



BUSINESS SCHOOL  
Te Kura Pakihi

ISSN 1178-2293 (Online)

**University of Otago  
Economics Discussion Papers  
No. 1311**

August 2013

---

## **Measurement of Competitive Balance and Uncertainty of Outcome**

P. Dorian Owen  
Department of Economics, University of Otago

Prepared for:

*Handbook on the Economics of Professional Football*, John Goddard and Peter J. Sloane (eds),  
Edward Elgar, Cheltenham

*Address for correspondence:*

Dorian Owen  
Department of Economics  
University of Otago  
PO Box 56  
Dunedin 9054  
New Zealand

Tel.: +64 3 479 8655

Email address: [Dorian.Owen@otago.ac.nz](mailto:Dorian.Owen@otago.ac.nz)

## **Introduction**

Competitive balance (CB) and uncertainty of outcome (UO) are central concepts in the economic analysis of sports leagues worldwide. As well as featuring prominently in the academic literature, CB provokes strongly held opinions among media commentators, sports administrators and fans. The aims of this chapter are to examine how these concepts have been measured in practice, to review recent attempts to address some of the problems with commonly used measures, and to consider some of the factors involved in selecting useful measures.

CB relates to how evenly teams are matched. A league in which teams display a high degree of variation in their playing strengths is considered to have a lower degree of CB (or, equivalently, a higher degree of ‘imbalance’) than a league in which teams’ strengths are closer to equality.

CB is considered important because it affects the degree of uncertainty surrounding the outcomes of individual matches and overall championships. A match involving teams with significant differences in their strengths is more likely, other things equal, to result in a win for the significantly stronger team. Similarly, the identity of the overall championship-winning team is more predictable if one or a small number of teams is significantly stronger than the rest of the teams in a league.

The most contentious step in the justification for the importance of CB is embodied in Rottenberg’s (1956) ‘uncertainty of outcome’ hypothesis (UOH). According to the UOH, a higher degree of predictability (i.e., reduced uncertainty) erodes spectators’ interest. This decreases consumer demand for the league’s output, reflected in lower match attendances, lower television audience ratings and reduced opportunities for sponsorship and merchandising. Indeed, Rottenberg (1956, p.258) characterizes sports leagues as “industries ... in which all firms [teams] must be nearly equal if each is to prosper”.

What constitutes an optimal level of CB is a contentious and under-researched question and the answer is likely to vary from one league to another (Fort and Quirk, 2010, 2011). Amongst economists, it is not a widely held view that the ‘ideal’ degree of CB involves perfect balance, with all teams of equal strength. As Zimbalist (2002, p.111) observes, “[e]veryone agrees [CB] is a good thing to have, but no one knows how much one needs”. Similarly, Szymanski (2006, p.597) emphasizes that “even if it were true that a completely predictable contest would be unattractive, it is hard to say what the effect of a small change in balance would be, starting from a given distribution of wins”. Despite the ambiguity over how much CB is optimal, disquiet is frequently expressed about the inadequacy of observed degrees of CB in football leagues, with the English Premier League (EPL), the “world’s most watched league” (Premier League, 2011), a frequent target:

It's the most competitive league in the world. Anybody can beat anybody on their day. There are no easy games in this league. The mantra of the Premier League apologists is well known. ... It's nonsense, of course: it's obvious the Premier League is a closed shop that can be opened only with the application of around a quarter of a billion pounds. (Wilson, 2011)

Wilson’s comments illustrate a widespread concern. If owners of high-revenue teams with a large fan base (or billionaire owners with deep pockets) are aiming to maximize their chances of winning competitions, they can, in a free market, use their greater resources to attract better-quality players to enhance the strength of their playing squads. Other things equal, this reduces CB and UO and, if the UOH is to be believed, will diminish interest in the league.

Perceptions of inadequate levels of CB lead to frequent calls for policy interventions, such as salary caps, various forms of revenue sharing and financial ‘fair play’ rules, to attempt to ‘improve’ the level of CB. Evaluation of such proposals requires empirical evidence on the relevance of the UOH and appropriate measurement of CB and UO.

Measurement of CB and UO has therefore attracted considerable attention in the sports economics literature and in policy debate.<sup>1</sup>

CB and UO are closely interrelated, and are often treated as interchangeable, but there are some differences. Following Kringstad and Gerrard (2007), UO is more sensibly treated as a forward-looking ex ante concept, defined in terms of a probability distribution of possible outcomes. In contrast, although perfect CB is theoretically better expressed in terms of equality of outcome probabilities rather than equality of actual outcomes, evaluation of CB commonly relies on backward-looking ex post measures (often based on the distribution of win ratios, points ratios or other characteristics of observed outcomes).

This difference in emphasis is related to Fort and Maxcy's (2003) characterization of the CB literature as consisting of two distinct strands. One strand, which they label the 'analysis of competitive balance' (ACB) literature, is concerned with assessing the extent of CB and tracking its movements over seasons. It also involves examining the effects on CB of regulatory and other institutional changes, thereby testing Rottenberg's (1956) 'invariance principle', which hypothesizes that outcomes such as the distribution of wins in a league are unaltered by reallocating property rights. Ex post CB measures are commonly used for such analyses.

The other strand of the literature examines the effects of CB on fan demand. Because demand is hypothesized to depend on the expected characteristics of future competition, UO is regarded as a key determinant. Correspondingly, the relevant UO metrics are ideally forward looking; however, in practice, these are often proxied by backward-looking measures, especially if taking a longer-run perspective.

This chapter examines measures of different dimensions of CB, used primarily in the ACB literature, followed by consideration of different approaches to characterizing UO in demand studies. The principal theme is that standard statistical measures of dispersion,

concentration and inequality are not always ideal CB measures because they ignore specific design characteristics of football leagues, in particular the details of playing schedules, differences in points assignment systems, the common occurrence of draws, the existence of home advantage and the structure of multiple prizes. Recent studies emphasizing the implications of these characteristics for CB measurement are reviewed. Potential avenues for future work to improve measurement of CB and OU are discussed and some broad principles to consider when selecting useful measures are suggested.

### **Static Measures of Competitive Balance**

CB measurement attempts to quantify how closely teams are matched. Even for simple league structures, it is widely accepted that CB (and UO) are multi-dimensional and that no single measure can encapsulate everything there is to know about CB (Sloane, 1971; Cairns et al., 1986; Vrooman, 1996; Szymanski, 2003). Kringstad and Gerrard (2007) identify three key dimensions of CB: ‘win dispersion’ (the distribution of wins, or more generally points, across teams within a single season), ‘performance persistence’ (the persistence over successive seasons of each team’s points outcomes), and ‘prize concentration’ (the distribution of prizes, such as championships, across teams over successive seasons).

Similarly, three dimensions of UO are commonly distinguished: match uncertainty, seasonal uncertainty and championship uncertainty (Szymanski, 2003).

#### *Variance-based measures of the dispersion of wins or points*

To evaluate the dispersion of wins in a single season, it is common to adopt standard measures of variation, inequality or concentration. For sports for which the result of each match is a win for one team and a loss for the other (i.e., there are no draws), the ex post standard deviation (*SD*) of teams’ win ratios within a single season is a natural measure of

variation. Because drawn matches are common in football, the standard deviation of teams' points ratios can be used instead, and can be represented as

$$SD = \sqrt{\sum_{i=1}^N (p_i - \bar{p})^2 / (N - 1)}$$

where  $N$  is the number of teams in the league,  $p_i$  is the points ratio for team  $i$  (the actual number of points accumulated by team  $i$  in a season divided by the maximum possible points attainable) and  $\bar{p} = \sum_{i=1}^N p_i / N$  is the league's mean points ratio.<sup>2</sup> Other things equal, a smaller standard deviation of win or points ratios across teams in a season indicates a more equal competition; the larger the value of  $SD$ , the more unbalanced is the competition.

