On outcome uncertainty and scoring rates in sport: the case of international rugby union

Phil Scarf, *Centre for Sports Business, Salford Business School, University of Salford, Salford, M5 4WT, UK. p.a.scarf@salford.ac.uk (corresponding author)*

Rishikesh Parma, *Department of Applied Mathematics, Indian Institute of Technology (ISM), Dhanbad 826004, India. rishikesh786@am.ism.ac.in*

Ian McHale, *Centre for Sports Business, University of Liverpool Management School, L69 7ZH, ian.mchale@liverpool.ac.uk*

Abstract: In this paper we investigate the relationship between outcome uncertainty and scoring rates in the framework of a "Poisson match". We argue that increasing scoring rates in the hope of increasing entertainment may have a detrimental impact on the popularity of sport. The basis of our argument is that higher scoring-rates decrease outcome uncertainty. We use international rugby to demonstrate our findings and show that scoring rates have indeed increased significantly over the previous half-century in this sport. Therefore, administrators should recognise our general point and we suggest that rugby union administrators in particular ought to consider the introduction of new laws to reduce scoring-rates. Scenarios in which the scoring-rate is radically reduced are illustrated through a simulation of the Rugby World Cup tournament.

1. Introduction

Sport holds a unique standing in society. Entire nations take joy from sporting success at Olympics Games and World Cup tournaments, whilst discussion of recent results and anticipated matches is a universal pastime. The fundamental reasons for the popularity of sport are unknown, but the outcome uncertainty hypothesis (Rottenburg, 1956) posits that outcome uncertainty is an important factor. It seems intuitive that interest in a competition will dwindle if the result is known before the competition takes place, and so, higher outcome uncertainty is deemed to be an attractive quality of a sport. Outcome uncertainty is closely related to competitive balance – a concept by which outcome uncertainty can be measured. The relationship between the popularity of a sport and outcome uncertainty has been well-studied (e.g. Forrest and Simmons, 2002; Alavy et al., 2010; Buraimo and Simmons, 2015; Hogan et al., 2013, 2017; Kuchar and Martin, 2016), and has been shown to even transcend national boundaries (Schreyer et al., 2017). In these cases, higher competitive balance is found to be associated with elevated interest, although Borland and Macdonald (2003) in a review paper found that the empirical results on this relationship are mixed.

The consumption of sport satisfies a desire for excitement (Mutz and Wahnschaffe, 2016), and an exciting sport combines elements of outcome uncertainty with other valuable goods that we bundle as "athleticism", including skill, speed, flair, power, determination, and teamwork, or indeed the opposite through errors. While Ely et al. (2015) propose suspense (variance of outcome) and surprise (the totality of within-match fluctuations in expected outcome) as distinct aspects of outcome uncertainty, and their relative importance to consumers of sport has been measured (Bizzozero et al., 2016; Mutz and Wahnschaffe, 2016), the preference for athleticism over outcome uncertainty is not known. Scoring events themselves conjoin athleticism with elements of outcome uncertainty (a score by a trailing team will increase suspense; a score by a weaker team will increase

surprise), and so disaggregating the effects of underlying factors is challenging. Nonetheless, it is our thesis that more scoring reduces (within-match) outcome uncertainty, and surprise and suspense by implication.

In particular, in this paper we show that as scoring rate increases the numbers of scores of opposing teams tend to diverge (Stefani, 2009) so that the outcome between two teams of differing strength will tend to become less uncertain. We establish this result in the context of a Poisson match.

The case of rugby union offers an opportunity to discuss the implications of this result. This is because in rugby there is more scoring now than in the past; we demonstrate this later in the paper. In rugby, there are also many ways of scoring (the try, the conversion, the penalty, and the drop-goal), with each type of score carrying a different points value (respectively 5, 2, 3, and 3). In part, the increase in scoring is a consequence of an increase in kicking skill and the quality of the ball and the playing surface, although this increase has been moderated by rule changes that have increased the benefit of not attempting a penalty-kick. Indeed, as penalty-kick success has increased, administrators have introduced rule changes (Wright, 2014) that allow more attacking rugby and more tries, maintaining the balance between open rugby (more tries, more excitement) and defensive rugby (less tries, more penalties, less exciting rugby), but neglecting the excitement derived from close outcomes. Thus scoring rates have increased in all aspects of scoring (except possibly drop goals which remain relatively rare although occasionally dramatically influential). We think policymakers have not considered the consequences of increasing scoring-rate sufficiently, and our notion is that in rugby the "goals are too wide". Furthermore, bonus points, for narrow defeats and high numbers of tries, may be making matters worse, by favouring the strong and incentivising yet more scoring (Winchester, 2008; Lenten and Winchester, 2015). Given the outcome uncertainty hypothesis, the increasing scoring rates observed over the last half a century may be reducing the popularity of rugby as a sporting contest.

We use a model of scoring in rugby to simulate matches in the Rugby World Cup tournament to illustrate the impact of alternative scoring rules and show how reverting to the original scoring rules upon which rugby was founded would increase surprise, suspense and outcome uncertainty, albeit using simpler definitions of surprise and suspense than those defined in Ely et al. (2015). Our model of a rugby match is based on the model Maher (1982) first presented for soccer, but extended in a similar way to Pledger and Morton (2011) and Baker and McHale (2013). The World Cup tournament simulation itself mirrors the approach taken in Scarf et al. (2009) and we use similar metrics for measuring tournament performance, e.g. the probability that the best team wins, and the probability that the *i*-th ranked team reaches round *j* (Scarf and Yusof, 2011). We assume team strengths reflect those of the tournament played in 2015.

This is the first paper to consider this relationship between scoring rates and outcome uncertainty and its results have wider implications than for just rugby. That this relationship has been overlooked may be because research on outcome uncertainty and competitive balance has focused on soccer (Utt and Fort, 2002; Owen et al., 2007; Manasis et al., 2013), where scoring rates are not increasing (Figure 1), and North American sports, where the draft plays a pivotal role in moderating competitive imbalance (Szymanski, 2003).

The structure of paper is a follows. First, we describe scoring-rate evolution in rugby since 1960. Then in Section 3, we prove that in a Poisson match, scores diverge as scoring rates increase. In Section 4 we further analyse scores in preparation for model fitting, we fit a model akin to a Poisson match, and carry out tournament simulations using the model. We conclude with a discussion of the implications of our results for sports administrators.
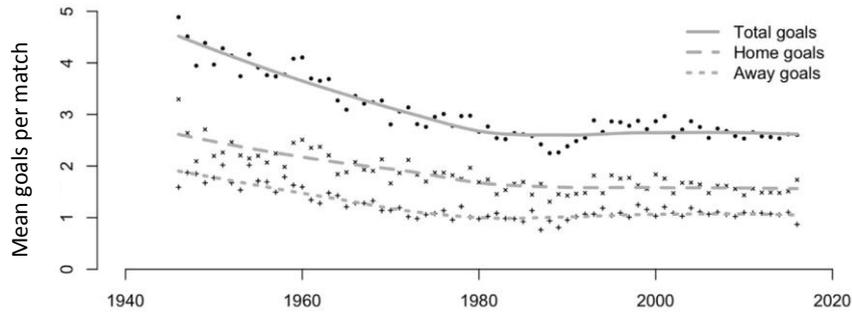
Figure 1. Mean number of goals scored in international soccer per match by year, reproduced from Baker and McHale (2018).

2. Introductory data analysis

For the analysis in this paper, data for all matches in the period from 1st January 1960 to 31st December 2016 between the teams that have been consistently ranked in the top 8 over this period (Australia, England, France, Ireland, New Zealand, Scotland, South Africa, Wales) were collected from http://stats.espnscrum.com/statsguru/rugby/stats, giving 1366 matches, at a mean rate of approximately 4 matches per team per year during the amateur era (to 1996) and 8 matches per team per year since then. We consider only these eight, elite teams so that the entry of newer weaker sides to the international scene does not distort measures of scoring rate. In the 60s and 70s, matches between these elite teams were 45% and 27% respectively of recorded matches played at international level. Aggregating the data across these teams, Figure 2 shows the development of scoring rates per match for tries, penalties and conversions by year. We can see that the numbers of scores has more than doubled. Drop-goal scoring rates are not shown in Figure 2, and in comparison to the other types of scoring rates, are very low (0.2 per team per match) and not increasing, although in an exceptional match in 1999 South Africa scored five against England.

