]$. If anything, the model fits *too* well, with fewer extreme points near 0 or 1 than we would expect.

The effect of differently-sized point deductions on Luton Town's probability of relegation is very marked (Figure 1d). The relationship between the number

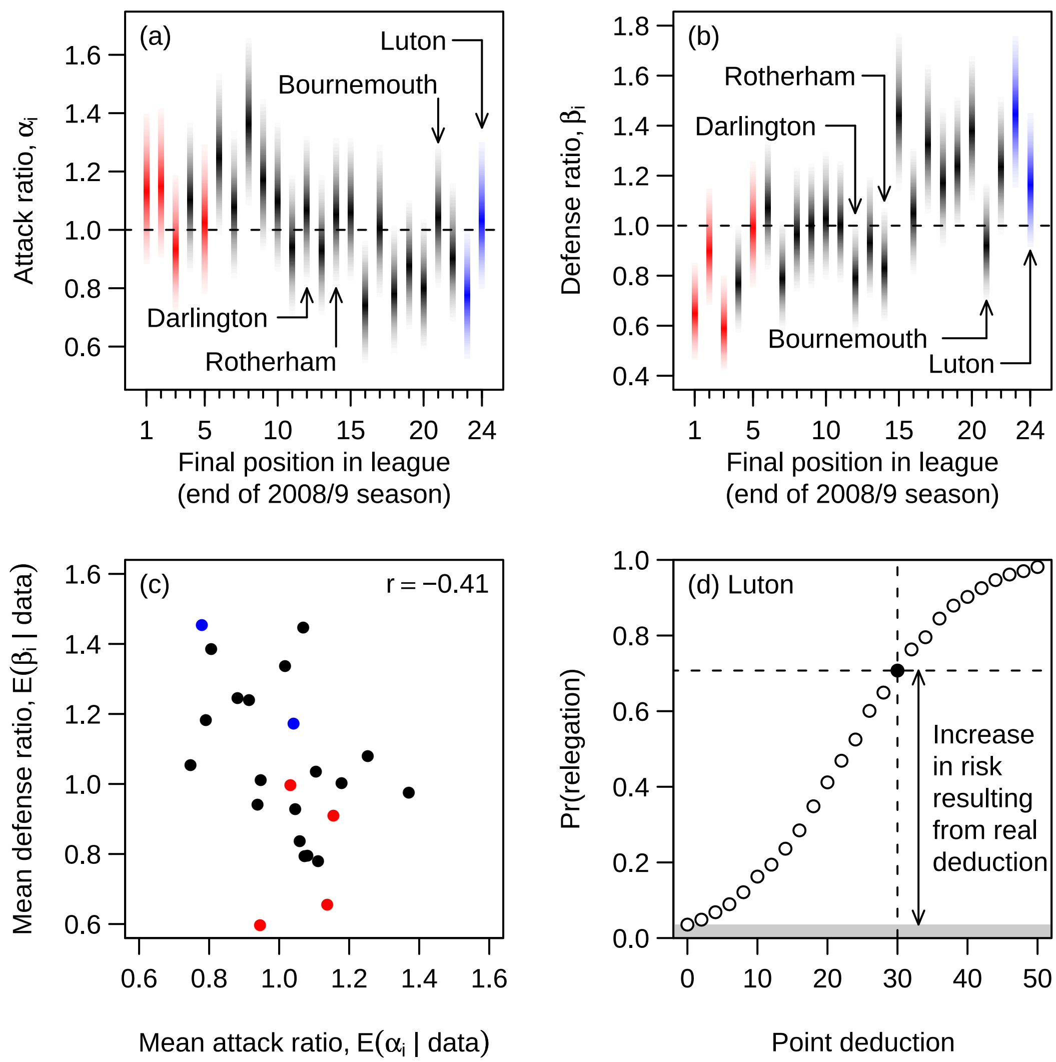

Figure 1: 2008–9 English League Two season. (a) Posterior distribution of attack ratios. Promoted teams are in red, relegated teams in blue, and teams with points penalties are indicated by arrows. Shades represent varying degrees of credibility, with the outermost regions relating to 95% credible intervals and the darkest shade the median. (b) Posterior distribution of defence ratios, as (a). (c) Correlation between posterior mean attack and defence ratios indicates that good teams tend to do well in both and bad teams in neither, although there are teams that are good at one skill only. Red points indicate promoted and blue relegated teams. Note that the spread of defence ratios is greater than that of attack. (d) The probability that Luton would have been relegated as a function of the imposed points deduction. With no points deduction, in a typical season, the probability of relegation is indicated by the gray box.