Comparisons of  $SD$  over time for the same league or across different leagues, however, may involve differences in the number of matches played by each team. Because  $SD$  tends to decrease if teams play more games (due to a reduction in the effect of random noise on the ratios), it is usual to compare  $SD$  to a benchmark 'idealized standard deviation',  $ISD$ . This represents the standard deviation of the win or points ratios in an (ex ante) perfectly balanced league in which each team has an equal probability of winning each game (Noll, 1988; Scully, 1989; Quirk and Fort, 1992; Fort and Quirk, 1995). Although it can be highly sensitive to outliers (Lenten, 2009) the ratio of standard deviations,  $RSD = SD/ISD$ , is the most widely used measure of CB in sports leagues.

For sports with no drawn matches,  $ISD$  is derived as the standard deviation of a binomially distributed random variable with a probability of success of 0.5 across independent trials; in this case,  $ISD = 0.5/G^{0.5}$ , where  $G$  is the number of games played by each team (Fort and Quirk, 1995). Although the 'no draws' assumption is not valid for football, the conventional  $RSD$  measure has been widely applied, including for comparisons

of the dispersion of wins in football leagues with the ‘big four’ North American sports leagues (Szymanski and Smith, 2002; Buzzacchi et al., 2003; Kringstad and Gerrard, 2007).

However, if draws are frequent, use of the binomial distribution to derive *ISD* is not appropriate. The trinomial nature of match outcomes (win, loss, draw) and the points assigned to each outcome need to be considered (Cain and Haddock, 2006; Fort, 2007; Owen, 2012). The appropriate *ISD* values for points ratios for the conventional allocation of three points for a win, one point for a draw (and zero points for a loss), represented as (3,1,0), and the historically relevant (2,1,0) assignment are given by:

$$ISD(\text{Points Ratio: } 3,1,0) = \sqrt{[(1-d)(d+9)/4]/9G}$$

$$ISD(\text{Points Ratio: } 2,1,0) = \sqrt{(1-d)/4G}$$

where the probability of a draw,  $P(\text{draw}) = d$ , and  $P(\text{win}) = P(\text{loss}) = (1 - d)/2$  (Owen, 2012, p.89). If  $d = 0$ , these expressions simplify to the conventional binomial-based expression for *ISD*. If  $d \neq 0$ , use of the binomial-based *ISD* as the denominator for *RSD* underestimates the degree of imbalance (by about 10% or more) compared to using the relevant trinomial-based *ISD*. It is important, therefore, to use the relevant *ISD* if comparing *RSD* values for football leagues with other sports that have no or few draws. Different points assignments with different ratios of points for wins compared to draws give numerically different results for *RSD* (even when based on identical sets of win, draw and loss results), although numerical differences are minor (Owen, 2012). Hence, if the aim is to track variation in *RSD* over a long run of seasons, adopting a common points assignment, as Lee and Fort (2012) do, provides a reasonable approximation even if the period includes a change from (2,1,0) to (3,1,0).<sup>3</sup>

Another feature of football, as with many other sports, is home advantage. In a match involving two teams of equal strength, the team playing at home will have a higher

probability of winning.<sup>4</sup> The *ISD* measures described so far do not take home advantage into account; instead they assume a constant probability of success for each team in each match (home or away). Home advantage leads the conventional *ISD* to be biased upwards, and hence *RSD* to be biased downwards, resulting in overestimation of the degree of CB (Trandel and Maxcy, 2011). Trandel and Maxcy (2011) derive a (computationally non-trivial) variant of *ISD* that allows for home advantage. For a given number of games played, the home-advantage-corrected *ISD* is a decreasing function of the home-team winning probability,  $h$ , but the extent of the bias is relatively unaffected by  $G$ , the number of games played. For a given value of  $G$ , the bias in using the conventional measure is relatively modest (approximately 2% for  $h = 0.6$  and  $G = 40$ ). Note, however, that this adjustment treats a draw as half a win, and uses a binomial rather than the more appropriate trinomial distribution. Currently, no adjustment to *ISD* has been formally proposed that allows for both home advantage and a non-zero draw probability, although an extension of the method used by Owen (2012) incorporating historical home-win and draw probabilities would be straightforward.

Other features of the distribution of points ratios can also be informative. The differences between the top and the bottom teams in the league can be examined by, for example, averaging the points ratios of, say, the top  $R$  teams and comparing with the corresponding average points ratio for the bottom  $R$  teams. Several variations of this type are possible, depending on the particular focus of interest, and some have been used in media commentary. Wilson (2011), for example, finds that the gaps in points per game between the top and bottom of the EPL and between the fourth and fourth-from-bottom teams are larger for the EPL than for any of the other football leagues he considers, which he interprets as indicating that the EPL is the least competitive league.

### *Inequality and concentration measures*

CB is essentially concerned with how differences in team strengths are reflected in the degree of inequality of match and championship outcomes; hence, an alternative source of CB measures is the literature on income inequality and firm concentration. Some studies, such as Michie and Oughton (2004), examine the Gini coefficient based on Lorenz curves plotting the cumulative percentage of wins (or points) against the cumulative percentage of teams. The Gini coefficient represents the extent of deviation from perfect equality in the outcomes, i.e., equal win or points ratios, with larger values representing greater imbalance.

The Herfindahl-Hirschman index, *HHI*, a widely used measure of concentration in the industrial organization literature, is based on the notion of ‘market share’. When applied to the distribution of wins across teams in a particular season, market share is interpreted as the number of wins by a team in a season as a proportion of total wins in the competition (Depken, 1999):

$$HHI = \sum_{i=1}^N (W_i / \sum_{i=1}^N W_i)^2$$

where  $W_i$  is the number of wins for team  $i$ . Using points instead of wins, this can be defined with total points representing the total of points actually accumulated by all teams or the feasible maximum of available points. *HHI* has a lower bound,  $HHI^{lb} = 1/N$ , representing the case of perfect balance in terms of the shares of wins or points; consequently, Depken (1999) emphasizes the importance of controlling for  $N$  when interpreting movements in *HHI* over time. He proposes an adjusted index,  $dHHI = HHI - 1/N$ , which measures the deviation of *HHI* from its lower-bound perfect-balance value. Increases in *HHI* and *dHHI* represent decreases in CB. Michie and Oughton (2004) adopt a variant of Depken’s measure, defining their ‘*H Index*’ of CB as  $HHI/(1/N)$ . Using this measure, they chart trends from the 1950s

indicating declines in CB in the EPL, Serie A in Italy and, to lesser degree, in the German Bundesliga. They interpret movements in the index for the French Ligue 1 as possessing no clear trend, and for Spain's La Liga as demonstrating cyclical behaviour.

Although inequality and concentration measures are frequently used in CB measurement, the two concepts are not identical (Hannah and Kay, 1977). For empirical work on firm size, concentration is the more relevant concept, as the existence of several very small firms with low market shares is of minor significance. However, for the measurement of CB, inequality is the more relevant concept because very low points ratios for a number of teams would be concerning, especially as  $N$  is not large.