These figures then capture the four means of scoring in rugby: the try, whereby the ball is grounded by the attacking team on or behind the goal line of the defending team; the conversion, whereby the ball is kicked from a tee to between the goal posts following a try; a penalty, whereby the ball is kicked from a tee to between the goal posts following an infringement of the rules by the defending team; and a drop-goal, whereby the ball is kicked between the goal posts in open play. Note, on the definition of terms, the reader should distinguish a penalty (the score event worth 3 points) from a penalty-kick (the attempt to score a penalty), and while one might use the term conversion to define a successful penalty-kick, we reserve the term conversion exclusively for the scoring event that follows the try. A fifth method of scoring has existed, the goal from a mark. This was discontinued in 1971. There was only one occurrence of this in the period 1960-1970 among these eight teams (NZ v Eng, 1st June, 1963) so we ignore this in the analysis.

Each score-type is assigned a points-value, and the team with the highest points wins the match. In 2017, the points-value assignments are 5 for a try, 2 for a conversion, and 3 each for a penalty and drop goal. In 1960, the respective values were 3,2,3,3. Until 1886, the respective values were 0,1,1,1, and the outcome of a match drawn on points, which occurred more frequently then than now, was determined by the number of unconverted tries. To accommodate the evolution of the points-value of a try, Figure 2 shows the number of scores per match versus time rather than the number of points per match versus time. Of course, these scoring rates are a reflection of the increasing relative

3

reward for scoring a try, so it is noteworthy that the rate of scoring penalties (which has decreased in value compared to a try) has increased at a faster rate than that of scoring tries.
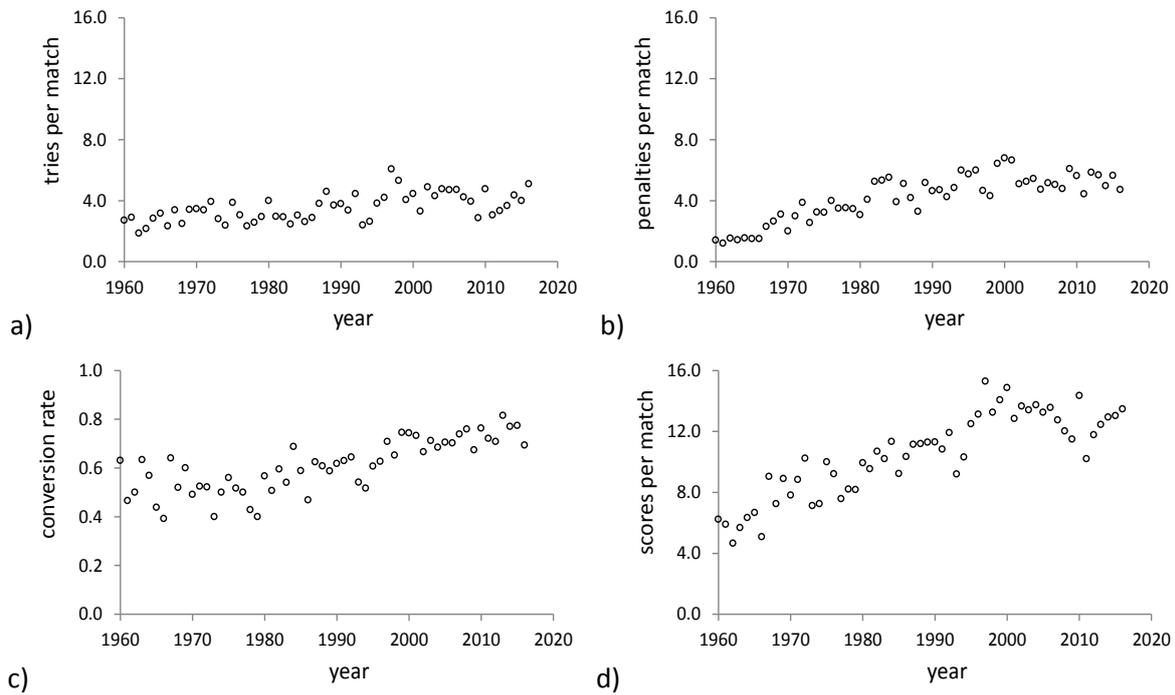
Figure 2. a) Mean number of tries scored per match by year, including penalty-tries, b) mean number of penalties scored per match by year, c) conversion success rate (proportion) by year, d) Mean total number of scores (tries plus penalties plus conversions plus drop-goals) per match by year , in all matches between the "top-8" since 1960.

We sought also data on domestic matches. Points scored are available from various sources for some interesting tournaments, but match results broken down by numbers of scores by type are not available. So it would be difficult to repeat our analysis using domestic matches. Also, the designs of domestic tournaments and the participating teams have varied over the decades. Nonetheless, a similar pattern of increasing scoring-rate can be observed (Figure 3).
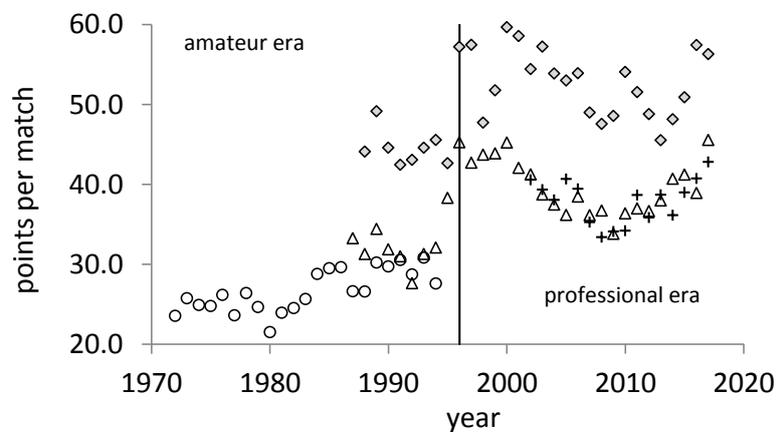
Figure 3. Mean total points per match in domestic tournaments by year and by tournament:
○ RFU Club Tournament (England) (1217 matches); Δ English Premiership (3914 matches);
+ Top 14 French league (2730 matches); ◊ Currie Cup (South Africa, 645 matches).

3. Outcome uncertainty and the Poisson match

In this section we prove some results that characterise outcome uncertainty in a Poisson match. We then discuss the implications of these results for competitive sport.

Proposition 1. Let $X_1 \sim \text{Po}(\lambda_1)$ and $X_2 \sim \text{Po}(\lambda_2)$ independent, with $\lambda_1 = \lambda$ and $\lambda_2 = \varepsilon\lambda$ for some $\varepsilon > 1$. Then $\text{Pr}(X_1 < X_2) \to 1$ as $\lambda \to \infty$.

Proof: Let Y be a random variable. Then for any $t > 0$, $\text{Pr}(Y \geq 0) = \text{Pr}(e^{tY} \geq 1) \leq E(e^{tY})$ by Markov's inequality (since e^{tY} is a non-negative random variable). Now setting $Y = X_1 - X_2$ we have that

$$\text{Pr}(X_1 - X_2 \geq 0) \leq E(e^{t(X_1 - X_2)}) = E(e^{tX_1})E(e^{-tX_2}) = \exp\{\lambda_1(e^t - 1) + \lambda_2(e^{-t} - 1)\}, \qquad (1)$$

since the moment generating function of X_i is $E(e^{tX}) = \exp\{\lambda_i(e^t - 1)\}$, for $i = 1, 2$. Now setting $t = \log\sqrt{\varepsilon}$ in (1), and noting that $t > 0$ because $\varepsilon > 1$, we obtain

$$\text{Pr}(X_1 - X_2 \geq 0) \leq \exp\{\lambda(\sqrt{\varepsilon} - 1) + \varepsilon\lambda(\frac{1}{\sqrt{\varepsilon}} - 1)\}$$

$$= \exp\{-\lambda(1 - 2\sqrt{\varepsilon} + \varepsilon)\}$$

$$= \exp\{-\lambda(1 - \sqrt{\varepsilon})^2\} \to 0$$

as $\lambda \to \infty$. This proves the result. $\qquad\qquad \square$

Therefore in a "Poisson match", a match between two teams whose scores are independent Poisson random variables, then no matter how close are the strengths (scoring rates) of the two teams, the probability that the stronger team wins approaches the value one as the scoring rate increases. Expressing this in terms of outcome uncertainty, as the scoring rate increases the outcome uncertainty decreases. In the limit for a very large scoring rate, the stronger team will always win and the outcome of the match is known with certainty.