of points deducted and the probability of relegation is sigmoidal, with the inflexion point corresponding to a deduction of approximately twenty-five points. Consequently, with the penalty that was actually applied, the probability of relegation is greatly increased, from 0.03 with no punishment to 0.71 after a thirty point deduction. Although relegation by no means became certain, such an extreme punishment arguably did not give Luton Town a sporting chance in the 2008–9 League Two season. Note that due to the sigmoidal shape, the probability of relegation with a thirty point deduction is higher than the sums of the probabilities corresponding to the individual ten and twenty point deductions issued to Luton Town. For such a large cumulative point deduction, it appears that the whole is greater than the sum of its parts.

3.2 Promotion of Juventus in 2006–7

Our investigation of the 2006–7 Italian Serie B season is illustrated in Figure 3. The trends in attack (Figure 3a) and defence (Figure 3b) ratios according to league position are unsurprising, although these results emphasise the strength of Juventus relative to other teams in Serie B. Unsurprisingly given Juventus' dominance, the patterns of expected probability of promotion and relegation with size of point deduction (Figure 3c) are very different to those for Luton Town. The nine point deduction which was actually levied leads to a rather small reduction in the probability of promotion, from 0.92 to 0.76, and does not lead to any increase in the (zero) probability of relegation. Comparing this with the impact of the same-sized nine point punishment on a random team (Figure 3d), where the probability of promotion is roughly halved and the probability of relegation similarly doubled, emphasises just how atypical Juventus were, while at the same time suggesting that the nine point penalty is actually a fairly strong penalty when applied generally in Serie B. Even with the punishment of seventeen points that was reduced on appeal, we estimate that Juventus would have had a 0.54 probability of winning promotion.

3.3 Designing fair penalties

As discussed in the Introduction, while a retrospective investigation of the fate of any particular team is an interesting application of mathematical modelling, more pressing and difficult is the design of fair penalties. Related results are summarised in Figure 4, which shows the impact of differently-sized point penalties upon randomly selected, weak and strong teams in English League Two 2008–9 Season. Comparing the response for a random team in League

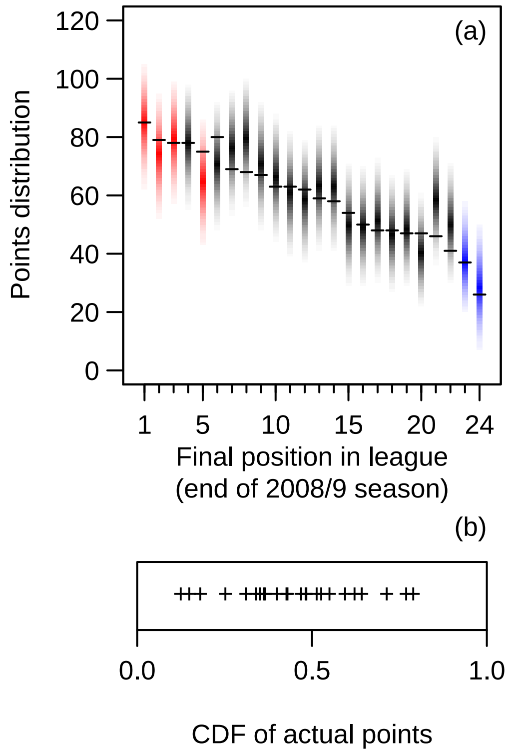

Figure 2: Goodness of fit of the model for English League Two in 2008–9. (a) Posterior predicted distribution of points for English League Two (gradated credible regions) and actual points (horizontal bars). (b) Cumulative posterior distribution function of actual points, which should be uniformly spread on the interval (0,1). With fewer outlying points than expected, our model fits arguably too well.