A concentration index, such as  $HHI$ , can be expressed as a function of both  $N$  and  $I$ , the size inequality in the variable of interest, such as market shares, wins or points (Davies, 1979; Bajo and Salas, 2002). In order to separate the effects of changes in  $N$  and  $I$  on concentration, Bajo and Salas (2002) show that the Hannah-Kay (1977) family of concentration indices, which includes  $HHI$ , is consistent with the family of generalized entropy measures of inequality (Cowell, 2011) defined, for different values of  $\theta$ , as

$$I_{GE}(\theta) = \frac{1}{\theta^2 - \theta} \left[ \frac{1}{N} \sum_{i=1}^N \left( \frac{W_i}{\bar{W}} \right)^\theta - 1 \right] \quad (\theta \neq 0, 1)$$

where  $\bar{W}$  is the mean of  $W$ . With  $\theta = 2$ ,  $HHI$  can be expressed as  $(1 + 2I_{GE}(2))/N$  (Bajo and Salas, 2002).  $HHI$  can also be expressed as  $(1 + CV^2)/N$ , where  $CV$  is the coefficient of variation (Hart, 1975); hence  $I_{GE}(2) = CV^2/2$ . This links  $CV$  to a pure inequality measure, independent of  $N$ .

Teams' shares of total points can also be used to define concentration ratios. The concentration ratio for the  $R$  most successful teams can be expressed as:

$$CR(R) = \frac{\sum_{i=1}^R P_i}{\sum_{i=1}^N P_i}$$

where  $P_i$  is team  $i$ 's points total in a season. The larger is  $CR(R)$  the greater the dominance of the top  $R$  teams, and hence the greater the degree of imbalance. Michie and Oughton (2004) compare this to the case of a perfectly balanced league in which the top  $R$  teams receive an 'equal share',  $R/N$ , of the total points. For  $R = 5$ , they define their C5 ratio as:

$$C5ICB = \frac{\sum_{i=1}^5 P_i / \sum_{i=1}^N P_i}{(5 / N)}$$

This has a value of unity if the top five teams, on average, achieve the same points total as the average for the rest of the teams. Compared to their  $H$  Index, this shows similar movements over time for the EPL, although the implied decline in CB is more marked for the C5 ratio. For a different perspective, Michie and Oughton (2005) also apply a similar measure to newly promoted teams (three in the EPL). Low values of this concentration ratio imply that promoted teams are relatively weak. For the EPL, this indicator shows a downward trend. This is interpreted to mean that promoted teams find it increasingly difficult to compete in the EPL, and that the promotion and relegation system may not be a particularly effective means of enhancing CB.

Different variants of concentration ratios are possible. Koning (2000) compares total points gained by the top  $R$  teams to the *maximum* number of points they could have won:

$$CRK(R) = \frac{\sum_{i=1}^R P_i}{\alpha R(2N - R - 1)}$$

where  $\alpha$  is the number of points for a win and assuming that each team plays every other team twice. A value of unity for  $CRK(R)$  indicates the maximum degree of dominance for the

top  $R$  teams, with decreasing values indicating that the top  $R$  teams lose more points against the weaker ( $N - R$ ) teams, consistent with a higher degree of CB.<sup>5</sup> In their examination of CB in the English Football League over a long period, Lee and Fort (2012) use  $CRK(R)$ , with  $R = 3$  and 4, to evaluate the dispersion of points for the top-ranking teams, as a complement to  $RSD$ , which reflects the *overall* dispersion of points. They also calculate  $CRK(R)$  for the *bottom*  $R$  teams in the league (again with  $R = 3$  and 4). They find that, for all three measures, CB in the EPL appears to be lowest in the most recent decade considered (2001-2007).

#### *Constraints implied by league playing schedules*

A drawback of measuring CB using standard ‘off-the-shelf’ measures of variation, inequality and concentration is that they ignore the specific design characteristics of sports leagues. A league’s playing schedules (i.e., the list of fixtures) impose an upper limit on the variance of the distribution of points and hence limit the range of feasible values of all these measures. This complicates the interpretation of CB measures, especially when comparisons (across different leagues or for the same league over time) involve differences in the numbers of teams or games played, a situation that is extremely common. For such comparisons, it is important to have a clear idea of how the range of feasible values of the chosen CB measure changes in response to different aspects of league design.

The implications of the league’s schedule of matches for interpretation of the Gini coefficient and the  $HHI$  applied to wins are examined by Utt and Fort (2002) and Owen et al. (2007) respectively. Given these measures’ emphasis on teams’ shares of wins, their focus is primarily on the fact that teams cannot win games in which they do not play; this is reflected in the existence of upper bounds for the conventional Gini and  $HHI$  measures. The upper bound occurs for the ‘most unequal distribution’ of match outcomes (Horowitz, 1997; Fort and Quirk, 1997; Utt and Fort, 2002). This involves one team winning all its games, the

second team winning all except its game(s) against the first team, and so on down to the last team, which wins none of its games. Utt and Fort (2002) numerically adjust Gini coefficients for Major League Baseball (MLB). To make the calculations for the most unequal distribution of wins tractable, they assume that playing schedules are balanced, i.e., each team plays the same number of games against each of its opponents. They show that the conventional Gini coefficient underestimates the degree of competitive imbalance, giving the impression that leagues are more balanced than they really are. For the MLB in the 1990s, decreases in CB, based on an adjusted Gini, are about double the decline in CB using the unadjusted Gini.

Adopting the same assumption of balanced playing schedules, Owen et al. (2007) derive an exact analytical expression for the upper bound for the *HHI* for wins (or points), denoted  $HHI^{ub}$ , given by:

$$HHI^{ub} = 2(2N - 1)/[3N(N - 1)]$$

This is less than unity (if  $N > 2$ ). A normalized version of *HHI*,  $HHI^*$ , can be defined (for balanced playing schedules) that adjusts for the lower and upper bounds, both of which decrease as  $N$ , the number of teams, increases:

$$HHI^* = (HHI - HHI^{lb})/(HHI^{ub} - HHI^{lb})$$

This normalized measure lies in the interval  $[0, 1]$ , with zero indicating perfect balance and one representing maximum imbalance. In an application to MLB,  $HHI^*$  very closely tracks the variation over time in the corresponding pure inequality measure,  $IGE(2)$ , even when changes in  $N$  occur.

Concentration ratios are also bounded. Manasis et al. (2011), assuming balanced playing schedules, demonstrate that Michie and Oughton's (2004) *C5ICB* measure has an

upper bound,  $C5ICB^{ub} = (2N - R - 1)/(N - 1)$ , which depends on  $N$  and  $R$ . Similarly, Koning's (2000)  $CRK(R)$  measure has a *lower* bound, which is the inverse of  $C5ICB^{ub}$ . Based on these results, and assuming two rounds of matches in a balanced schedule, Manasis et al. (2011) propose a normalized concentration ratio:

$$NCR(R) = \frac{\sum_{i=1}^R P_i - \alpha R(N - 1)}{\alpha R(N - R)}$$

where  $\alpha$  is the number of points for a win. This ratio lies in the interval  $[0, 1]$ , with zero indicating perfect balance and one representing maximum domination by the top  $R$  teams.

Because the Gini coefficient,  $HHI$  and concentration ratios are share-based measures, concern about them focused on teams not being able to win games they did not play. The perception remained that this does not affect the standard deviation measures, such as  $RSD$ . Indeed, Utt and Fort (2002, p.373) cite the downward bias of conventional Gini coefficients, and the problems of adjusting for this bias in more complex competition designs (for which the assumption of balanced schedules is only an approximation), as justification for sticking with “the tried and true standard deviation of winning percentages (and their idealized values) for within-season competitive balance analysis”.

However, *all* measures of within-season CB are affected by the constraints imposed by the league's playing schedules. Both  $SD$  and  $RSD$  also have upper bounds,  $SD^{ub}$  and  $RSD^{ub}$  respectively (Groot, 2008; Owen, 2010). For the case of balanced playing schedules in which each of the  $N$  teams plays every other team  $K$  times (so each team plays  $G = K(N - 1)$  games), and with draws treated as half a win: <sup>6</sup>

$$SD^{ub} = [(N + 1)/\{12(N - 1)\}]^{0.5}$$

$$RSD^{ub} = 2[K(N + 1)/12]^{0.5}$$

$RSD^{ub}$  is a strictly increasing function of both the parameters that affect season length,  $N$  and  $K$ . It is much more sensitive to variation in these parameters than  $SD^{ub}$ , which depends on  $N$  only and has a finite limit; as  $N$  increases,  $SD^{ub}$  tends to  $(1/12)^{0.5} = 0.289$ . Although these results apply specifically to the upper bounds and do not consider other values in the distribution, they do suggest that comparing  $RSD$  values for different leagues with different values of  $N$  and  $K$  may be problematical.

