ISSN: 1756-8811

Birkbeck Sport Business Centre

Research Paper Series

A review of measures of competitive balance in the 'analysis of competitive balance' literature

Richard Evans

Birkbeck, University of London

Volume 7, Number 2, May 2014

Abstract[1]

This paper presents a review of measures of the competitive balance of sports leagues proposed in the line of literature which Fort and Maxcy (2003, p. 155) call "… the analysis of competitive balance (ACB) literature itself". They describe this as the literature which "… focuses on what has happened to competitive balance over time or as a result of changes in the business practices of pro sports." A number of books and articles have provided surveys of these measures of competitive balance. This review consolidates that work and augments it by adding details to facilitate the appropriate application and interpretation of the measures. This paper also considers the applicability of the measures of competitive balance in the literature to the professional leagues of association football.

1. This paper was first published in May 2014 as part of the Birkbeck Sports Business Centre Research Paper series. This revised version, published in March 2015, incorporates comments received on the original paper.

Introduction

Fort and Maxcy (2003, p. 155) state that the empirical literature on competitive balance is easily characterised along two distinct lines. One they call "the analysis of competitive balance (ACB) literature itself". They describe this as the literature which "… focuses on what has happened to competitive balance over time or as a result of changes in the business practices of pro sports." This paper presents a review of the measures of competitive balance of sports leagues proposed in the ACB line of literature.

The other line of literature on competitive balance, following Cairns, Jennett and Sloane (1986), addresses the effect of competitive balance on fans and tests the uncertainty of outcome hypothesis. This line of literature, including assessments of competitive balance at points during a competition will be covered in a separate review of the literature on the demand for football, as will be the issue of the competitive balance of matches since, in each case, the focus is on the effect competitive balance has on demand as a function of the uncertainty of outcome rather than on the competitive balance of the sports league 'per se'.

Zimbalist (2002, p. 112) comments that "[t]here are almost as many ways to measure competitive balance as there are to quantify the money supply." This review distinguishes between two aspects of competitive balance which the ACB literature aims to measure:

a) The extent of the closeness between teams in a league in a season (i.e. level of concentration)
b) The extent to which the same teams persist in winning over a number of seasons (i.e. level of dominance)

The essential difference is whether the identity of the team matters to the measure. It does not matter for measures of concentration but it does matter for measures of dominance.

This literature also provides measures that combine these aspects to examine the entire distribution of performance of teams in a league and how they evolve over time. These have the advantage of providing a single statistic as a measure for both aspects. However, they do not fully replace the independent measures of concentration and dominance. For example, from a policy perspective it may be important to be able to differentiate between, and assess, the two aspects of competitive balance separately.

This review therefore categorises the measures in the literature according to whether they are:

a) Measures of concentration
b) Measures of dominance
c) Measures combining concentration and dominance

The literature sometimes uses the terms 'static' and 'dynamic' to classify measures. In some cases the terms are used in a corresponding way to the terms

‘concentration’ and ‘dominance’ (for example, Szymanski, 2003). But the literature also has uses of the term ‘static’ for measures that relate to competitive balance within a particular season (for example, Koning, 2000, p. 426). To avoid this confusion, this review has not adopted the terms ‘static’ and ‘dynamic’.

Following from papers for the symposium ‘Competitive balance in sports leagues’ published in the Journal of Sports Economics (2002, Vol. 2, No. 2) a number of books and articles have provided recent surveys of these measures of competitive balance. See, for example, Dobson and Goddard (2011, Chapter 3), Downward, Dawson and Dejonghe (2010, Chapter 8), Leeds and von Allmen (2008, Chapter 5) and Goosens (2006). This review consolidates that work and augments it by including additional measures from the literature and adding some additional details to facilitate the application and interpretation of the measures to form a consolidate body of reference.