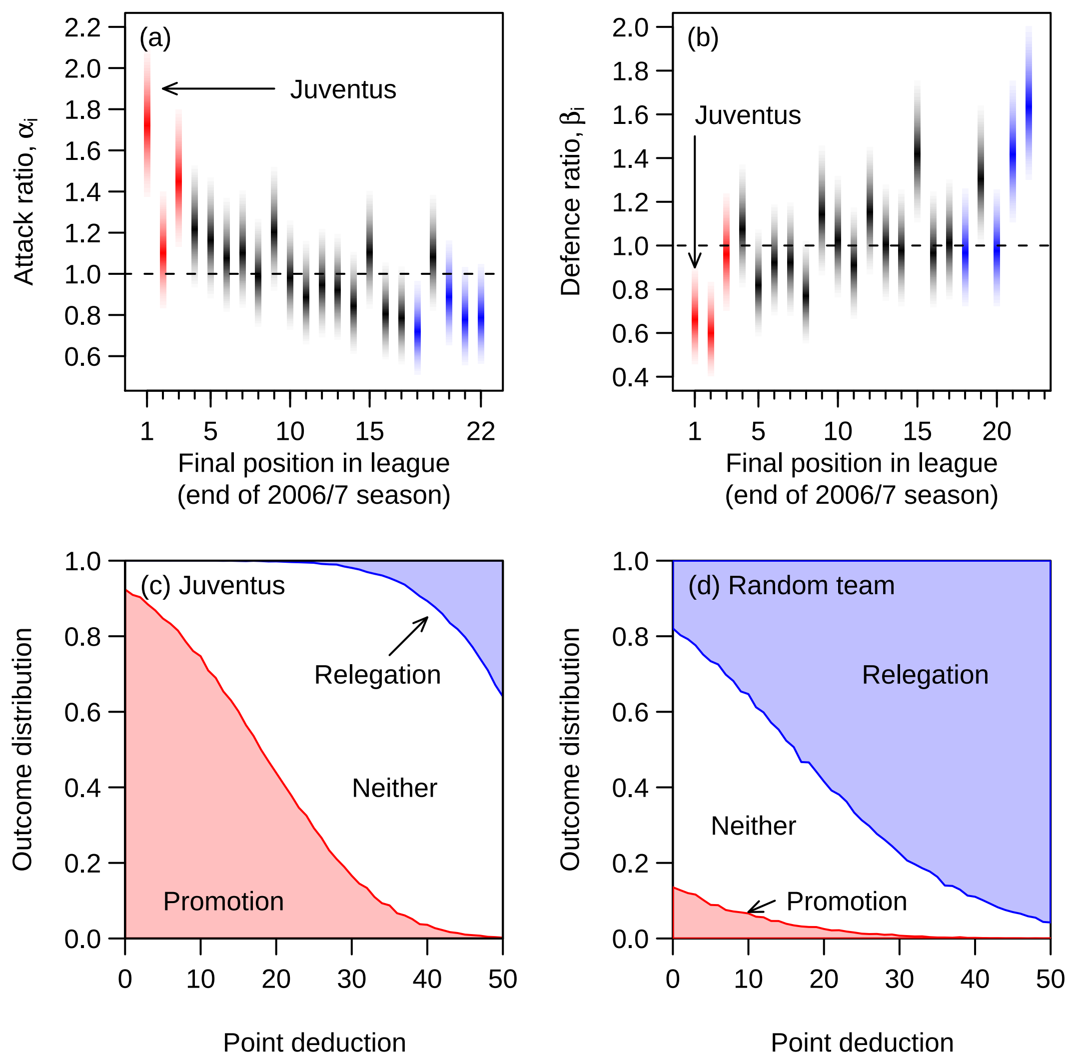

Figure 3: 2006–7 Italian Serie B season. (a) and (b) Posterior distribution of attack and defence ratios, with the same colourscheme as Figure 1. The relative strength of Juventus is clear. (c) The probability that Juventus would be promoted or relegated as a function of the imposed points deduction. The actual nine point penalty does little to diminish the very high probability of promotion. (d) The probability that a random team would be promoted or relegated from a typical league composed of randomly sampled opponents. In this case a nine point deduction leads to the probability of promotion being almost halved, and the probability of relegation being almost doubled.