Note, the literature on outcome uncertainty (see e.g. Leeds and von Allmen, 2014) distinguishes three types of outcome uncertainty: within-match or intra-match; intra-tournament; and inter-tournament. Our results here consider the first of these, although we will consider the second by implication later in the paper. Also, by modelling strengths relatively ($\lambda_2 = \varepsilon\lambda_1$), the model considers the relative quality of play, which is the concern of the outcome uncertainty hypothesis, rather than the absolute quality of play, which is not.

Proposition 2. Let $X_1, X_2 \sim \text{Po}(\lambda)$ independent, then (a) $\text{Pr}(X_1 = X_2) \to 0$ as $\lambda \to \infty$, and (b) $\text{Pr}(X_1 < X_2) = \text{Pr}(X_1 > X_2) \to \frac{1}{2}$ as $\lambda \to \infty$.

Proof: Now

$$\text{Pr}(X_1 = X_2) = \sum_{x=0}^{\infty} e^{-2\lambda}\lambda^{2x} / (x!)^2 = e^{-2\lambda}I_0(2\lambda)$$

where $I_0(.)$ is a modified Bessel function of the first kind with index zero. Using an integral representation of $I_0(.)$, the result (a) follows since

$$e^{-2\lambda}I_0(2\lambda) = e^{-2\lambda}\frac{1}{\pi}\int_0^{\pi} \exp(2\lambda\cos\theta)\,d\theta = \frac{1}{\pi}\int_0^{\pi} \exp\{-2\lambda(1 - \cos\theta)\}\,d\theta \to 0$$

as $\lambda \to \infty$ because the integrand $\exp\{-2\lambda(1 - \cos\theta)\} \to g(\theta)$ as $\lambda \to \infty$ where

$$g(\theta) = \begin{cases} 1, & \theta = 0, \\ 0, & 0 < \theta < \pi, \\ 1, & \theta = \pi. \end{cases}$$

Result b) follows from result (a) by symmetry. □

These propositions establish the asymptotic behaviour of $\Pr(X_1 > X_2)$ and $\Pr(X_1 = X_2)$. For exact calculation of $\Pr(X_1 > X_2)$ and $\Pr(X_1 = X_2)$ for $X_1 \sim \text{Po}(\lambda)$ and $X_2 \sim \text{Po}(\varepsilon\lambda)$ we use

$$\Pr(X_1 < X_2) = \sum_{y=1}^{\infty} \sum_{x=0}^{y-1} \{e^{-\lambda}\lambda^x / x!\}\{e^{-\varepsilon\lambda}(\varepsilon\lambda)^y / y!\},$$

and

$$\Pr(X_1 = X_2) = \sum_{x=0}^{\infty}\{e^{-\lambda}\lambda^x / x!\}\{e^{-\varepsilon\lambda}(\varepsilon\lambda)^x / x!\} = \sum_{x=0}^{\infty} e^{-\lambda(1+\varepsilon)}\lambda^{2x}\varepsilon^x / (x!)^2 .$$

We use a finite approximation to these exact formulae to illustrate the asymptotic results graphically in Figure 4.
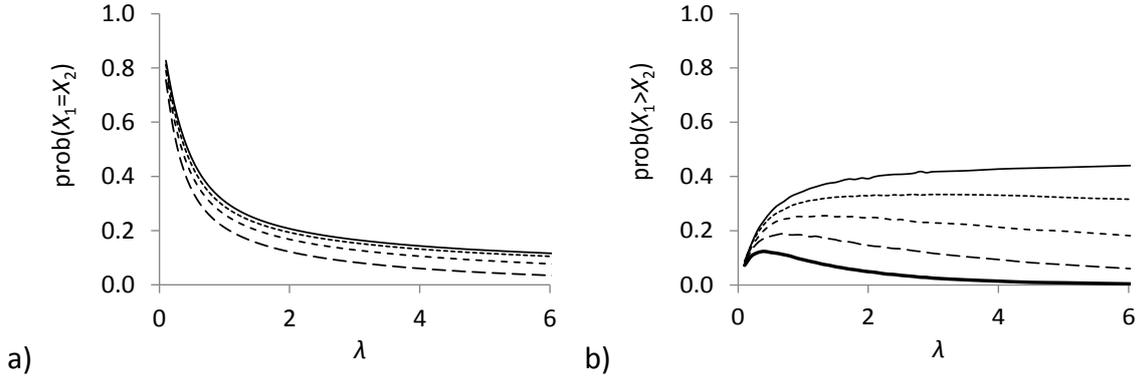

Figure 4. For $X_1 \sim \text{Po}(\lambda)$ and $X_2 \sim \text{Po}(\varepsilon\lambda)$ independent: a) $\Pr(X_1 = X_2)$, b) $\Pr(X_1 > X_2)$ as a function of λ for various ε (solid line $\varepsilon = 1$, short dash $\varepsilon = 1.2$, medium dash $\varepsilon = 1.5$, long dash $\varepsilon = 2$, bold solid line $\varepsilon = 3$, (b) only).

Corresponding results can be proved for similar models with normal, log-normal, and binomial margins. Thus, we can obtain a further proposition.

Proposition 3. If $(X_1, X_2) \sim \text{N}(\lambda, \lambda\varepsilon, \sigma_1^2, \sigma_2^2, \rho)$ (dependent bivariate normal distribution) $\varepsilon > 1$, then $\Pr(X_1 < X_2) \rightarrow 1$ as $\lambda \rightarrow \infty$.

Proof: The result follows because $X_1 - X_2 \sim \text{N}(\lambda(1-\varepsilon), \sigma^2)$, where $\sigma^2 = \sigma_1^2 - \rho\sigma_1\sigma_2 + \sigma_2^2 > 0$, and so $\Pr(X_1 < X_2) = \Pr(X_1 - X_2 < 0) = \Phi(\lambda(\varepsilon-1)/\sigma)$, where $\Phi(.)$ is the cummulative distribution function of the standard normal distribution, and $\Phi(\lambda(\varepsilon-1)/\sigma) \rightarrow 1$ as $\lambda \rightarrow \infty$ because $\varepsilon > 1$. □

We make some remarks about these propositions.

As Propostion 3 holds in the dependent case, it also holds in the independent case ($\rho = 0$).

A result for a bivariate binomial can be proved in a similar way to Proposition 1.

The result (divergence of scores) in Propositions 1-3 is not universally true. For example, when the marginal distributions are both zero-inflated, $\Pr(X_1 = X_2 = 0) > 0$, always, due to the zero-inflation of the bivariate distribution. Crucially, a general result depends on two assumptions in the model: firstly, that strengths differ relatively ($\lambda_2 = \varepsilon\lambda_1$) rather than absolutely ($\lambda_2 = \delta + \lambda_1$); and secondly, that a Bernoulli process underlies the score distributions, which is the basis of the Poisson match. Thus, regarding the first, we can see that if $(X_1, X_2) \sim \text{N}(\lambda, \lambda + \delta, \sigma_1^2, \sigma_2^2, \rho)$ then $X_1 - X_2 \sim \text{N}(-\delta, \sigma^2)$ and so $\Pr(X_1 < X_2)$ does not depend on λ.

When continuous distributions are used to approximate discrete distributions (of scores in sports), probabilities may be calculated using a continuity correction. A result similar to Proposition 3 can be proved for this case (see Appendix 1).

The corresponding result for the bivariate log-normal distibution (which has support on the upper right quandrant) can be proved by using the log-transformation in Proposition 3. The discretized bivariate log-normal distribution is then a model that is suitable for scores that must be positive. As an aside, to accommodate heavier tails, one could discretize elliptical models (e.g. Bingham et al., 2003) that are used in financial modelling.

Moving then from the model to the reality, the practical consequences are that if scores are approximately Poisson (e.g. soccer, see McHale and Scarf, 2007) then increasing the scoring rate decreases the outcome uncertainty. Thus, picking up a soccer example here, while the size of the goal is sacrosanct, penalising defenders for impeding attackers in the penalty area more stringently (than has been the case in the past) is not and would likely increase the scoring rate.