In cases where measures were not presented with a mathematical formulation in the original article, but I considered that this would add clarification to the derivation of the measure, I have expressed measures in mathematical form. For some measures it is useful to have reference values for the upper and lower limits and I have also calculated and added those as appropriate to this review (if not included in the literature). To provide a level of consistency to this review, and to facilitate comparison between measures proposed in the literature I have, on occasion, changed the mathematical notation from the original articles and introduced notation for clarification.

Some measures that have been applied to Major US sports leagues in the literature are not applicable to leagues such as soccer leagues in Europe (and elsewhere). As my thesis specifically relates to the English Premier League for football, I have indicated the applicable measures and also noted related evidence presented together with reference to the source literature.

Much of the literature proposing measures of competitive balance has applied the measures to the major sports leagues in North America. Whilst the measures are the subject of this review, the findings relating to these leagues are not and consequently they are not included in this review. Instead, references to the original source literature are provided for these findings. The measures reviewed are as follows:

A) Measures of concentration

1. Range	
2. Standard deviation	(i) Standard deviation of points or share of maximum possible wins (σ_L) (ii) Ratio of observed standard deviation to 'idealised' standard deviation (σ_R) a) Binomial-based ISD b) Possible draw outcome ISD c) 'Home advantage' ISD (iii) Normalised standard deviation of win percentages (σ_L^*) (iv) Standard deviation of team strength
3. Coefficient of variation	
4. "Excess tail frequencies" for win percentages	
5. Relative entropy $(\boldsymbol{R})$	
6. Herfindahl-Hirschman Index (HHI) – for concentration in a league season	(i) Share of all wins (games or points) in a season (HHI^C) (ii) Deviated Herfindahl-Hirschman Index $(dHHI)$ (iii) Normalised Herfindahl-Hirschman Index $(nHHI)$
7. Concentration ratios	(i) 'Attainable' concentration ratio (ii) Five club concentration ratio (C5 Ratio)
8. Index of dissimilarity $(\boldsymbol{ID})$	
9. Lorenz curve and Gini coefficient – for teams within season $(\boldsymbol{G_t^C})$	
10. Surprise index for leagues $(\boldsymbol{S_L})$	

B) Measures of dominance

1. Descriptive statistics (counts, percentages, frequency, averages etc.)	(i) Number (percentage etc.) of league titles per team (ii) Consecutive title wins (iii) 'Lifetime' achievement of teams (A) (iv) Number of different title winners (v) Number of 'top' teams (vi) Identity of 'top' teams (vii) Frequency of failure to win a league title
2. Time series association	(i) Correlation coefficient (r) (ii) Adjusted churn (AC_t) (iii) Autoreggressive win percentage (iv) 'Top 4' recurrence
3. Herfindahl-Hirschman Index (HHI) – for teams over a number of seasons $(\boldsymbol{HHI^D})$	(i) Titles won (ii) Number of top (or bottom) positions (iii) 'Virtual' league appearances
4. Lorenz curve and Gini coefficient – for leagues over a number of seasons	
5. Surprise index for teams $(\boldsymbol{D_i})$	

C) Measures combining concentration and dominance

1. Distribution of 'lifetime' win percentages in a league $(\boldsymbol{L})$	
2. ANOVA-type measures	(i) Variance decomposition $(\%\, Time)$ (ii) Competitive Balance Ratio (CBR)
3. Mobility gain function $(\boldsymbol{MGF_t})$	
4. Markov models	(i) Transitional probability tests (ii) Gini-type measure $\left(G(T^*)\right)$

Measures

A) Measures of concentration

These measures focus on the aspect of competitive balance related to the closeness of a league competition rather than to the relative performance of particular teams or groups of teams. Ten types of measure that have been proposed in the literature are presented below.

1. Range

This is a simple measure of the spread in a league. For a sports league, it is generally calculated as the difference in the win percentage of the best and worst teams in the league. The measure has limiting values of one, in the case where the best team has a perfect record and the worst team loses every game and a value of zero when all teams have the same record (indicating perfect competitive balance). See, for example, Noll (1991, Figure 1.1, p.41) for NBA, 1946-1989 and Quirk and Fort (1992, Table 7.1, p.247) for all the US sports major leagues, AL, NL, NBA, NFL and NHL, with average values for each of the relevant decades from 1901 to 1990.