Two (Figure 4a) with that for a random team in Serie B (Figure 3d), we see that although they are qualitatively very similar, the actual probabilities vary quantitatively over leagues. While part of this is due to the different baselines caused by number of teams in each league relative to the number of teams promoted/relegated, and there may also be some influence corresponding to Juventus' presence in certain realisations of the average Italian league, we note that any fair penalties would probably only make sense for particular leagues, and so that a single standard across Europe may be impossible to achieve.

From Figure 4c we see that a standardised penalty has a disproportionate effect on very weak and very strong teams, relative to the "average" team, since these are the teams that would be battling to avoid relegation or win promotion in the absence of a penalty anyway. If we associate plausible payoffs to promotion and relegation, we obtain *fine equivalents* of the point deduction (Figures 4b and d). The exact details of these depend on the value associated to promotion or relegation—the middle echelons of English football conform roughly to a revenue ratio of 2:1 in leagues one higher to one lower [Deloitte, 2009]—but some broad messages come out: weaker teams given a large points burden are invariably relegated, so that there is a diminishing marginal burden upon them from increasing penalties; stronger teams are affected most of all by sizeable fines since these effectively rule out obtaining the larger benefits of promotion for at least one year; for low points penalties, ϕ is approximately linear in points, so that if the league administration were able to evaluate the benefit and loss to promotion and relegation more precisely, they could devise a rule of thumb such as "one point deducted is worth X monetary units to the typical team" to assist in deciding the best penalty for a given infraction.

4 Discussion

Applying a well-accepted stochastic model of individual games of football [Dyte and Clarke, 2000, Karlis and Ntzoufras, 1999, Lee, 1997, Maher, 1982], we have presented a methodology to determine how the point deductions that are widely-applied as punishments in European football impact upon position in final league tables. This allowed us to estimate the increased probability of relegation and the decreased probability of promotion that is associated with a given penalty. By appropriately parameterising the model, we illustrated our method in the context of two case studies: relegation of Luton Town from English League Two in 2008–9, and promotion of Juventus from Italian Serie B in 2006–7. We found that the thirty point penalty imposed on Luton greatly increased that team's probability of relegation, from less than 0.05 to more