Notice conversely that, as $\lambda \to 0$, a contest becomes perfectly (competitively) balanced but the outcome becomes certain, since in the limit no team scores and the match outcome is 0-0 with probability 1. Therefore competitive balance and outcome uncertainty are not synonymous.

We might anticipate that the effect of interest is moderated as the dependence between scores increases, so that the scoring-rate effect is exaggerated if scores are negatively dependent and somewhat reduced if scores are positively dependent. We investigate the dependence of scores in international rugby in the next section. We also consider whether the Poisson distribution is a good model for international rugby scores.

4. Implications for international rugby union

So far, we have i) proved that for the Poisson match an increasing scoring-rate decreases outcome uncertainty, ii) shown that the scoring-rate in matches between elite international rugby teams is increasing (more than doubled since 1960), and iii) stated the outcome uncertainty hypothesis that outcome uncertainty is an important contributor to popularity of a sport. Now we do three things. Firstly we consider whether a Poisson match is a reasonable approximation to an international rugby match. If it is then it may be that increasing scoring rate in rugby is detrimental to the popularity of the sport. Secondly we fit a model that is akin to the Poisson match to all results of matches played in the previous ten years between the twenty teams who contested the 2015 Rugby World Cup tournament. Thirdly, we use this model to simulate the tournament and we illustrate through simulation how the outcome of the tournament may look when scoring rates are modified. In this way, we can explore the effect on tournament outcome of changes to the rules of the game. In particular, we consider a revised scoring rule in which effectively a team scores a point if and only if it converts a try.

4.1. *Poisson match approximation*

We performed tests of goodness-of-fit (GOF) of the Poisson distribution, decade by decade (because of the increase in the mean scoring rate over time), and by first-named and second-named team (because of the potential for a higher scoring-rate for the first-named team and bearing in mind that where applicable the first-named is the home team). Summary results are shown in Table 1. Overall, there is evidence of zero inflation and some tendency for larger numbers of scores than is described by the Poisson distribution (Figure 5), but not in all decades. Contributions to the chi-squared GOF statistics, which are not reported for brevity, reflect this. The higher frequency of zeros and high

7

numbers of scores than expected implies that the Poisson match is less dispersed than the reality and so, in the Poisson-match modelling environment, matches will be somewhat more competitive in reality. Also, given that conversions can only follow tries and the different means of scoring do not have the same points-value, there are structural reasons why the reality cannot be a Poisson match.

Furthermore, the case for accommodating zero-inflation is not strong because 0-0 would arise in 0.05% of matches (64 of 2732 scores (2.3%) are zero), and so the effect of zero-inflation on outcome uncertainty will be negligible. In fact, five of the 1366 matches in our dataset ended 0-0 (0.37%), which is five times more than expected under the assumption of independence. However, all these 0-0 matches occurred before 1965.

Table 1. Chi-squared goodness-of-fit statistics for tests of the Poisson distribution, by decade and for first-named and second-named teams. All tests on 7 degrees of freedom. $\chi^2_{7,0.99} = 18.5$.

	60s	70s	80s	90s	00s	10s
First-named	36.4 **	5.2	6.6	28.0 **	42.0 **	9.5
Second-named	45.7 **	13.2	11.9	49.7 **	45.7 **	19.2 **

** significant at 1% level
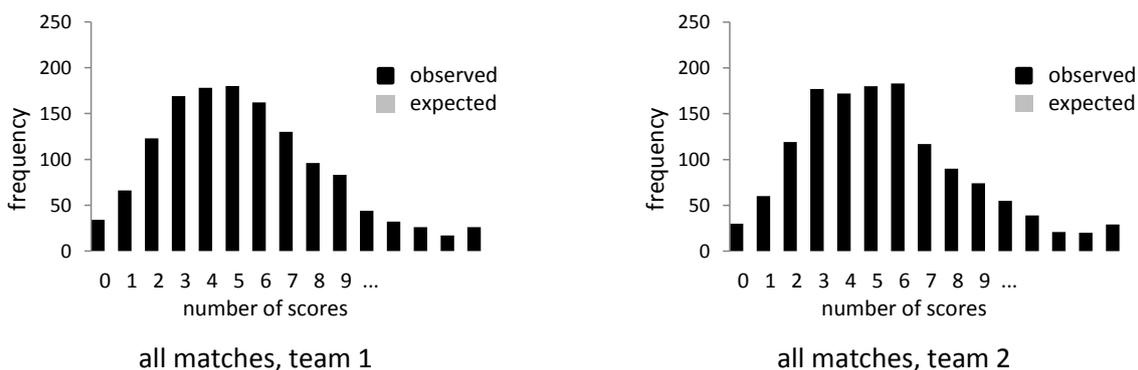
Figure 5. Observed frequencies and expected Poisson frequencies of numbers of scores (tries plus conversions plus penalties plus drop-goals) for team 1 (first-named in match) and team 2 (second-named) for all matches between the "top-8" since 1960.

Indeed, it is important now to consider dependence between scores. This is because the Poisson match will tend to exaggerate the competitive balance, scoring rate effect if real scores are positively dependent. Then as an alternative model, we might use the bivariate Poisson distribution (Karlis and Nzoutfras, 2003). For all matches, the correlation between team 1 and team 2 scores is slightly positive, but not statistically significant (0.047, $p > 0.1$). Looking year-wise (Figure 6), the correlations are mostly negative. This reversal is most likely due to the increase in scoring rates over time. Regardless of this, independence is a useful simplifying assumption even if it is not completely, empirically justifiable.

For the model development in Section 4, it will be simplest also to assume number of tries and number of penalties by a team within-match are independent. We investigate this dependence here. Dependence, if it exists, may be at team level, so we consider correlations of number of tries and number of penalties by team and decade, shown in Table 2. All the significant correlations there are

negative. There is a good reason for the negative correlation: when the try-count becomes high for one team in a one-sided match, the dominant team will tend to forego penalty-kicks. Nonetheless, for matches in which try-scoring is lower, correlation may be less evident. Actually, this appears to be true but the correlation does not decrease appreciably, so that for example if one looks at just scores where the number of tries < 6, the correlation is -0.080. Team-decade-wise, the results in Table 2 make intuitive sense. For example, in the case of New Zealand as their dominance has grown over time they have tended to play an open game (scoring tries) instead of a tight game (kicking penalties). And England in the 90s and 00s were noted for playing a tight, cautious game.
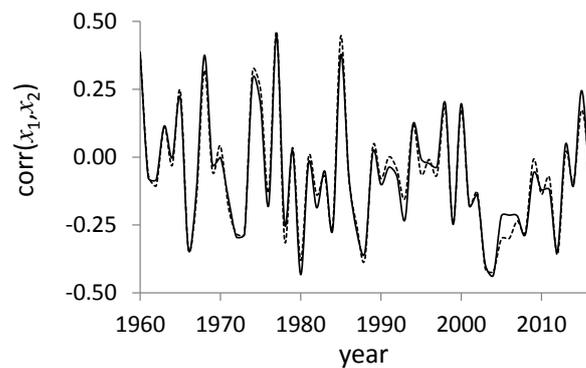

Figure 6. Correlation of team 1 score, x_1, and team 2 score, x_2, for all matches played in year i for all years since 1960 to date; dashed line for numbers of scores; solid line for points.

Table 2. Pearson correlations between number of tries and number of penalties by team and by decade, with significance levels for 2-sided test.

	60s	70s	80s	90s	00s	10s
Australia	-0.091	-0.189	-0.004	-0.252 *	-0.324 **	-0.145
England	0.146	-0.067	-0.327 *	-0.200	-0.225 *	-0.140
France	-0.308 *	-0.207	-0.038	-0.084	-0.086	-0.141
Ireland	-0.214	-0.261	-0.328 *	-0.443 **	-0.080	-0.266
New Zealand	-0.024	-0.269	-0.316 *	-0.286 *	-0.432 **	-0.520 **
Scotland	-0.075	-0.098	-0.127	-0.179	-0.340 **	-0.243
South Africa	0.044	0.119	0.038	-0.355 **	-0.223 *	-0.371 **
Wales	0.189	-0.232	-0.239	0.089	-0.093	-0.383 **

* significantly different from zero at 5% level, ** at 1% level

Overall then the Poisson match provides a reasonable approximation to the scores in an international rugby match, not least because scores show low correlation. Next, we estimate team strengths in a Poisson-match model.