Goosens (2006) advocates the use of a normalized standard deviation measure (which she labels the ‘National Measure of Seasonal Imbalance’), here denoted  $SD^*$ . This compares  $SD$  with its maximum feasible value; i.e.,  $SD^* = SD/SD^{ub}$ . Her motivation is mainly to avoid  $RSD$  taking values below unity (i.e., if  $SD < ISD$ ), which, taken at face value, imply “a competition that is more equal than when the league is perfectly balanced” (Goosens, 2006, p.87). Owen (2010) argues that a potentially more fundamental motivation for the use of a normalized standard deviation measure, such as  $SD^*$ , is that, by standardizing the measure to lie in the interval  $[0, 1]$ , it adjusts for the variation in the feasible range of values for  $RSD$  due to different numbers of games played.<sup>7</sup>

Apart from their upper and lower bounds, very little else is known about the properties of the sampling distributions of standard measures of variation, inequality and concentration in the specific context of comparing CB in different sports leagues. To provide a clearer picture, Owen and King (2013) undertake a simulation study to examine the distributional properties of  $RSD$ ,  $SD$ ,  $SD^*$  and  $CV$ , applied to points ratios, for a range of different assumptions about the distribution of teams’ strengths. Their results suggest that  $RSD$  performs as expected (with its sampling distribution centred on unity) in cases of perfect balance, i.e., equal team strengths. However, if there is imbalance in team strengths, the location of its distribution is very sensitive to changes in season length.  $RSD$  is usually advocated for comparisons involving scenarios with different numbers of teams and/or games

played, but the simulation results suggest that such comparisons of *RSD* could lead to highly misleading conclusions about differences in the underlying degrees of CB. The other standard-deviation-based measures, *SD*, *CV* and *SD\**, are subject to upward bias if the number of games played is not large, but their distributions are far less sensitive to variations in season length and so provide a more reliable basis for cross-league CB comparisons. More research into the distributional properties of different CB measures is clearly required.<sup>8</sup>

### *Unbalanced schedules*

The results discussed so far apply to end-of-season measures in the simplest case of balanced schedules, the norm in most football leagues. However, in tournament designs based on the conference system, or other forms of unbalanced scheduling, teams do not play every other team the same number of times. Consequently, the ‘strength of the schedule’ is not the same for each team in the league; some teams can end up with an easier set of games than others. In such situations, one approach would be to try to characterize the ‘overall’ CB for the competition, reflecting not only the underlying dispersion in teams’ strengths but also the effects on outcomes stemming from the design of the playing schedules. However, the literature currently contains no analysis of the feasible range of values, let alone the distributional properties, of standard CB measures for such designs. Schedules that vary across teams introduce more analytical complications than the simple balanced schedules case. Simulation methods appear to be the most promising way to proceed if the aim is to develop modified CB measures fine-tuned to the peculiarities of a league’s specific design.

If the focus of attention is on obtaining a CB measure reflecting the dispersion of actual team strengths (stripping out the effects due to the peculiarities of the scheduling), then an alternative approach is to adjust the observed values of standard CB measures to correct for the biases inherent in the end-of-season win or points ratios due to the unbalanced schedules.

These adjusted measures can then, other things equal, be more meaningfully compared with comparable measures for leagues with balanced schedules. The variation in the strength of the schedule in the Scottish Premier League (SPL), for example, is explicit: the league splits in half for the last five games of the season with the top six teams playing one more game against each of the other top six, and the same for the bottom six, providing a more balanced set of games. Lenten (2008) develops an adjustment that specifically deals with the SPL setup, using the end-of-season win ratios for all the teams to estimate win probabilities for the ‘missing’ matches required to balance the schedules. He then calculates *RSD*, concentration ratios, *HHI* and a Gini coefficient based on the adjusted win ratios. As expected, measures based on the adjusted win ratios imply that the underlying degree of CB in the SPL is considerably lower than is apparent from the raw win ratios in the ability–matching unbalanced schedules design.<sup>9</sup>

#### *Multiple prizes and competitive balance*

Another important characteristic of many football leagues is the existence of multiple prizes within a single league competition (Kringstad and Gerrard, 2007). Although winning the championship is the ultimate prize, finishing in the top  $R$  positions leads to qualification for lucrative continental club competitions, such as the UEFA Champions League (with, for example,  $R = 4$  for the EPL) or the UEFA Europa League. In addition, avoiding relegation is in itself a major prize, especially for recently promoted or smaller-market teams.

Standard measures of CB that focus on the dispersion of points ratios or championship concentration do not reflect these important features. This may not matter for assessment of how evenly teams are matched, but is likely to be important for demand analysis. Even if traditional measures of CB suggest relatively high levels of imbalance, the existence of

several competitions within a single league will, other things equal, have the effect of increasing fans' interest (Brandes and Franck, 2007).<sup>10</sup>

Manasis et al. (2012) provide the first attempt to construct a CB measure that explicitly recognizes the existence of multiple prizes in European football leagues. They construct separate normalized concentration ratio measures (similar to  $NCR(R)$ ) for the championship and teams qualifying for Europe, and for relegated teams. These are combined into a 'Special Concentration Ratio', which ranges between zero (a perfectly balanced league) and one (a completely unbalanced league). Application to the EPL reveals, in line with most other measures, a marked reduction in CB after the mid-1990s. The component ratios suggest that this may be due to less competition among the top teams rather than increasing overall domination by these teams; also, in contrast to Michie and Oughton's (2005) conclusion, the relegation sub-index suggests that promotion and relegation contribute significantly to a more balanced league.

One limitation of the measure proposed by Manasis et al., which the authors recognize, is that the component reflecting European qualification does not incorporate information on the competition from teams after the  $R$ th position considered; similarly, the relegation teams' component does not provide any information for the behaviour of the remaining teams. Hence, the closeness of competition around these key cut-off points is not fully captured. However, this is an important attempt to address the relevant issue of multiple prizes and is likely to stimulate further work.

Borooah and Mangan (2012), one of the few studies to emphasize the potential for linking the measurement of CB to the broader inequality literature, address the multiple prizes feature by exploiting the decomposability properties of generalized entropy measures of inequality. They apply  $IGE(2)$ , essentially  $CV$ , to the 2006/07 season of the EPL to examine how much of the observed inequality in the distribution of points is due to inequality

*within* predefined groups and how much is due to inequality *between* groups. They find that, splitting the end-of-season-ranked EPL teams into equal groups of five (top five, next five, etc), 84% of the imbalance in 2006/07 arises from inequality between the four groups, whereas only 16% is due to inequality within the four groups. Although CB for the league as a whole may be relatively low, within-group CB is relatively high.

Hence, in terms of the effects of CB on fan interest, the existence of multiple prizes within the EPL, combined with a high degree of competition among subsets of teams in the league for these different prizes, may more than offset the relatively low level of CB as judged by conventional measures. As Robson (2010) sums it up, somewhat more prosaically:

[I]t's been another great weekend for the Premier League. There's a wafer-thin mint between the top three, a fag paper between the next four, and the thickness of a butterfly's wing between 13th and 19th. Who's complaining?