DOI: 10.2202/1559-0410.1214

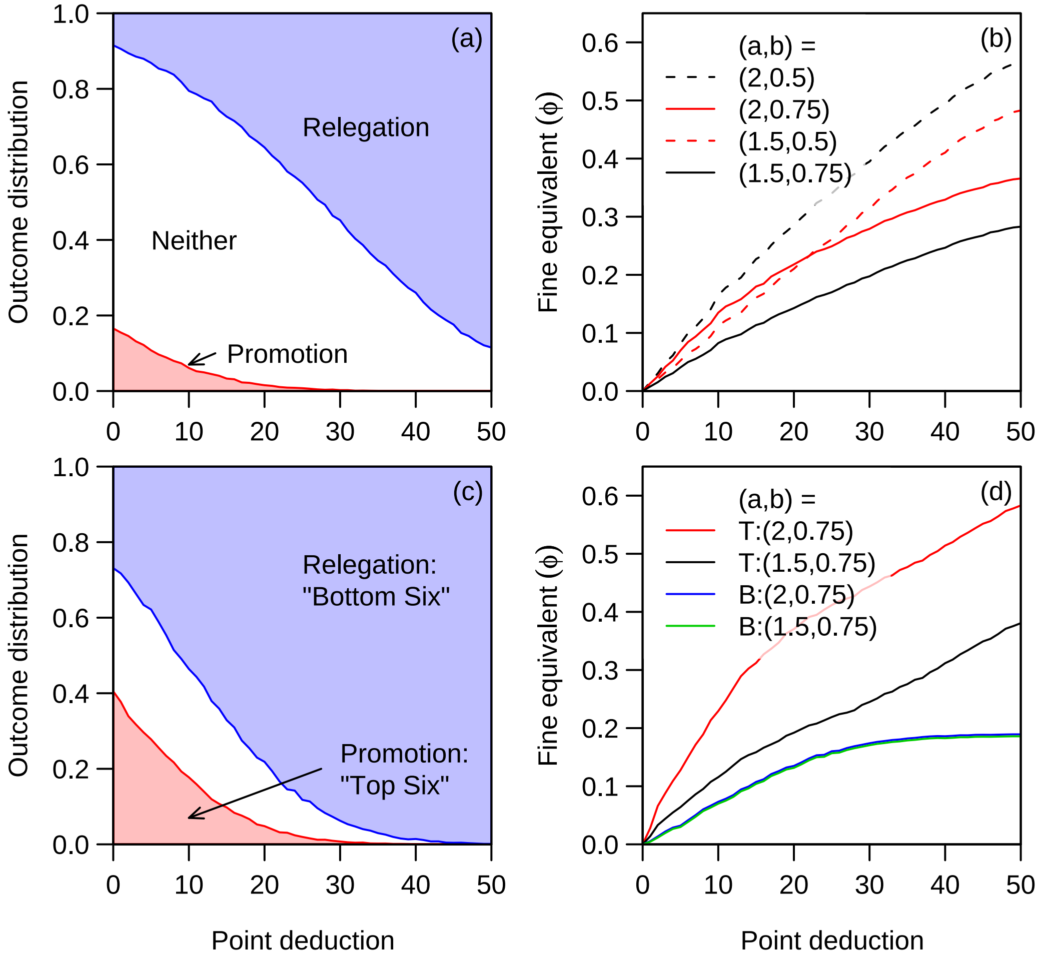

Figure 4: Fair penalties in English League Two. (a) Posterior predictive distribution of outcomes for a random team as a function of points deduction. Points deductions of 20 or more lead to almost no probability of promotion, while more than half of teams given a penalty of 30 points would be relegated. (b) Fine equivalent of points deductions for random teams under different scenarios on costs (the parameters a and b). Any penalty is most severe when $a = 2.0$ and $b = 0.5$ (black dotted line), corresponding to the payoff of being promoted being double that of remaining in situ, and the payoff of being relegated being half. (c) Impacts on probabilities of promotion and relegation for random teams from the top and bottom six respectively. (d) Associated payoff under different scenarios for top (T) and bottom (B) six teams. Beyond a 30 point penalty, bottom teams are effectively guaranteed relegation, reflected by the asymptote; top teams, however, become at-risk of relegation for very large penalties.

than 0.70. However the nine point penalty imposed on Juventus led to a only very small decrease in that team's probability of promotion, from 0.92 to 0.76. The difference in impact between the punishments imposed on this pair of teams is clearly very marked.

To address the question of designing standardised punishments, we investigated the effect of variously-sized point deductions on a "typical" team. This is much more difficult, as the probability of any unfavourable event depends not only on the strength of the team that is penalised, but also the quality of its opposition. We normalised for this by populating a typical realisation of a particular league by randomly sampling with replacement from the teams within it, and randomly choosing our focus team. In both leagues we considered, there was a smooth response to the number of points deducted, with the probability of relegation/promotion increasing/decreasing sigmoidally. However the numerical responses were different. Taking an arbitrary reference value of an increase of 0.25 in the probability of relegation as our threshold, we found that 18 and 13 points would have to be deducted from random teams in League Two and Serie B respectively to achieve this outcome. This indicates that any notion of a "fair" punishment would have to be matched to the league in question.