4.2. Estimation of team strengths

We now identify the 20 teams who played in the Rugby World Cup tournament (RWC) in 2015. These teams are labelled 1,...,20. We collected data on all matches between these teams over an 11 year period (942 matches from 4th February 2006 to 26th February 2017). We fit a model like the Maher (1982) model to these scores, but without the sophistication of time varying strengths (Owen, 2011)

or zero-inflation (McHale and Scarf, 2011) or inflation for particular score combinations (Dixon and Coles, 1997) or alternatives to the Poisson distribution (Boshnakov et al., 2017) that have been used to model soccer. However, note, such development in the context of matches with complicated scoring systems would make an interesting study.

Consider match i played between team i_1 and i_2, for $i = 1,...,M$. Thus the pair (i_1, i_2) indexes the teams playing in match i. Let $(Y_{i_1}^j, Y_{i_2}^j)$ be the number of scores of type j by teams i_1 and i_2 respectively in match i, where $j = 1$ for tries, $j = 2$ for penalties and $j = 3$ for drop-goals. Let (Z_{i_1}, Z_{i_2}) be the number of conversions by teams i_1 and i_2 in match i. Assume: $\mathbf{Y}_i = (Y_{i_1}^1, Y_{i_1}^2, Y_{i_1}^3, Y_{i_2}^1, Y_{i_2}^2, Y_{i_2}^3)$ is a 6-variate Poisson with all marginal distributions mutually independent; $\mathbf{Z}_i = (Z_{i_1}, Z_{i_2})$ is bivariate independent binomial with parameters $(Y_{i_1}^1, p_{i_1}, Y_{i_2}^1, p_{i_2})$ and independent of $\mathbf{Y}_i$, where (p_{i_1}, p_{i_2}) are the conversion rates (probabilities) of teams (i_1, i_2) in match i. Let the mean (or rate) of $(Y_{i_1}^j, Y_{i_2}^j)$ be

$$(\mu_{i_1}^j = \delta_{i_1}^j \alpha_{i_1}^j / \beta_{i_2}^j, \mu_{i_2}^j = \alpha_{i_2}^j / \beta_{i_1}^j)$$

for all $j = 1, 2, 3$, where $\alpha_{i_1}^j$ is the type j attacking strength of team i_1 and $\beta_{i_2}^j$ is the type j defensive strength of team i_2, and $\delta_{i_1}^j$ specifies the home advantage of team i_1 for score type j, noting that in the data if there is a home team this is always the first named team (indexed i_1), and when there is no home team this parameter is absent from the parameter specification.

The expected scores, μ_{i_1} and μ_{i_2}, of teams i_1 and i_2 in this match are then given by

$$\mu_{i_1} = (5 + 2p_{i_1}) \delta_{i_1}^1 \alpha_{i_1}^1 / \beta_{i_2}^1 + 3\delta_{i_1}^2 \alpha_{i_1}^2 / \beta_{i_2}^2 + 3\delta_{i_1}^3 \alpha_{i_1}^3 / \beta_{i_2}^3,$$

and

$$\mu_{i_2} = (5 + 2p_{i_2}) \alpha_{i_2}^1 / \beta_{i_1}^1 + 3\alpha_{i_2}^2 / \beta_{i_1}^2 + 3\alpha_{i_2}^3 / \beta_{i_1}^3,$$

where the terms here correspond to the expected points-value of tries, penalties and drop-goals, respectively. In the first term in each equation, the expected number of tries scored (the try-strength ratio) is multiplied by the expected points-value of a try (5 plus the expected value of the conversion, $2p$).

The log-likelihood for the data $(\mathbf{y}, \mathbf{z})$ (the M by 8 matrix of scores) is

$$
\begin{aligned}
l(\mathbf{y}, \mathbf{z}) = \sum_{i=1}^{M} \Bigg\{ & \sum_{j=1}^{3} \left\{ -\mu_{i_1}^j - \mu_{i_2}^j + y_{i_1}^j \log \mu_{i_1}^j + y_{i_2}^j \log \mu_{i_2}^j - \log(y_{i_1}^j!) - \log(y_{i_2}^j!) \right\} \\
& + \log(y_{i_1}^1!) - \log(z_{i_1}!) - \log\{(y_{i_1}^1 - z_{i_1})!\} + z_{i_1} \log p_{i_1} + (y_{i_1}^1 - z_{i_1}) \log(1 - p_{i_1}) \\
& + \log(y_{i_2}^1!) - \log(z_{i_2}!) - \log\{(y_{i_2}^1 - z_{i_2})!\} + z_{i_2} \log p_{i_2} + (y_{i_2}^1 - z_{i_2}) \log(1 - p_{i_2}) \Bigg\} \\
= C + \sum_{i=1}^{M} \Bigg\{ & \sum_{j=1}^{3} \left\{ -\mu_{i_1}^j - \mu_{i_2}^j + y_{i_1}^j \log \mu_{i_1}^j + y_{i_2}^j \log \mu_{i_2}^j \right\} \\
& + z_{i_1} \log p_{i_1} + (y_{i_1}^1 - z_{i_1}) \log(1 - p_{i_1}) + z_{i_2} \log p_{i_2} + (y_{i_2}^1 - z_{i_2}) \log(1 - p_{i_2}) \Bigg\}.
\end{aligned}
$$

So in the first row of this equation, we have the Poisson scores for each team and for each of the three score-types, in the second row we have the binomial number of conversions for team 1 (since $Z_{i_1} \sim \text{bin}(Y_{i_1}^1, p_{i_1})$), and in the third row the same for team 2.

For the simplest parameterisation, there are distinct parameters (α_k^1, β_k^1) for the try-attacking (type 1) and try-defensive (type 1) strengths of all teams $k = 1,...,n$, up to the one required constraint (e.g. $\alpha_1^1 = 1$); the (type 2) penalty-strengths are (α_k^2, β_k^2) with e.g. $\alpha_1^2 = 1$. The simplification $\alpha_k^2 = \theta \alpha_k^1$ and $\beta_k^2 = \theta \beta_k^1$, so that for every team in every match the penalty scoring

rate is proportional to the try-scoring rate, is appealing from an parameter reduction point of view but not justified by the exploratory data analysis in Section 4.1. Likewise for drop goals, the parameter reduction that $\alpha_k^3 = \phi\alpha_k^1$ and $\beta_k^3 = \phi\beta_k^1$ would imply is not justified because the number of drop goals is not likely to be linearly related to the number of tries. This is because, if a dominant team is scoring many tries, it may not attempt drop-goals, and conversely, in a tight match, drop-goal attempts may be more frequent particularly in the latter stages of a match. So, with a conversion parameter for each team, and home advantage specified for each score type j as $\delta_{i_1}^j = \delta$ if team i_1 is at home and $\delta_{i_1}^j = 1$ otherwise, we estimate 140 parameters (Table 3).

Table 3. Maximum likelihood estimates of team strength parameters and conversion rate; 942 matches in total in 2006-2017; the five-number-summary for the numbers of matches played by each team is (22-64-92-132-147).

TEAM	Try-attack strength α_k^1	Try-defence strength β_k^1	Penalty-attack strength α_k^2	Penalty-defence strength β_k^2	Drop-goal-attack strength α_k^3	Drop-goal defence-strength β_k^3	Conversion rate p_k
New Zealand	1.000	0.468	1.000	0.498	1.000	19.204	0.737
Wales	0.470	0.273	1.060	0.412	1.499	12.934	0.787
England	0.546	0.348	1.064	0.387	3.027	9.943	0.720
Australia	0.717	0.365	0.950	0.410	1.596	11.366	0.722
South Africa	0.719	0.358	1.044	0.450	2.788	11.978	0.757
France	0.493	0.297	1.037	0.454	2.415	10.590	0.744
Ireland	0.519	0.327	0.902	0.473	1.322	11.888	0.755
Scotland	0.327	0.246	1.034	0.422	2.490	14.890	0.781
Italy	0.294	0.172	0.813	0.401	2.380	27.390	0.732
Argentina	0.419	0.248	1.073	0.397	2.799	14.096	0.767
Japan	0.299	0.103	0.904	0.529	0.774	62.519	0.676
Fiji	0.374	0.132	0.722	0.464	0.010	13.728	0.733
Romania	0.151	0.098	0.924	0.401	1.823	24.471	0.674
Canada	0.258	0.106	0.817	0.408	1.188	30.963	0.696
Samoa	0.328	0.177	0.930	0.397	0.741	13.799	0.665
USA	0.235	0.097	0.730	0.430	2.156	35.468	0.668
Georgia	0.157	0.135	0.927	0.366	1.102	11.546	0.770
Tonga	0.313	0.139	0.984	0.442	0.311	22.399	0.681
Uruguay	0.107	0.069	0.668	0.633	0.519	56.646	0.600
Namibia	0.203	0.063	0.556	0.509	2.667	10.506	0.655
HOME ADV.	1.212		1.227		0.874		

It is apparent in Table 3 that New Zealand has the largest try-attack strength and try-defence strength. The other southern hemisphere teams are the next strongest, followed by the "home nations" and Argentina. Penalty-attack and penalty-defence strengths are more narrowly distributed across the top nine teams, indicating that these teams expect to score similar numbers of penalties in matches between them. So the utility of the penalty as a discriminator in matches is questionable. Teams outside the top nine teams are some way behind on all these measures. Home advantage appears to be quite large for tries and penalties, and the home disadvantage with respect to drop-goals may be an artefact of low numbers of such events. However, for prediction purposes, one might specify team-related home advantages, since travelling around the world is a different prospect to travelling between countries in the UK.