This measure is applicable to 'open' leagues and could be applied more generally to other defined pairs of points in a league. It could also be based on other performance data, such as points scored. In this case, the limiting values depend on the points allocation system. This measure takes no account of concentration between (or beyond) the observed values.

2. Standard deviation

The standard deviation is a statistical measure of dispersion related to the mean as a measure of central tendency. Applied to a sports league, it provides a measure of the concentration of the teams in the league for a competition period. In general, the measure (σ_L) is calculated as follows:

$$\sigma_L = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i^* - \frac{\sum_{i=1}^{N} X_i^*}{N}\right)^2}{N}}$$

Where:
$X_i^* =$ Selected variable
$N =$ Number of teams (i) in the league

Note that if the divisor $(N-1)$ is used (instead of N) it makes the statistic σ_L an unbiased estimator of the standard deviation of the population of outcomes. The definition of the variable (X_i^*) makes the measure specific.

Four applications of this statistical measure as a measure of competitive balance of leagues are presented below. The first covers the direct application of the basic statistic (as defined above) in the literature with the case of the share of maximum possible wins or absolute points as the selected variable. The second, the ratio to an 'idealised' standard deviation, addresses the issue of the basic statistic as a

comparative measure of competitive balance if there are differences between the numbers of games played between the teams in the comparative leagues. Modified versions of this approach, which allow for games with possible draw outcomes and the different scoring systems of leagues and other influences such as differences in home advantage, are also presented. The third, a normalised measure, addresses the issue of the scale effects of the league on the basic statistic as a comparative measure of competitive balance. The fourth, with a derived variable to represent the difference in strength between teams in a league, allows for differences in the scoring system of the leagues and is also invariant to changes in home advantage.

(i) *Standard deviation of share of maximum possible wins or absolute points*

The literature includes the following variables:

a) Win percentage

In this case the selected variable is derived as: $\frac{\text{Wins achieved by team } i}{\text{Maximum possible wins for team } i}$

This produces a share based measure known as the "standard deviation of win percentages" (σ_L). Note that although this terminology is used in the literature, and accordingly it is retained in this review, it is incorrect as a mathematical description of the measure as generally applied. The measure generally used in the literature is actually the standard deviation of win '*proportions*'.

This measure has a lower limit value of zero which corresponds to a perfectly balanced league. The maximum value, which occurs if the top team wins all of their games, the second team wins all of their other games etc., depends on the number of teams in the league. For example, a league with 20 teams has a maximum value of approximately 0.3.

It is used extensively in the literature related to the competitive balance of Major League sports in the US. See, for example, Scully (1989) for the US baseball National League, 1876-1987 (ibid, Figure 4.4, p. 90) and American League, 1901-1987 (ibid, Figure 4.5, p. 90) and Quirk and Fort (1992, Table 7.1, p.247) for average standard deviations of win percentages for all the major US sports leagues, AL, NL, NBA, NFL and NHL, for each of the relevant decades from 1901 to 1990.

Scully (1989) also calculates the ratio of these values for the American League to the National League to produce a measure of the relative competitive balance of the two MLB leagues in each season from 1947 to 1987 (see Scully, 1989, Figure 4.8, p.92).

Humphreys (2002) uses this measure for a comparison of measures of competitive balance with the Herfindahl-Hirschman Index measure and the Competitive Balance Ratio (CBR). He calculates values for the American League and National League of MLB in decades from the 1900s to the 1990s (see Humphreys, 2002, Table 2, p. 139). His data excludes teams that did not play in the league in every year of the decade and consequently the values for the standard deviation of the win percentages of teams that he calculates differ from those reported by Quirk and Fort (1992).

b) Absolute points

In this case the selected variable is the absolute number of points awarded to each team in the league. This points based measure produces a more natural measure of competitive balance for sports that use point systems (rather than win shares) to determine league standing.