In our investigation of English League Two we also examined the influence of team strength on the size of fair deduction and the cost. Although we believe that two teams of different ability found guilty of the same misdemeanour should receive the same penalty, clearly that penalty will have a different effect on the two teams, and it is important to quantify this differential in setting a standard penalty. For this league, the effect of small penalties (less that 10 points, say) would be much more pronounced on teams vying for promotion or to avoid relegation (Figures 4a vs 4c). A larger penalty of 30 points is nearly enough to guarantee a weaker team be relegated or a stronger team not win promotion. When (arbitrary) costs are accounted for, though, the effect of a point deduction is more keenly felt on stronger teams, since revenues between a higher and a lower league are roughly in the ratio 2:1 [Deloitte, 2009], and so the loss from not being promoted is more than the loss from being relegated.

An obvious criticism of our approach is that the results depend upon the underlying model of individual games of football. A similar objection could be levelled at any model, and as our general framework is agnostic to the particular details of the model used for individual games, this component could be easily replaced. In our modelling we followed Lee [1997] and Maher [1982], taking the number of goals scored by each team as independent Poisson processes, and parameterising each team using goal scoring rates that were constant over the entire season. We feel that the combination of par-

DOI: 10.2202/1559-0410.1214

simony, simplicity and goodness-of-fit (Figure 2b) provide strong support for our results, although there are various ways that the model could become more sophisticated/complex if so desired. The independence assumption in the number of goals scored by two competing teams has been repeatedly challenged, mainly because of suggestions that it underestimates the probabilities of low scoring games [Dixon and Coles, 1997, Dixon and Robinson, 1998] and of draws [Karlis and Ntzoufras, 2003, McHale and Scarf, 2007]. A number of corrections have been proposed, including artificially increasing the probability of low scores using diagonal-inflated mixture models [Dixon and Coles, 1997, Dixon and Robinson, 1998], and using bivariate Poisson distributions which model the correlation between scores in a single match more explicitly [Karlis and Ntzoufras, 2003, McHale and Scarf, 2007]. It has also been noted that teams that run up large numbers of goals against low scoring opponents are possibly overweighted in our framework, whereas those which settle into a defensive shell when they take the lead are underweighted [Lee, 1997]. The assumption of constant goal scoring rates over the entire season has also been criticised [Dixon and Robinson, 1998], with the obvious objections that the relative strength of any particular team will differ from game to game due to fluctuations associated with possible injuries, transfers, changes in management and team morale, leading to over-dispersal relative to a Poisson model. In fact in addition to allowing the parameters corresponding to a particular team to change over the season, these authors used a more complex model in which the rate at which goals are scored changed according to the current score (although their focus was mainly in improving estimates of the probability of particular scores for "in-play" betting markets rather than modelling an entire season). In the case of a large pre-season point deduction, it could certainly be argued that team morale may be reduced after the players realised a sequence of good results had had no impact on a poor position in the league table.

Compared to previous Poisson models, we used a different parameterisation for the strength of each team. The original models using this approach [Lee, 1997, Maher, 1982], together with the majority of the various extensions [Dixon and Coles, 1997, Dixon and Robinson, 1998, Karlis and Ntzoufras, 1999, 2003], parameterise the underlying Poisson model using a pair of parameters per team (one related to the number of home goals, one related to the number of away goals), together with a home advantage that remains constant across all teams (and which is the mean number of additional goals expected by the home team). Although relative increases and decreases in the home and away parameters have the obvious interpretation in terms of relative team performance, the parameters themselves often do not actually mean anything as their numerical values depend on the normalisation used (for example, the decision to set the home parameter of one arbitrary team to one in Lee [1997]). By replacing the home advantage by two separate parameters representing the mean number of expected goals playing at home and away, and using the appropriate summation constraints on the home and away parameters ($\log \alpha$ and $\log \beta$), we ensured that our numerical values were actually ratios, relating the performance of each team's attack and defence to that of the average team in that league. We feel this is more easily interpreted.