4.3. *Simulation of the RWC tournament*

We use the model above and its parameter estimates to simulate the RWC 2015 tournament, in a similar manner to that used in soccer tournament simulations (e.g. Leitner et al., 2010; Suzuki et al., 2010; Scarf and Yusof, 2011; Goossens et al., 2012). Details of the tournament are here: https://en.wikipedia.org/wiki/2015_Rugby_World_Cup. In summary, the first round was a round robin in each of 4 pools of 5 teams, with the top two teams qualifying for the seeded knock-out phase with three rounds: quarter finals, semi-finals, and final. The simulation replicated the tournament rules and circumstances:

- The assignment of teams to pools was the same as RWC 2015 and fixed over repetitions in the simulation.

- England was the home team and played all its games at home; other teams played all games away from home except Wales which played Fiji and Uruguay at home in the pool stage.

- Match results were simulated using the model described in Section 4.2 and the parameters in Table 3.

- In the pool phase, 4 points were awarded for a win, 2 for a draw, 1 for scoring four tries or more, and 1 point for the losing side if the points difference was 7 or less.

- Progression from the pool when two teams were tied on points used a succession of tie breakers: i) the winner of the match between the teams, ii) the points-difference (points-for minus points-against in all pool matches, iii) the try-difference, iv) points scored, v) tries scored, and vi) the IRF ranking as at 12.10.2015.

- The assignment of teams to the knock-out rounds was deterministic, with the winner of pool A playing the runner-up of pool B in the Q1, etc.

- The only departure from the tournament rules was that for determining the winner when a match was tied at full-time, with the winner determined purely at random (coin-toss).

- The tournament simulation was repeated a number of times that was sufficient for the proportions to reach stable values (10,000 repetitions).

We proxy suspense using the proportion of matches tied at half-time. To calculate this measure, matches were simulated as the sum of two-halves, with a match-half simulated by setting the attack parameters to half the value in Table 3.

In addition, a home and away round-robin (H&A RR) tournament was simulated 1000 times in order to calculate team ratings. Using these ratings, a surprise result was defined as a match won by a team 20 points weaker than its opponent. In this way we use a much simpler notion of surprise than that defined by Ely et al. (2015).

We acknowledge that a more sophisticated tie-breaker (e.g. Koning et al., 2003) could have been used, but again our view is that this is not necessary in the context of our study.

4.4. *Scoring scenarios*

We simulated the RWC 2015 tournament in each of four scenarios:

- Scenario 1, the status quo, with a try worth 5 points, a conversion 2, a penalty 3, and a drop goal 3 points, and scoring parameters as in Table 3.

- Scenario 2, akin to 1870s scoring rules, with a try worth 0, a conversion 3, a penalty 0, and a drop goal 1 points, and scoring parameters as in Table 3.

- Scenario 3, similar but mimicking a drop-kicked conversion rather than a place-kicked conversion, with points as scenario 2, and scoring parameters as in Table 3, but with the conversion-rates halved.

- Scenario 4, mimicking 1960s scoring rates, with points as scenario 1, but with attack parameters for each score type halved, and conversion rates reduced by 25%.

Effectively, in scenario 2, the team with the larger number of converted tries in a match wins, and typical expected scores are half those of the status quo, with e.g. 4-1 for NZ v Wales, 2-1 for Australia v England, 2-2 for France v Ireland. Under scenario 3, expected scores take values that are half these. In scenario 4, expected scores lie closer to scenario 2 than scenario 3 since penalties still count but overall scoring rates are less than half those of the status quo.

Returning to the point about modelling refinements discussed in the opening paragraph of Section 4.1, we contend that these are not necessary because if there were no points for a penalty and 1 point for a conversion the game itself would change, and we would anticipate that try scoring rates would change in a manner that is not known. Such a reaction would likely modify scoring rates to a greater extent than a modelling refinement. Thus, we are not attempting to predict RWC tournament outcomes under a points-rules change. Instead, we are using this tournament to illustrate how rule changes can influence outcomes in an idealised tournament.

4.5. *Tournament simulation results*

Table 4 presents our results, which shows for each team for each scoring scenario:

- The value of the win percentage in the home and away round-robin ("Win % in H&A RR").

- The proportion of repetitions in which the team qualified for the knock out phase ("Qualify for KO").

- The proportion of repetitions in which the team reached the semi-final round ("Reach semis").

- The proportion of repetitions in which the team won the tournament ("Win tournament").

Also shown are the tournament win probabilities calculated using the pre-tournament betting odds ("starting prices").

We can see in the H&A RR ratings that in the scenarios with reduced scoring rates the dominant teams remain dominant but to a lesser degree, and remain dominant but to the least degree in the scenario that mimics the drop-kicked conversion. This is apparent in Figure 7. Thus, team rankings are not changed by reducing the scoring rates, and for example New Zealand is the best team in the world in each scenario. In scenario 1, we can see the effect of home advantage for England since England has higher progression probabilities than Australia who are ranked higher. Under the scenarios with no penalty scoring, England does less well relative to Australia who have a larger try attack strength. These are just some examples of effects that are apparent at the top of the table.

Table 4. Estimated tournament outcome probabilities for 10000 repetitions of the RWC 2015 simulation (= 470,000 matches) and 1000 repetitions of the home and away round robin simulations (= 380,000 matches).