The maximum value, which occurs if the top team wins all of their games, the second team wins all of their other games etc., depends on the number of teams in the league and the number of games played by each team in the competition period. For example, a league with 20 teams in which every team plays every other team twice in the competition period, has a maximum value of approximately 3.8.

It is used, for example, in the literature related to European football leagues. Szymanski and Kuypers (1999) calculate this measure for each of the professional football leagues in England from 1946 to 1995. They then calculated the average of these statistics for each league for consecutive ten year periods (see Szymanski and Kuypers, 1999, Figure 7.2, p.261). They find that "[o]ver time there has been an increasing level of standard deviation, suggesting a decrease in the competitive balance. In each of the divisions the standard deviation in the most recent decade [1986-95] is bigger than that in the decade immediately after the [Second World] war [1946-55], and this trend is particularly noticeable in the top division." (ibid p. 261)

Szymanski and Kuypers (1999, Figure 7.3, p.262) shows average values for this measure for the top professional football leagues in Italy, Spain and the Netherlands for the same decades. They find that Italy and Spain have lower average standard deviations than in England, suggesting a higher competitive balance in those leagues.Koning (2000) adopts this variable and applies it to the top league of football in the Netherlands (since the current competition was introduced in the 1955/56 season).

There are, however, two important limitations to this use of this basic statistic as a comparative measure of the competitive balance of sports leagues. The first limitation relates to the above data variables used with the statistic. Both 'win percentages' and 'absolute points' for teams represent performance over a number of games (typically a league season). The standard deviation measure itself is independent of the number of games played. However, a league with fewer games is statistically more likely to have a higher standard deviation (due to more 'random noise' in the final outcomes) than a league with more games. Hence, the measure for the league does not reflect the level of certainty of the underlying performance data.

The second limitation relates to the statistic in this context. A fundamental issue with this statistic as comparative measure of competitive balance between sports leagues is that, since the upper bound depends on the number of teams in the league, it also captures the scale effect of the league (which is unrelated to the competitive balance of the league). An increase (decrease) in the number of teams in the league reduces (increases) the maximum possible value *even if there is no change in the competitive balance in the league*.

These limitations have led to the following two adapted measures of competitive balance of sports leagues based on the basic standard deviation statistic being proposed in the literature.

(ii) *Ratio of observed standard deviation to 'idealised' standard deviation*

This approach weights the standard deviation measure by a factor known as the 'idealised' standard deviation. The compensated (ratio) measure (σ_R), for when there is a difference in the number of games played by the teams in the league, is given by the ratio of the observed standard deviation to an 'idealised' standard deviation. It can be expressed as:

$$\sigma_R = \frac{\sigma_L}{\sigma_I}$$

Where:
σ_L = The observed standard deviation of a league
σ_I = The 'idealised' standard deviation of a league

The minimum value for this measure, corresponding to a perfectly balanced league is zero. The maximum value, which occurs if the top team wins all of their games, the second team wins all of their other games etc., depends on the specified variable (absolute value or share) and, if the variable is the share of maximum possible wins, on the number of games played per team. Note that consequently in both cases an increase (decrease) in the number of teams in the league, or in the number of games per team, reduces (increases) the maximum possible value *even if there is no change in the competitive balance in the league*. Furthermore, if the probability of a draw exists, the maximum value depends on the points system used by the league.

'Idealised' standard deviation (ISD)

Following the suggestion by Noll (1988), applied by Scully (1989), an assumption made in the literature to derive the 'idealised' standard deviation is that teams have "*equal playing strengths*". Quirk and Fort (1992) introduced the additional assumptions that:

I. there are *only two* possible outcomes, and
II. each outcome is *equally probable*.

This forms the basis for the binomial-based ISD approach.

Both of the assumptions added by Quirk and Fort (1992), underpinning the binomial-based ISD calculation, have been fundamentally challenged in the literature. Cain and Haddock (2006) argue that the first assumption is incorrect for sports with the 'trinomial case' where a draw outcome is possible (i.e. with more than two possible outcomes). Trandel and Maxcy (2009) argue that the second assumption is incorrect if there are other influences, such as an advantage for the 'home' team (i.e. even with symmetrical pairs of outcomes but in which each outcome does not have equal probability).