In attempting to quantify the impact of point penalties in financial terms, we arbitrarily assigned costs of promotion or relegation, and investigated the response relative to these choices, and began to investigate the impact of these choices on the form of the response. The alternative would be to determine the actual costs of relegation or promotion. This is clearly very difficult, and indeed has itself been the subject of research by economists [Noll, 2002]. We did consider attempting to define a baseline by comparing the turnovers of teams in the leagues directly above and below those in question, using information that is available in certain cases [Deloitte, 2009]. Unfortunately financial information for teams in the leagues directly below English League Two and Italian Serie B proved difficult to obtain. However even if we have been able to find the relevant figures, the approach would have been flawed, since newly promoted or relegated teams are not typical of their new league; in particular, they may have different wage bills or attendances and possibly parachute payments (which certain leagues offer to aid the transition to a lower league). More fundamental is that the future economic costs due to relegation (or missed promotion) must be accrued over multiple seasons, in which further promotions or relegations may occur. This variation in sojurn time makes it very difficult to identify costs unambiguously. However assigning payoffs to these outcomes, even if they do not necessarily correspond exactly to financial payoffs, allows the impact of differently-sized point reductions to be compared in terms of one well-defined metric according to the desirability of remaining in the current league.

Our work offers a quantitative method to investigate the potential impact of point penalties on team sports organised in leagues. Our basic approach could easily be extended to other sports with similar penalties, such as rugby football or basketball. We contend that this approach allows an unambiguous quantification of the potential consequences of the point penalties that have been widely, but not necessarily scientifically, implemented by the governing bodies of European football leagues.

DOI: 10.2202/1559-0410.1214

References

Deloitte. *Annual Review of Football Finance*. Manchester: Sports Business Group, 2009.

M. J. Dixon and S. G. Coles. Modelling association football scores and inefficiencies in the football betting market. *Applied Statistics*, 46:265–280, 1997.

M. J. Dixon and M. E. Robinson. A birth process model for association football matches. *The Statistician*, 47:523–536, 1998.

S. Dobson and J. Goddard. *The Economics of Football*. Cambridge University Press, 2001.

S. Dobson and J. Goddard. Revenue divergence and competitive balance in a divisional sports league. *Scottish Journal of Political Economy*, 51:359–376, 2004.

D. Dyte and S. R. Clarke. A ratings based Poisson model for World Cup soccer simulation. *Journal of the Operational Research Society*, 51:993–998, 2000.

P. Everson. Good offense vs. poor defense: the 2007 Women's World Cup. *Chance*, 21:46–54, 2008. doi: 10.1007/s144-008-0029-z.

P. Everson and P. S. Goldsmith-Pinkham. Composite Poisson models for goal scoring. *Journal of Quantitative Analysis in Sports*, 4, 2008. doi: 10.2202/1559-0410.1107.

W. R. Gilks, S. Richardson, and D. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Taylor and Francis, 1995.

D. Karlis and I. Ntzoufras. Statistical modelling for soccer games: the Greek league. In E.A. Lipitakis, editor, *Proceedings of the Fourth Hellenic-European Conference on Computer Mathematics and its Applications*, pages 541–548, 1999.

D. Karlis and I. Ntzoufras. Analysis of sports data using bivariate Poisson models. *The Statistician*, 52:381–393, 2003.

A. J. Lee. Modelling scores in the premier league: is Manchester United *really* the best? *Chance*, 10:15–19, 1997.

P. M. Lee. *Bayesian Statistics: an introduction*. Wiley, 2009.

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

P. McCullagh and J. Nelder. *Generalized Linear Models*. Chapman and Hall, 1989.

I. G. McHale and P. A. Scarf. Modelling soccer matches using bivariate discrete distributions. *Statistica Neerlandica*, 61:432–445, 2007.

R. G. Noll. The economics of promotion and relegation in sports leagues: the case of English football. *Journal of Sports Economics*, 3:169–203, 2002.

S. Szymanski and R. Smith. The English football industry: profit, performance and industrial structure. *International Review of Applied Economics*, 11: 135–153, 1997.

DOI: 10.2202/1559-0410.1214