|	Scenario 1: try 5, con. 2, pen. 3, drop 3					Scenario 2: try 0, con. 3, pen. 0, drop 1				Scenario 3: try 0, con. 3, pen. 0, drop 1 con.rate x 0.5				Scenario 4: try 5, con. 2, pen. 3, drop 3 try x 0.5, pen. x 0.5 con.rate x 0.75			
	Win % in H&A RR	Qualify for KO	Reach semis	Win tournament	Starting "prices"	Win % in H&A RR	Qualify for KO	Reach semis	Win tournament	Win % in H&A RR	Qualify for KO	Reach semis	Win tournament	Win % in H&A RR	Qualify for KO	Reach semis	Win tournament
New Zealand	92.6	1.000	0.853	0.503	0.369	91.5	0.999	0.828	0.457	86.0	0.990	0.758	0.355	88.2	0.993	0.790	0.397
South Africa	83.7	0.990	0.569	0.126	0.160	83.2	0.985	0.576	0.149	78.5	0.956	0.562	0.143	80.0	0.958	0.553	0.149
Australia	81.6	0.812	0.543	0.116	0.080	81.4	0.834	0.578	0.151	76.7	0.791	0.493	0.141	78.4	0.762	0.484	0.121
England	76.0	0.824	0.558	0.130	0.175	75.5	0.771	0.473	0.095	71.2	0.693	0.423	0.099	72.9	0.781	0.518	0.130
Ireland	74.3	0.929	0.418	0.052	0.097	73.7	0.912	0.411	0.058	69.7	0.822	0.394	0.086	69.8	0.842	0.388	0.065
France	72.8	0.910	0.394	0.041	0.042	70.8	0.873	0.377	0.045	67.2	0.786	0.367	0.059	68.9	0.837	0.385	0.064
Wales	69.0	0.351	0.152	0.013	0.052	69.3	0.378	0.172	0.015	65.6	0.457	0.222	0.042	67.4	0.409	0.185	0.023
Argentina	64.6	0.804	0.304	0.013	0.011	64.6	0.836	0.338	0.022	62.6	0.777	0.362	0.046	61.8	0.723	0.329	0.033
Scotland	59.1	0.631	0.144	0.006	0.003	58.8	0.630	0.157	0.008	57.3	0.582	0.204	0.019	56.7	0.576	0.178	0.014
Samoa	47.4	0.287	0.029	0	0.006	46.1	0.306	0.041	0.000	46.5	0.300	0.067	0.004	47.1	0.294	0.061	0.003
Italy	43.7	0.140	0.007	0	0.001	46.7	0.185	0.021	0.001	47.8	0.294	0.060	0.004	44.4	0.241	0.041	0.001
Tonga	42.9	0.174	0.022	0	0.001	40.1	0.133	0.021	0	41.3	0.154	0.041	0.002	44.6	0.224	0.053	0.001
Fiji	42.8	0.013	0.001	0	0.001	44.6	0.017	0.002	0	44.5	0.059	0.012	0.001	43.9	0.048	0.010	0
Japan	34.4	0.077	0.002	0	0	31.6	0.056	0.002	0	34.9	0.098	0.009	0	36.6	0.127	0.011	0
Canada	28.8	0.018	0	0	0	29.9	0.026	0.001	0	33.2	0.075	0.007	0.001	31.3	0.061	0.003	0
Georgia	24.7	0.022	0.001	0	0	27.5	0.030	0.002	0	31.6	0.064	0.010	0	30.1	0.053	0.009	0
USA	24.2	0.016	0	0	0	25.4	0.024	0	0	29.9	0.064	0.009	0	26.4	0.045	0.001	0
Romania	18.1	0.003	0	0	0	18.4	0.004	0	0	24.1	0.025	0.001	0	23.2	0.020	0.001	0
Namibia	11.6	0	0	0	0	13.3	0.002	0	0	17.7	0.015	0.001	0	15.5	0.007	0	0
Uruguay	7.6	0	0	0	0	7.4	0	0	0	13.4	0	0	0	12.9	0	0	0
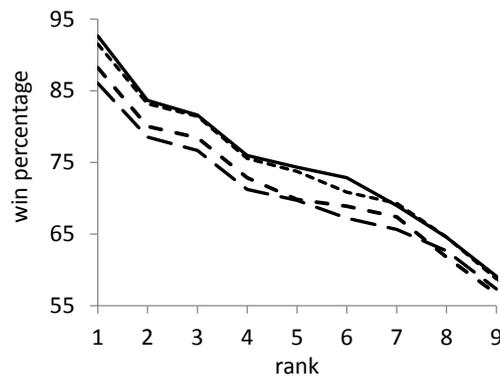
Figure 7. Team strengths (win percentage in H&A RR) of best 9 teams for each scoring system. Solid line, scenario 1 (status quo); short dash, scenario 2 (3 points for conversion, 1 for drop goal); long dash, scenario 3 (drop-kicked conversion); medium dash, scenario 4 (1960s scoring rates).

More importantly perhaps, weaker teams have greater progression probabilities in the scenarios with reduced scoring rates, and there are correspondingly more surprises, more tied matches, and more matches tied at half time, suggesting more suspense (Table 5). We qualify these points by noting that surprise and suspense can be measured in different ways. For example, matches in-play that are close rather than just tied might be deemed suspenseful, and smaller or larger "strength" differences might be used in the definition of surprise. Nonetheless, we would expect the effects to remain the same. Also important is that the try-scoring rate in scenarios 2 and 3 is same as the status quo, so that more surprises and more suspense, and greater outcome uncertainty, does not necessarily imply fewer tries. Thus, the administrator may not have to trade-off exciting play (i.e. tries) against outcome uncertainty, two goods that are valued by consumers of sport. Nonetheless, the direction that the game would take in response to such a radical rule-change is unknown, and scenarios can be envisaged in which the try-scoring rate increases and in others it decreases.

Table 5. Proportion of matches tied at half time, at full time (c.f. 7 out of 20 soccer World Cup finals have been tied at full time = 0.350), and proportion of matches with surprise result (when a team beats another whose H&ARR win% is 20 points greater)

		Scenario 1	Scenario 2	Scenario 3	Scenario 4
All matches	tied at half time	0.045	0.236	0.403	0.139
	tied at full time	0.016	0.111	0.213	0.045
	surprise results	0.050	0.090	0.172	0.081
Final	tied at half time	0.050	0.304	0.418	0.160
	tied at full time	0.030	0.149	0.248	0.053

5. Discussion

We study the relationship between outcome uncertainty and scoring rates and focus on the particular case of international rugby union. In this sport, scoring rates have increased over the last 50 years, with more tries scored, more penalty-kicks attempted and more kicks successful. However, this increase in scoring rate may be at the expense of outcome uncertainty, and we show that the outcome of a Poisson match (an idealised model of a ball game between two teams) becomes more certain in favour of the stronger team as the overall scoring rate increases.

Furthermore, less outcome uncertainty appears to imply less suspense and fewer surprises. We argue therefore that a high scoring rate may be detrimental to the development of a sport that approximates to the Poisson match, and may be detrimental to the development of rugby union. The implication is that it is harder for new international teams to break into the top flight if scoring rates are higher, and that very often weak international teams will be beaten convincingly, by large margins, by strong teams in tournaments. A significant proportion of matches (close to half of all games) in these tournaments are hopelessly one-sided. It can be argued that low outcome uncertainty is not a systemic issue but the result of few teams competing at the highest level. But this is a vicious circle, and may be preventing the entry of the full game to the Olympic programme, and may even perpetuate issues of welfare of players of the sport. And on the contrary, that soccer is so popular to watch is not a surprise given that scoring rates in this sport are close to those which maximise outcome uncertainty in the Poisson match. That said, there exist studies that show outcome uncertainty is not necessarily desired by fans (e.g. Buraimo and Simmons, 2015).

We suggest that the crux of the issue in rugby union is that, as kicking has improved, the size of the target has remained the same. Thus more kicks are successful now than in the past. We further suggest that the response of administrators has been to incentivise more exciting play and hence more try-scoring. It is our thesis that a policy to reduce scoring rates should be considered. This could be achieved by redefining the try as just that, a try for goal, an attempt for a point. We study this policy by modelling a scenario in which the try effectively counts for a point only if it is converted. We then illustrate, through simulation of the Rugby World Cup tournament under various reduced-scoring-rate scenarios, that fewer scores leads to more suspense and more surprises, but that the rank-order of teams remains unchanged. In this way, the best would appear to remain the best while winning less often. Furthermore, and most importantly, under the most extreme reduced scoring-rate scenario, more outcome uncertainty may not occur at the expense of less exciting play because while the points-scoring rate is lower the try-scoring rate is not. Finally, an alternative to scoring-rule change that might be considered is a smaller tournament of say 12 or 16 teams. However, this would negate widening participation.

There are limitations to our claims. Firstly, we assume try-scoring rates are fixed across the various scenarios considered. In reality, the effect we desire (more outcome uncertainty) is likely to be moderated by an increase in try-scoring that a decrease in penalty scoring (to zero) will imply. Next, rugby union is not closely approximated by a Poisson match. Nonetheless, we might expect it to approximate to the Poisson match somewhat better under the "one point for a conversion" scenario. Also, penalty- and try-scoring events are negatively dependent, and so we would expect try-scoring to increase if penalty-scoring decreases even without a change in playing-behaviour in response to rule-changes. Next, fewer points scored implies more tied matches, although there is the suggestion that even under the most extreme score-reduction scenario, the proportion of tied matches would not be as great as occurs in soccer. Finally, changing the point-scoring rules will likely change the game (Hogan and Massey, 2017) in a way that is unforeseen and unintended (Hon and Parinduri, 2016; Kendall and Lenten, 2017). However, this is not an issue that an operational research analysis of the game can address, because inevitably a model can only provide an approximation to a future reality. Modellers have to persuade administrators about the quality of approximation, and administrators have to use their judgement. The role of this paper is merely to use analysis to inform discussion of matters of importance to sport in general and outcome uncertainty in particular.

References

Alavy K, Gaskell A, Leach S and Szymanski S (2010) On the edge of your seat: demand for football on television and the outcome uncertainty hypothesis. *International Journal of Sport Finance* 5, 75-95.