Neither of these challenges is to the approach of a standard deviation measure relative to an ISD measure. However, they produce different ISD (for different circumstances) and hence different quantifications of competitive balance. The three measures of ISD are presented below.

a) Binomial-based ISD

Quirk and Fort (1992, p. 245) state that:
"Using the Noll-Scully approach, we can evaluate the degree of competitive balance in a league by comparing the realized values of the standard deviation of the W/L percentages for a league to an idealized measure, namely, the standard deviation of W/L percentages for a league in which, for every team, the probability of winning any game is one-half." Quirk and Fort (1992) infer from the assumption of "*equal playing strengths"* the nested assumption that the event of a draw can be treated as a 'half win, half loss'. Then, as there are only two ('binary') possible outcomes for the measure (i.e. there is no draw outcome) it is assumed that the number of wins (W) follows a binomial distribution with a constant probability of winning (0.5) over independent trails; hence the winning percentage (W/G) follows a binomial distribution with an average of 0.5 and a variance of $0.25/G$.

The 'idealised' standard deviation of win percentages, on this binomial basis (i.e. if there is no draw outcome) is given by:

$$\sigma_I^{(w,l)} = \frac{0.5}{\sqrt{G}}$$

Where:
$\sigma_I^{(w,l)} =$ The 'idealised' standard deviation of win percentages on a binomial (win, loss) outcome basis
$G =$ Number of games per team

This approach has been widely adopted to compare Major US sports leagues for different sports. See, for example, Quirk and Fort (1992, Table 7.1, p.247) for average 'idealised' standard deviations of won-lost percentages for all the major US sports leagues, AL, NL, NBA, NFL and NHL, for each of the relevant decades from 1901 to 1990, and Vrooman (1995, Table II, p. 984) for US MLB AL and NL leagues, NBA and NFL leagues, 1970-1992).

Quirk and Fort (1992) also use this approach to consider the effect of league expansion on competitive balance in the US major sports leagues. They note that "[f]or the first few years of their history, expansion teams are manned primarily by players acquired in the expansion draft, which pretty much ensures very weak teams with low W/L records. Thus, periods of league expansion tend to be periods of highly dispersed W/L percentages." … "The extremely weak W/L records of expansion teams are matched by unusually strong W/L records of existing teams during a period of expansion, both of which act to increase the dispersion of W/L percentages for a league." (ibid p.250)

Berri, Schmidt and Brook (2007) also use this measure to assess competitive balance in the two leagues of the US Major League baseball from 1901 to 2006 (see Berri, Schmidt and Brook, 2007, Table 4.1, p. 51 and Table 4.2, p. 52). The period they cover includes 1995-99 which was the period considered by the Blue Ribbon Panel on Baseball Economics (Levin, R. et al, 2000). Berri, Schmidt and Brook (2007) point out that both the Blue Ribbon period provides a very small sample sizes which does not support strong statistical inferences.

b) Possible draw outcome ISD

Cain and Haddock (2006) highlight a limitation of the nested assumption that a draw can be treated the same as a 'half win, half loss' to produce the 'idealised' standard deviation weighting for the measure of competitive balance that can be used to compare leagues for different sports where a draw outcome is permitted. They point out that, if the rules of a game require it to be played to a positive (win, loss) conclusion, the relative strength of the teams is represented by the result. However, if the rules of a game permit it to end without a positive conclusion the result, if it were to have been played to a conclusion, is unknown.

As the Major US sports provide a positive result in (almost all) games the approach of Quirk and Fort (1992) can be viewed as reasonable to make comparisons of competitive balance between these sports leagues. However, this is not the case for other sports, notably European football, where Cain and Haddock (2006) found that approximately 25 per cent of games played in both of the top two league of professional football in England from 1888/89 to 2003/04 were drawn. The findings of Cain and Haddock (2006) are very similar to those of Koning (2000) who found that 26 per cent of all games played in the Dutch professional football league from 1956/57 to 1996/97 were drawn.