### **Dynamic Measures of Competitive Balance**

When considering the dynamics of CB over time, one possible approach is to select static end-of-season measures and track their movement over time. For example, Lee and Fort (2012) examine the evolution of *RSD* and several of Koning's (2000) concentration measures for the English First Division/EPL since its first season in 1888/89. They apply break-point tests to identify structural breaks and stationary subsets of the data. The most recent break is identified around the mid 1990s, soon after the old English First Division broke away from the Football League with the formation of the EPL in 1992 and the redesign of the Champions League in 1994/95, after which there is increasing imbalance, as measured by *RSD*, *CRK*(3) and *CRK*(4). In contrast, other measures are specifically designed to incorporate information on outcomes from more than one season, and so have an explicit time dimension.

### *Measures of prize concentration*

One approach is to consider ‘prize concentration’ over time, using the *HHI* or the Gini coefficient applied to shares of championship wins over a selected number of successive seasons. Unlike the application of these measures to the dispersion of points in a single season, prize concentration is more closely analogous to the conventional market concentration scenario. The same team could, in principle, win the championship in every season, as in the case of a monopolist’s market share, so the *HHI* could equal unity. In a league with  $N$  teams and perfect balance in terms of championship concentration, each team would win the championship, on average, every  $N$  years. Hence, the interpretation of the upper and lower bounds carries over from market concentration analysis.

Kringstad and Gerrard (2007), for example, report decadal average *HHI* values for the five largest European leagues. In comparison with North American major leagues, championship wins in the European leagues are more highly concentrated. For the decade 1996/97 to 2006/06, Germany had the highest level of championship concentration ( $HHI = 0.52$ ), reflecting the dominance of Bayern Munich, and Spain (with four different championship winners in the decade) the lowest ( $HHI = 0.30$ ).

Some studies have compared the number of different teams appearing in the top  $R$  positions over successive comparable time periods (Buzzacchi et al., 2003; Goossens, 2006). A decrease in this count indicates increased imbalance. Buzzacchi et al. (2003) derive the expected values of the number of teams appearing in the top  $R$  positions for open leagues (with promotion and relegation) and for closed leagues, assuming equal win probabilities for the teams. Comparisons of theoretical (perfect balance) and observed frequencies show that European football’s open leagues are dynamically less balanced and deviate more from the theoretical frequencies than North American closed leagues.

### *Measures of performance persistence*

‘Performance persistence’ is concerned with the degree of persistence of *individual* teams’ points outcomes over successive seasons. From this perspective, the identities of teams occupying different positions in the final ranking are important. This calls for different measures, especially if the aim is to characterize the overall mobility of teams in a league, not just the champions or the top few teams.

The Spearman rank correlation coefficient (*SRC*), which measures the correlation between end-of-season team rankings in successive seasons, is one measure that captures team mobility in the ordinal rankings, although it does not consider the dispersion of points across teams (Kringstad and Gerrard, 2007). If the rank order is unchanged,  $SRC = 1$ , if the rank order is completely reversed,  $SRC = -1$ , whereas if the ranks are randomly allocated, *SRC* will be close to zero. Kringstad and Gerrard’s (2007) application of *SRC* to the ‘big five’ European football leagues shows no major differences compared to the North American major leagues for selected pairs of seasons and there is little evidence of any trend in performance persistence for the European leagues.

Groot (2008, Ch. 5) uses an alternative rank correlation measure, Kendall’s tau, to evaluate the mobility of teams in the league’s ranking. Tau is defined on the interval  $[-1, 1]$ . If the rank order remains unchanged over time, tau equals one, corresponding to dynamically perfect imbalance. From this mobility perspective, dynamically perfect balance corresponds to a situation where the rankings are independent from one season to the next, reflected in tau taking a value near zero. Groot (2008) calculates this statistic for different sub-periods over the life of the English First Division/EPL and finds that the extent of mobility in teams’ rankings declines considerably (tau increases) after the Second World War and especially for the period 1995 to 2006.

Mizak et al. (2007) propose a ‘churn’ index calculated as the average (over teams) of the absolute change in teams’ rankings from one season to the next. As the maximum possible value of the churn index varies with  $N$ , they divide by the maximum value to give an ‘adjusted churn’ index, which is thereby normalized to lie in the  $[0, 1]$  interval. A similarly constructed ‘fluidity’ measure is used by Flores et al. (2010) in an event study to assess the effects of the Bosman ruling on CB.

Lenten (2009) constructs what he describes as a ‘mobility gain function’, specified either as a quadratic or linear objective function defined on changes in each team’s win ratio in adjacent seasons. These measures appear to pick up aspects of CB that are not captured in within-season measures such as *RSD*, *HHI* or the Gini coefficient.<sup>11</sup>

Measures have also been proposed to link within-season win dispersion and between-season performance persistence. Eckard (1998) decomposes the variation in win ratios for  $N$  teams over  $T$  seasons into a time component (the mean of teams’ variances of annual win ratios around their individual period mean) and a cumulative component (the variance of cumulative win ratios across all teams during the period). In a similar vein, Humphreys (2002) proposes his ‘competitive balance ratio’, which compares the mean (across teams) of each team’s standard deviation of win ratios across seasons in the numerator to the mean (across a selected number of seasons) of the standard deviation of win ratios across teams in each season in the denominator. This provides a single summary value in the range  $[0, 1]$ , with an increase in CB reflected in an increase in the ratio, due either to a decrease in the average spread of within-season win ratios and/or an increase in mobility in the ranking of win ratios. However, applications to football leagues are rare because of changes in the number of teams in the league over time, changing identities of the teams due to promotion and relegation, and the need to decompose the measure into its component dimensions in order to identify the sources of any changes (Goossens, 2006; Kringstad and Gerrard, 2007).

Other approaches to assessing dynamic season-to-season CB draw on the literature on time-series econometrics and convergence in the growth literature. For example, Szymanski and Smith (2002) formulate a first-order autoregressive model of win percentages including team-specific fixed effects to proxy for innate ability. A high value of the autoregressive parameter indicates the persistence of teams' performances between seasons. Their results suggest that European football leagues exhibit more persistence in teams' win ratios and more predictability compared to North American sports leagues.

### **Measures of Uncertainty of Outcome**

UO is hypothesized to be an important (some would say *the* most important) determinant of the demand for sport. It is explicitly concerned with fans' perceptions of future outcomes and is therefore inherently an ex ante concept. As with CB, different time dimensions are relevant in measuring UO. However, with UO, there is also a more pronounced shorter run focus on match uncertainty, which is concerned with the predictability of individual matches, as well as medium-run within-season uncertainty and long-run championship uncertainty (Cairns et al., 1986; Borland and Macdonald, 2003; Szymanski, 2003). As with CB, there are no clear 'best ways' to measure the different unobservable ex ante dimensions of UO. As a result, lack of consistent empirical evidence identifying significant effects of UO in demand analysis may be partly due to the difficulties in measuring the different dimensions of UO, exacerbated by the likely relevance of interaction effects between these different dimensions (Sloane, 2006).

Match uncertainty is the most frequently examined dimension of OU. Its measurement draws on two main sources of information: statistics on the two teams' relative performances prior to a match or betting odds. The former approach may consider teams' league positions (Hart et al., 1975; Dobson and Goddard, 1992; Baimbridge et al., 1996; Falter and Pérignon, 2000; García and Rodríguez, 2002; Benz et al., 2009; Madalozzo and Villar, 2009) or points

totals (Wilson and Sim, 1995). The most important drawback with past relative performance measures is that they are “entirely backward-looking” (Downward and Dawson, 2000, p.134); past results are relevant but other information also matters. Also, recent results do not account for the difficulty of the teams’ playing schedules up to that point in the season or for home advantage (Forrest and Simmons, 2002; Buraimo et al., 2008) and may not adequately capture current form (Sandy et al., 2004).