Baker RD and McHale IG (2013) Forecasting exact scores in National Football League games. *International Journal of Forecasting* 29, 122-130.

Baker RD and McHale IG (2018) Time-varying ratings for international football teams. *European Journal of Operational Research* 267, 659-666.

Bingham NH, Kiesel R and Schmidt R (2003) A semi-parametric approach to risk management. *Quantitative Finance* 3, 426-411.

Bizzozero P, Flepp R and Franck E (2016) The importance of suspense and surprise in entertainment demand: Evidence from Wimbledon. *Journal of Economic Behavior and Organization* 130, 47-63.

Boshnakov G, Kharrat T and McHale IG (2017) A bivariate Weibull count model for forecasting association football scores. *International Journal of Forecasting* 33, 458–466.

Borland J and Macdonald R (2003) Demand for sport. *Oxford Review of Economic Policy* 19, 478-502.

Buraimo B and Simmons R (2015) Outcome uncertainty or star quality? Television audience demand for English premier league football. *International Journal of the Economics of Business* 22, 449-469.

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

Ely J, Frankel A and Kamenica E (2015) Suspense and surprise. *Journal of Political Economy* 123, 215-260.

Forrest D and Simmons R (2002) Outcome uncertainty and attendance demand in sport: The case of English soccer. *Journal of the Royal Statistical Society: Series D* 51, 229–241.

Goossens DR, Beliën J and Spieksma FCR (2012) Comparing league formats with respect to match importance in Belgian football. *Annals of Operational Research* 194, 223-240.

Hogan V, Massey P and Massey S (2013) Competitive balance and match attendance in European rugby union leagues. *The Economic and Social Review* 44, 425-446.

Hogan V and Massey P (2017) Team responses to changed incentives: evidence from rugby's Six Nations Championship. *International Journal of Sport Finance* 12, 140-159.

Hogan V, Massey P and Massey S (2017) Analysing match attendance in the European Rugby Cup: Does outcome uncertainty matter in a multinational tournament? *European Sport Management Quarterly* in press.

Hon LY and Parinduri RA (2016) Does the three-point rule make soccer more exciting? Evidence from a regression discontinuity design. *Journal of Sports Economics* 17, 377-395.

Karlis D and Nzoutfras I (2003) Analysis of sports data by using bivariate Poisson models. *Journal of the Royal Statistical Society Series D* 52, 381-393.

Kendall G and Lenten LJA (2017) When sports rules go awry. *European Journal of Operational Research* 257, 377-394.

Koning RH, Koolhaas M, Renes G and Ridder G (2003) A simulation model for football championships. *European Journal of Operational Research* 148, 268-276.

Kuchar R and Martin A (2016) The Comparison of competitive balance between Super Rugby (Sanzar) and English Premiership rugby: a case study from 1996-2014 season or not attractive – no people – no money. *International Journal of Entrepreneurial Knowledge* 4, 112-129.

Leeds MA and von Allmen P (2014) *The Economics of Sports*, 5th ed. Boston: Pearson.

Leitner C, Zeileis A and Hornik K (2010) Forecasting sports tournaments by ratings of (prob)abilities: A comparison for the EURO 2008. *International Journal of Forecasting* 26, 471-481.

Lenten LJA and Winchester N (2015) Secondary behavioural incentives: bonus points and rugby professionals. *Economic Record* 91, 386-398.

Maher M (1982) Modelling association football scores. *Statistica Neerlandica* 36, 109-118.

Manasis V, Avgerinou V, Ntzoufras I and Reade JJ (2013) Quantification of competitive balance in European football: development of specially designed indices. *IMA Journal of Management Mathematics* 24, 363-375.

McHale IG and Scarf PA (2007) Modelling soccer matches using bivariate discrete distributions with general dependence structure. *Statistica Neerlandica* 61, 432-445.

McHale I and Scarf PA (2011) Modelling the dependence of goals scored by opposing teams in international soccer matches. *Statistical Modelling* 11, 219–236.

Mutz M and Wahnschaffe K (2016) The television viewer's quest for excitement – does the course of a soccer game affect TV ratings? *European Journal for Sport and Society* 13, 325-341.

Owen A (2011) Dynamic Bayesian forecasting models of football match outcomes with estimation of the evolution variance parameter. *IMA Journal of Management Mathematics* 22, 99–113.

Owen PD, Ryan M and Weatherston CR (2007) Measuring competitive balance in professional team sports using the Herfindahl-Hirschman Index. *Review of Industrial Organization* 31, 289- 302.

Pledger MJ and Morton RH (2011) Modelling the 2004 Super 12 rugby union competition. *Australian and New Zealand Journal of Statistics* 53, 109-121.

Rottenberg, S. (1956) The baseball players' labor market. *Journal of Political Economy* 64, 242-258.

Scarf PA, Yusof MM and Bilbao M (2009) A numerical study of designs for sporting contests. *European Journal of Operational Research* 198, 190-198.

Scarf PA and Yusof MM (2011) A numerical study of tournament structure and seeding policy for the soccer World Cup Finals. *Statistica Neerlandica* 65, 43-57.

Schreyer D, Schmidt SL and Torgler B (2017) Game outcome uncertainty in the English Premier League. *Journal of Sports Economics* in press.

Stefani R (2009) Predicting score difference versus score total in rugby and soccer. *IMA Journal of Management Mathematics* 20, 147-158.

Suzuki AK, Salasar LEB, Leite JG and Louzada-Neto F (2010)A Bayesian approach for predicting match outcomes: The 2006 (Association) Football World Cup. *Journal of the Operational Research Society* 61, 1530-1539.

Szymanski S (2003) The economic design of sporting contests. *Journal of Economic Literature* 41, 1137-1187.

Utt J and Fort R (2002) Pitfalls to measuring competitive balance with Gini coefficients. *Journal of Sports Economics* 3, 367-373.

Winchester, N (2008). Shifting the 'goal posts': optimizing the allocation of competition points for sporting contests. *Journal of Quantitative Analysis in Sports*, 4(4).

Wright M (2014) OR analysis of sporting rules–a survey. *European Journal of Operational Research* 232, 1-8.

Appendix 1

Let $X_1 \sim \text{Po}(\lambda)$ and $X_2 \sim \text{Po}(\lambda\varepsilon)$ independent with $\varepsilon > 1$. The score difference $X_1 - X_2$ can be approximated by a random variable Y such that $Y \sim \text{N}(\lambda(1-\varepsilon), \lambda(1+\varepsilon))$. This is because X_1 can be approximated by a random variable $Y_1 \sim \text{N}(\lambda, \lambda)$ and X_2 can be approximated by a random variable $Y_2 \sim \text{N}(\lambda\varepsilon, \lambda\varepsilon)$, and so $Y = Y_1 - Y_2$ is normally distributed with mean

$$E(Y) = E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = \lambda - \lambda\varepsilon = \lambda(1-\varepsilon)$$

and variance

$$\text{var}(Y) = \text{var}(Y_1 - Y_2) = \text{var}(Y_1) + \text{var}(Y_2) = \lambda + \lambda\varepsilon = \lambda(1+\varepsilon).$$

Then, using the continuity correction when calculating probabilities, we have

$$\Pr(X_1 - X_2 = 0) \simeq \Pr(-\tfrac{1}{2} < Y < \tfrac{1}{2}) = \Phi\left(\frac{\tfrac{1}{2} - \lambda(1-\varepsilon)}{\sqrt{\lambda(1+\varepsilon)}}\right) - \Phi\left(\frac{-\tfrac{1}{2} - \lambda(1-\varepsilon)}{\sqrt{\lambda(1+\varepsilon)}}\right) \to 0,$$

as $\lambda \to \infty$, and

$$\Pr(X_1 - X_2 < 0) \simeq \Pr(Y < -\tfrac{1}{2}) = \Phi\left(\frac{-\tfrac{1}{2} - \lambda(1-\varepsilon)}{\sqrt{\lambda(1+\varepsilon)}}\right) \to 1$$

as $\lambda \to \infty$ because $\varepsilon > 1$.

   Notice that when $\varepsilon = 1$,

$$\Pr(X_1 - X_2 < 0) \simeq \Pr(Y < -\tfrac{1}{2}) = \Phi\left(-\frac{1}{2\sqrt{2\lambda}}\right) \to \Phi(0) = \tfrac{1}{2}$$

as $\lambda \to \infty$, as we would expect since then the outcome is perfectly balanced.