In the event of a draw outcome the relative strengths of the teams is not revealed and the nested assumption that a draw can be treated the same as a 'half win, half loss' is not valid since, were the rules to have required drawn games to be played to a positive conclusion, the stronger team would (by definition) be more likely to win more than half of these games. Hence leagues with draw outcomes will appear to be more competitively balanced than would be revealed by a league that played games to a conclusion. Consequently they argue that it is erroneous to use the binomial assumption for sports that permit a draw outcome even if a draw has half the points value of a win.

The approach proposed by Cain and Haddock (2006) is to provide a more general formulation of the 'idealised' standard deviation based on an exogenously specified probability of a draw outcome for the sport (league). This equates to the nested assumption of Quirk and Fort (1992) only in the case where the probability of a draw is zero.

When the (non-zero) probability of a draw is included, the weight given to a win relative to a draw outcome can vary from the weight given to a draw relative to a loss outcome. Consequently, for the approach proposed by Cain and Haddock (2006), the 'idealised' standard deviation is dependent on the system used by the league to allocate points for each outcome.

Owen (2012) shows the derivation of the 'idealised' standard deviation both when the variable for each team is:

(i) the absolute number of points awarded
(ii) the share of the maximum possible points

In both cases, the scoring systems for the league is represented in the format (w,d,l) denoting the allocated points for the respective outcome of win, draw and loss.

In all cases, let:
X_i = The number of points awarded to team i in any game
G = The number of games per team per season
Y_i = The number of points awarded to team i in a season
y_i = The share of wins relative to the maximum possible number of wins for team i in a season
α = The number of points awarded for a win
d = Probability of a draw
Probability of a win = Probability of a loss = $\left(\frac{1-d}{2}\right)$
$E(.)$ and $V(.)$ denote the expected value and variance, respectively

(i) <u>Absolute points based standard deviation measures</u>

The points based measures are normally applied to leagues with the scoring systems $(2,1,0)$ or $(3,1,0)$. The $(2,1,0)$ derivation is also applicable more generally to any league with a scoring system of the form $(\alpha, 0.5\alpha, 0)$.

a) System: $(\alpha, 0.5\alpha, 0)$

For each game:

$$E(X_i) = \left[\left(\frac{1-d}{2}\right) \times \alpha\right] + \left[d \times \frac{\alpha}{2}\right] + \left[\left(\frac{1-d}{2}\right) \times 0\right]$$

$$E(X_i) = \frac{\alpha}{2}$$

And:

$$V(X_i) = \left\{\left[\frac{1-d}{2}\right] \times \left[\alpha - \left(\frac{\alpha}{2}\right)\right]^2\right\} + \left\{d \times \left[\left(\frac{\alpha}{2}\right) - \left(\frac{\alpha}{2}\right)\right]^2\right\} + \left\{\left[\frac{1-d}{2}\right] \times \left[0 - \left(\frac{\alpha}{2}\right)\right]^2\right\}$$

$$V(X_i) = \frac{\alpha^2(1-d)}{4}$$

For a season:

$$V(Y_i) = V(X_i) \times G$$

$$V(Y_i) = \frac{\alpha^2(1-d)}{4} \times G$$

The 'idealised' standard deviation is:

$$\boldsymbol{\sigma_I}^{(\alpha,0.5\alpha,0)} = \sqrt{\frac{\boldsymbol{\alpha^2(1-d)}}{\boldsymbol{4}} \times \boldsymbol{G}}$$