Betting odds accord more closely with the ex ante nature of OU. They can incorporate a wider range of relevant information than past outcome statistics, for example, on suspensions, injuries, player transfers and managerial changes. Odds-based measures of match uncertainty are generally expressed in the form of the probability of a home-team win (Peel and Thomas, 1988; Czarnitzki and Stadtmann, 2002; Benz et al., 2009) or the ratio of the probability of a home win to the probability of an away win (Forrest and Simmons, 2002). However, the major problem with using betting odds, if available, is concern about biases in setting the odds (Forrest and Simmons, 2002; Forrest et al., 2005; Dawson and Downward, 2005; Buraimo et al., 2008) and a timing mismatch between when odds are set and fans make decisions about match attendance (Szymanski, 2006).

Seasonal uncertainty is the degree of within-season uncertainty associated with teams winning the championship, qualifying for international club competitions and avoiding relegation. From an overall league perspective, fans are expected to prefer higher levels of uncertainty with many teams in contention and the final outcomes in the balance until the end of the season. In many sports (and some football leagues, such as the A-league in Australia, Major League Soccer in the US, and the lower leagues in English football), seasonal uncertainty is enhanced by the introduction of playoffs. However, in the top European and South American football leagues, seasonal uncertainty is enhanced by multiple prizes, as

discussed previously. This may help explain why conventional CB and UO measures seem to have less of an effect on attendance than might be expected (Brandes and Franck, 2007).

Measures of seasonal uncertainty are usually based on the number of games a team needs to win in order to win the championship (or qualify for other competitions or avoid relegation) (Jennett, 1984), the number of games (wins) or points behind the leading team (García and Rodríguez, 2002; Benz et al., 2009), or the significance of the match for the championship or relegation (Dobson and Goddard, 1992; Madalozzo and Villar, 2009). Normally, these are used to help explain attendance at individual matches; however average league-level seasonal uncertainty measures, based on aggregating team-specific information on games behind the leader, can be constructed for the analysis of average attendances (Lee, 2009).

Although widely used in attendance studies, these variables are relatively crude. Any measure based on the number of games or points a team needs in order to win the championship or avoid relegation requires information that is not available at the time spectators decide whether to attend a match (Dawson and Downward, 2005). Counting the number of ‘games behind’ provides a guide to the *feasibility* of a team remaining in contention, but relies entirely on the current league table, does not consider the difficulty of remaining fixtures either for that team or for others in contention, and usually assesses how teams rank relative only to the current leader.

An alternative that avoids these problems is to use forecasting or simulation methods to derive ex ante measures of different aspects of UO (such as the probabilities of each team winning a particular match, the overall championship, being relegated, etc). Forecast or simulation-based probabilities can evolve as the season unfolds and information on actual outcomes becomes available. Such methods are computationally more demanding than

counting games behind and, so far, have been used in relatively few empirical studies (Dawson et al., 2007; Bojke, 2008; King et al., 2012).

Consecutive-season or long-run championship uncertainty refers to the absence of long-run domination by one team or a small number of teams. There are relatively few empirical demand studies that incorporate this dimension of UO; those that do tend to use historic series for some of the ex post CB measures discussed above (Borland and Macdonald, 2003).

## **Concluding Comments**

Although there is general agreement that CB and UO are multi-dimensional, analysts are clearly faced with a disconcertingly large set of possible indicators. As Zimbalist (2003, p.162) has observed, “[t]he competitive balance literature should be more than an endless cataloguing of these metrics”. In making choices about which measures to use in practice, this review suggests some broad principles.

Firstly, measures that explicitly take into account the implications of specific design characteristics of football leagues, such as playing schedules, drawn matches, home advantage and multiple prizes are likely to avoid well-documented biases and be more informative than standard off-the-shelf statistics.

Secondly, measures for which there is evidence of well-behaved distributional properties for a range of feasibly relevant scenarios are preferable. Surprisingly little work has been done in this area, but there are concerns about the reliability of the popular ratio of standard deviations of win or points ratios, especially for comparisons involving different season lengths.

Thirdly, the purpose of the analysis can make a difference. If the aim is to evaluate the underlying degree of inequality in teams’ strengths, the choice of CB measures may be

different from an application in which the aim is to describe the degree of CB of the overall league design, including all other influences, such as scheduling and the structure of prizes.

Fourthly, for the selection of UO measures, relevance for fans and demonstrable sensitivity of demand to changes in the measures are crucial. Arguably, this also applies to CB measurement (Zimbalist, 2002, 2003; Fort and Maxcy, 2003). As Zimbalist (2003, p.162) observes, “[i]f fans were indifferent to the uncertainty of outcome and fairness of competition, competitive balance would be no more relevant ... than the choice of typescript on the balls used in competition”.

## References

- Baimbridge M., Cameron S. and Dawson P. (1996), Satellite Television and the Demand for Football: A Whole New Ball Game, *Scottish Journal of Political Economy*, 43, 317-333.
- Bajo O. and Salas R. (2002), Inequality Foundations of Concentration Measures: An Application to the Hannah-Kay Indices, *Spanish Economic Review*, 4, 311-316.
- BBC Sport, Scudamore defends 'boring' League, 7 May 2008,  
<http://news.bbc.co.uk/sport2/hi/football/7388360.stm> (retrieved 21 December 2012).
- Ben-Naim E., Vazquez F. and Redner S. (2006), Parity and Predictability of Competitions, *Journal of Quantitative Analysis in Sports*, 2(4), Article 1,  
<http://www.bepress.com/jqas/vol2/iss4/1>
- Benz M.-A., Brandes L. and Franck E. (2009), Do Soccer Associations Really Spend on a Good Thing? Empirical Evidence on Heterogeneity in the Consumer Response to Match Uncertainty of Outcome, *Contemporary Economic Policy*, 27, 216-235.
- Bojke C. (2008), The Impact of Post-Season Play-Off Systems on the Attendance at Regular Season Games, in Albert J. and Koning R.H. (Eds), *Statistical Thinking in Sports*, Chapman & Hall/CRC Press, Boca Raton, FL, pp. 179-202.
- Borland J. and Macdonald R. (2003), Demand for Sport. *Oxford Review of Economic Policy*, 19, 478-502.
- Borooh V.K. and Mangan J. (2012), Measuring Competitive Balance in Sports Using Generalized Entropy with an Application to English Premier League Football, *Applied Economics*, 44, 1093-1102.
- Brandes L. and Franck E. (2007), Who Made Who? An Empirical Analysis of Competitive Balance in European Soccer Leagues, *Eastern Economic Journal*, 33, 379-403.
- Brizzi M. (2002), A Class of Indices of Equality of a Sport Championship: Definition, Properties and Inference, in Mrvar A. and Ferligoj A. (Eds), *Developments in Statistics*, Faculty of Social Sciences (FDV), University of Ljubljana, Ljubljana, Slovenia, pp. 175-195.

- Buraimo B., Forrest D. and Simmons R. (2008), Outcome Uncertainty Measures: How Closely Do They Predict a Close Game?, in Albert J. and Koning R.H. (Eds), *Statistical Thinking in Sports*, Chapman & Hall/CRC Press, Boca Raton, FL, pp. 167-78.
- Buzzacchi L., Szymanski S. and Valletti T.M. (2003), Equality of Opportunity and Equality of Outcome: Open Leagues, Closed Leagues and Competitive Balance, *Journal of Industry, Competition and Trade*, 3, 167-186.
- Cain L.P. and Haddock D.D. (2006), Measuring Parity: Tying Into the Idealized Standard Deviation, *Journal of Sports Economics*, 5, 169-185.
- Cairns J.A., Jennett N. and Sloane P.J. (1986), The Economics of Professional Team Sports: A Survey of Theory and Evidence, *Journal of Economic Studies*, 13, 3-80.
- Cowell F.A. (2011), *Measuring Inequality* (Third Edition), Oxford University Press, Oxford.
- Czarnitzki D. and Stadtmann G. (2002), Uncertainty of Outcome Versus Reputation: Empirical Evidence for the First German Football Division, *Empirical Economics*, 27, 101-112.
- Davies S. (1979), Choosing Between Concentration Indices: The Iso-Concentration Curve, *Economica*, 46, 67-75.
- Dawson A. and Downward P. (2005), Measuring Short-Run Uncertainty of Outcome in Sporting Leagues: A Comment, *Journal of Sports Economics*, 6, 303-313.
- Dawson P., Dobson S., Goddard J. and Wilson, J. (2007), Are Football Referees Really Biased and Inconsistent?: Evidence on the Incidence of Disciplinary Sanction in the English Premier League, *Journal of the Royal Statistical Society: Series A*, 170. 231-250.
- Depken C.A., II (1999), Free-agency and the Competitiveness of Major League Baseball, *Review of Industrial Organization*, 14, 205-217.
- Dobson S.M. and Goddard J.A. (1992), The Demand for Standing and Seated Viewing Accommodation in the English Football League, *Applied Economics*, 24, 1155-1163.
- Dobson S. and Goddard J. (2011), *The Economics of Football*, 2<sup>nd</sup> Edition, Cambridge University Press, Cambridge.

- Downward P. and Dawson A. (2000), *The Economics of Professional Team Sports*, Routledge, London.
- Eckard E.W. (1998), The NCAA Cartel and Competitive Balance in College Football, *Review of Industrial Organization*, 13, 347-369.
- Falter J.-M. and Pérignon C. (2000), Demand for Football and Intramatch Winning Probability: An Essay on the Glorious Uncertainty of Sports, *Applied Economics*, 32, 1757-1765.
- Flores, R., Forrest, D. and Tena, J.D. (2010), Impact on Competitive Balance from Allowing Foreign Players in a Sports League: Evidence from European Soccer, *Kyklos*, 63, 546-557.
- 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 (The Statistician)*, 51, 229-241.
- Forrest D., Simmons R. and Buraimo B. (2005), Outcome Uncertainty and the Couch Potato Audience, *Scottish Journal of Political Economy*, 52, 641-661.
- Fort R. (2007), Comments on “Measuring Parity”, *Journal of Sports Economics*, 8, 642-651.
- Fort R. and Maxcy J. (2003), Comment on “Competitive Balance in Sports Leagues: An Introduction”, *Journal of Sports Economics*, 4, 154-160.
- Fort R. and Quirk J. (1995), Cross-Subsidization, Incentives, and Outcomes in Professional Team Sports Leagues, *Journal of Economic Literature*, 33, 1265-1299.
- Fort R. and Quirk J. (1997), Introducing a Competitive Economic Environment Into Professional Sports, in Hendricks W. (Ed.), *Advances in the Economics of Sport*, Volume 2, JAI Press, Greenwich, CT, pp. 3-26.
- Fort R. and Quirk J. (2010), Optimal Competitive Balance in Single-Game Ticket Sports Leagues, *Journal of Sports Economics*, 11, 587-601.
- Fort R. and Quirk J. (2011), Optimal Competitive Balance in a Season Ticket League, *Economic Inquiry*, 49, 464-473.

- García J. and Rodríguez P. (2002), The Determinants of Football Match Attendance Revisited: Empirical Evidence from the Spanish Football League, *Journal of Sports Economics*, 3, 18-38.
- Goossens K. (2006), Competitive Balance in European Football: Comparison by Adapting Measures: National Measure of Seasonal Imbalance and Top 3, *Rivista di Diritto ed Economia dello Sport*, 2, 77-122.
- Groot L. (2008), *Economics, Uncertainty and European Football: Trends in Competitive Balance*, Edward Elgar, Cheltenham.
- Hannah L. and Kay J.A. (1977), *Concentration in Modern Industry: Theory, Measurement and the U.K. Experience*, Macmillan, London.
- Hart P.E. (1975), Moment Distributions in Economics: An Exposition, *Journal of the Royal Statistical Society, Series A*, 138, 423-434.
- Hart R.A., Hutton J. and Sharot T. (1975), A Statistical Analysis of Association Football Attendances, *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 24, 17-27.
- Horowitz I. (1997), The Increasing Competitive Balance in Major League Baseball, *Review of Industrial Organization*, 12, 373-387.
- Humphreys B.R. (2002), Alternative Measures of Competitive Balance in Sports Leagues, *Journal of Sports Economics*, 3, 133-148.
- Jennett N. (1984), Attendances, Uncertainty of Outcome and Policy in Scottish League Football, *Scottish Journal of Political Economy*, 31, 176-98.
- King N., Owen P.D. and Audas R. (2012), Playoff Uncertainty, Match Uncertainty and Attendance at Australian National Rugby League Matches, *Economic Record*, 88, 262-277.
- Koning R.H. (2000), Balance in Competition in Dutch Soccer, *The Statistician*, 49, 419-431.
- Kringstad M. and Gerrard, B (2007), Beyond Competitive Balance, in Parent M.M. and Slack T. (Eds), *International Perspectives on the Management of Sport*, Butterworth-Heinemann, Burlington, MA, pp. 149-172.

- Lee Y.H. (2009), The Impact of Postseason Restructuring on the Competitive Balance and Fan Demand in Major League Baseball, *Journal of Sports Economics*, 10, 219-235.
- Lee Y.H. and Fort R. (2012), Competitive Balance: Time Series Lessons from the English Premier League, *Scottish Journal of Political Economy*, 59, 266-282.
- Lenten L.J.A. (2008), Unbalanced Schedules and the Estimation of Competitive Balance in the Scottish Premier League, *Scottish Journal of Political Economy*, 55, 488-508.
- Lenten L.J.A. (2009), Towards a New Dynamic Measure of Competitive Balance: A Study Applied to Australia's Two Major Professional 'Football' Leagues, *Economic Analysis and Policy*, 39, 407-428.
- Lenten, L.J.A. (2011), The Extent to Which Unbalanced Schedules Cause Distortions in Sports League Tables, *Economic Modelling*, 28, 451-458.
- Lenten, L.J.A. (2013), Measurement of Competitive Balance in Conference and Divisional Tournament Design, *Journal of Sports Economics*, doi:10.1177/1527002512471538
- Madalozzo R. and Villar R.B. (2009), Brazilian Football: What Brings Fans to the Game?, *Journal of Sports Economics*, 10, 639-650.
- Manasis V., Avgerinou V., Ntzoufras I. and Reade J.J. (2011), Measurement of Competitive Balance in Professional Team Sports Using the Normalized Concentration Ratio, *Economics Bulletin*, 31, 2529-2540.
- Manasis V., Avgerinou V., Ntzoufras I. and Reade J.J. (2012), Quantification of Competitive Balance in European Football: Development of Specially Designed Indices, *IMA Journal of Management Mathematics*, doi:10.1093/imaman/dps014
- Michie J. and Oughton C. (2004), *Competitive Balance in Football: Trends and Effects*, The Sports Nexus.
- Michie J. and Oughton C. (2005), *Competitive Balance in Football: An Update*, The Sports Nexus.

- Mizak D., Neral J. and Stair A. (2007), The Adjusted Churn: An Index of Competitive Balance for Sports Leagues Based on Changes in Team Standings Over Time, *Economics Bulletin*, 26(3), 1-7. <http://economicsbulletin.vanderbilt.edu/2007/volume26/EB-07Z00005A.pdf>
- Noll R.G. (1988), Professional Basketball, Studies in Industrial Economics Paper No. 144. Stanford University, Stanford, CA.
- Owen P.D. (2010). Limitations of the Relative Standard Deviation of Win Percentages for Measuring Competitive Balance in Sports Leagues, *Economics Letters*, 109, 38-41.
- Owen P.D. (2012), Measuring Parity in Sports Leagues With Draws: Further Comments, *Journal of Sports Economics*, 13, 85-95.
- Owen P.D. and King N. (2013), Simulating Distributions of Competitive Balance Measures in Sports Leagues: The Effects of Variation in Season Length, University of Otago, Discussion Papers in Economics No. 1309.
- Owen P.D. et al. (2007), Measuring Competitive Balance in Professional Sports Using the Herfindahl-Hirschman Index. *Review of Industrial Organization*, 31, 289-302.
- Peel D. and Thomas D. (1988), Outcome Uncertainty and the Demand for Football: An Analysis of Match Attendances in the English Football League, *Scottish Journal of Political Economy*, 35, 242-249.
- Premier League (2011), The World's Most Watched League, PremierLeague.com, 16 November 2011, <http://www.premierleague.com/en-gb/about/the-worlds-most-watched-league.html> (retrieved 20 December 2012).
- Quirk J. and Fort R.D. (1992), *Pay Dirt: The Business of Professional Team Sports*, Princeton University Press, Princeton, NJ.
- Robson D. (2010), Bridge over the River Cry?, BBC Sport, 1 March 2010, [http://www.bbc.co.uk/blogs/roborobson/2010/03/bridge\\_over\\_the\\_river\\_cry.html](http://www.bbc.co.uk/blogs/roborobson/2010/03/bridge_over_the_river_cry.html) (retrieved 21 December 2012).

- Rottenberg S. (1956), The Baseball Players' Labor Market, *Journal of Political Economy*, 64, 242-258.
- Sandy R., Sloane P.J. and Rosentraub M.S. (2004), *The Economics of Sport: An International Perspective*, Palgrave Macmillan, Basingstoke.
- Scully G.W. (1989), *The Business of Major League Baseball*, University of Chicago Press, Chicago, IL.
- Sloane P.J. (1971), The Economics of Professional Football: The Football Club as a Utility Maximiser, *Scottish Journal of Political Economy*, 18, 121-146.
- Sloane P.J. (2006), Rottenberg and the Economics of Sport After 50 Years: An Evaluation, in Rodríguez P., Kácsenne S. and García J. (Eds), *Sports Economics after Fifty Years: Essays in Honour of Simon Rottenberg*, University of Oviedo Press, Oviedo.
- Szymanski S. (2003), The Economic Design of Sporting Contests, *Journal of Economic Literature*, 41, 1137-1187.
- Szymanski S. (2006), Uncertainty of Outcome, Competitive Balance and the Theory of Team Sports, in Andreff W. and Szymanski S. (Eds), *Handbook on the Economics of Sport*, Edward Elgar, Cheltenham, pp. 597-600.
- Szymanski S. and Smith R. (2002), Equality of Opportunity and Equality of Outcome: Static and Dynamic Competitive Balance in European and North American Sports Leagues, in Barros C.P., Ibrahimo M. and Szymanski S. (Eds), *Transatlantic Sport: the Comparative Economics of North American and European Sports*, Edward Elgar, Cheltenham, pp. 109-124.
- Trandel G.A. and Maxcy J.G. (2011), Adjusting Winning-Percentage Standard Deviations and a Measure of Competitive Balance for Home Advantage, *Journal of Quantitative Analysis in Sports*, 7(1), Article 1, <http://www.bepress.com/jqas/vol7/iss1/1>
- Utt J. and Fort R. (2002), Pitfalls to Measuring Competitive Balance with Gini Coefficients, *Journal of Sports Economics*, 3, 367-373.

- Vrooman J. (1996), The Baseball Players' Labor Market Reconsidered, *Southern Economic Journal*, 63, 339–360.
- Wilson J. (2011), The Question: Just How Competitive is the Premier League?, *The Guardian*, 19 October 2011, <http://www.guardian.co.uk/sport/blog/2011/oct/19/the-question-how-competitive-premier-league> (retrieved 20 December 2012).
- Wilson P. and Sim B. (1995), The Demand for Semi-Pro League Football in Malaysia 1989-91: A Panel Data Approach, *Applied Economics*, 27, 131-138.
- Zimbalist A.S. (2002), Competitive Balance in Sports Leagues: An Introduction, *Journal of Sports Economics*, 3, 111-121.
- Zimbalist A. (2003), Competitive Balance Conundrums: Response to Fort and Maxcy's Comment, *Journal of Sports Economics*, 4, 161-163.

---

## Notes

<sup>1</sup> The focus of this chapter is on measurement. Consideration of theoretical and empirical analysis of the effects of CB and UO on the demand for football and the welfare of football's fans, players and club owners, although central to the economic analysis of football leagues, is outside the scope of this review.

<sup>2</sup> For scenarios in which draws are worth half a win (or for sports in which draws are not possible),  $\bar{p}$  always equals 0.5, so the mean points ratio does not need to be estimated and  $N$  is often used as the divisor, instead of  $(N - 1)$ , in calculating the standard deviation. In situations in which  $\bar{p} \neq 0$ , the coefficient of variation of points proportions,  $CV = SD/\bar{p}$  can be considered to allow for variation in  $\bar{p}$  across seasons.

<sup>3</sup> These observations are based on comparing different *ISD* and corresponding *RSD* values for identical sets of win, draw and loss results. They do not imply that changing the points assignment from (2,1,0) to (3,1,0) will have minimal effects on CB, as, in practice, this would affect teams' playing strategies and hence, among other things, the ratios of wins and draws.

<sup>4</sup> Dobson and Goddard (2011, Table 3.1) report season-by-season percentages of home wins, draws and away wins for the English league from 1970. These range from 43% to 52.4% for home wins and 24.3% to 30.9% for draws. Using the (2,1,0) points assignment scheme, home teams' points percentage varies from 56.6% to 65.9%. Dobson and Goddard (2011, Table 3.2) also document a decline in the extent of home advantage, especially in the lower tiers of the English league.

<sup>5</sup> Despite the plethora of measures of within-season CB reviewed, the coverage is not exhaustive and concentrates on those most widely used. Other measures worthy of note include Horowitz's (1997) relative entropy measure, Groot's (2008) surprise index, Ben-Naim et al.'s (2006) 'frequency of upsets' index and Koning's (2000) approach based on standard errors from an ordered probit estimation.

<sup>6</sup> These expressions are based on *SD* calculated with  $N$  as the divisor. Using  $(N - 1)$ ,  $SD^{ub} = [N(N + 1)/\{12(N - 1)^2\}]^{0.5}$  with corresponding amendment to  $RSD^{ub}$ .

<sup>7</sup> The normalized variant of *RSD*, defined as  $RSD^* = RSD/RSD^{ub}$ , is identical to  $SD/SD^{ub} = SD^*$  because the *ISD* terms in the definition of the numerator and denominator of the former cancel.

---

<sup>8</sup> Another study that examines the distributional properties of CB measures is Brizzi (2002). He examines normalized measures based on the standard deviation, the Gini coefficient, the mean absolute deviation and the mean letter spread, all applied to points totals, but he considers simulation results for the case of exactly equally matched teams only.

<sup>9</sup> Lenten (2011, 2013) extends the adjustment method to allow for a richer range of unbalanced schedules (with applications to the Australian Football League and the National Football League).

<sup>10</sup> This observation has been emphasized by, for example, Richard Scudamore, Chief Executive of the EPL: “There are a lot of different tussles that go on in the Premier League depending on whether you're at the top, in the middle or at the bottom that make it interesting” (BBC Sport, 2008).

<sup>11</sup> Kringstad and Gerrard (2007) suggest other mobility measures based on the changing composition of membership of strategic groups or changes in absolute win ratios.