Then, for example, if $\alpha = 2$, $d = 0.25$ and each team plays 38 games in the season

$$\sigma_I{}^{(\alpha,0.5\alpha,0)} = \sqrt{28.5} \cong 5.338$$

b) System: $(3,1,0)$

For each game:

$$E(X_i) = \left[\left(\frac{1-d}{2}\right) \times 3\right] + [d \times 1] + \left[\left(\frac{1-d}{2}\right) \times 0\right]$$

$$E(X_i) = \left(\frac{3-d}{2}\right)$$

And:

$$V(X_i) = \left\{\left[\frac{1-d}{2}\right] \times \left[3 - \left(\frac{3-d}{2}\right)\right]^2\right\} + \left\{d \times \left[1 - \left(\frac{3-d}{2}\right)\right]^2\right\} + \left\{\left[\frac{1-d}{2}\right] \times \left[0 - \left(\frac{3-d}{2}\right)\right]^2\right\}$$

$$V(X_i) = \frac{(1-d)(d+9)}{4}$$

For a season:

$$V(Y_i) = \left[\frac{(1-d)(d+9)}{4}\right] \times G$$

The 'idealised' standard deviation is:

$$\boldsymbol{\sigma_I^{(3,1,0)} = \sqrt{\left[\frac{(1-d)(d+9)}{4}\right] \times G}}$$

Then, for example, if $d = 0.25$ and each team plays 38 games in the season

$$\sigma_I^{(3,1,0)} \cong \sqrt{65.9} \cong 8.118$$

(ii) <u>Share of maximum possible wins based standard deviation measures</u>

The share based measures are normally applied to leagues, such as the Major US sports leagues, with the scoring system $(1,0.5,0)$. This is a specific case of a league with the more general scoring system of the form $(\alpha, 0.5\alpha, 0)$.

a) System: $(\alpha, 0.5\alpha, 0))$

For a season:

$$V(y_i) = V\left(\frac{Y_i}{\alpha \times G}\right) = V(Y_i) \times \frac{1}{(\alpha \times G)^2}$$

$$V(y_i) = \left\{\frac{\alpha^2(1-d)}{4} \times G\right\} \times \left\{\frac{1}{(\alpha \times G)^2}\right\}$$

$$V(y_i) = \frac{(1-d)}{4 \times G}$$

The ‘idealised’ standard deviation is:

$$\boldsymbol{\sigma_I}^{(\boldsymbol{\alpha},\mathbf{0.5}\boldsymbol{\alpha},\mathbf{0})} = \sqrt{\left[\frac{(\mathbf{1}-\boldsymbol{d})}{\mathbf{4}\times\boldsymbol{G}}\right]}$$

Then, for example, if $d = 0.25$ and each team plays 38 games in the season

$$\sigma_I{}^{(\alpha,0.5\alpha,0)} \cong \sqrt{0.0049} \cong 0.0702$$

Note that the binomial-based ISD for win percentages (actually ‘proportions’), with the assumption that a draw can be treated the same as a ‘half win, half loss’, is:

$$\sigma_I{}^{(w,l)} = \frac{0.5}{\sqrt{G}} \cong 0.0811$$

Hence, in this case, (i.e. with a share of maximum possible wins variable, a $(\alpha, 0.5\alpha, 0)$ system and a probability of a draw result of 0.25), the binomial-based ISD assumption that a draw can be treated the same as a ‘half win, half loss’ produces a ratio of standard deviation measure of competitive balance, $\sigma_R{}^{(Q\&F)}$, which understates the level of competitive balance by approximately 15% compared to the Cain and Haddock (2006) measure, $\sigma_R{}^{(C\&H)}$.

More generally, for any assumed level of draws, with a share of maximum possible wins variable and a $(\alpha, 0.5\alpha, 0)$ system, the binomial-based ISD assumption that a draw can be treated the same as a ‘half win, half loss’ understates the level of competitive balance compared to the Cain and Haddock (2006) approach by a factor of $\left(\frac{1}{\sqrt{1-d}} - 1\right)$. The Cain and Haddock (2006) approach only produces the same result as the binomial-based ISD when the possibility of a draw result does not exist (i.e. when $d = 0$).

b) System: $(3,1,0)$

It is not normal to adopt the ‘share of maximum possible wins’ variable to a league with the $(3,1,0)$ scoring system but, for completeness, in this case the ‘idealised’ standard deviation for a season is derived as follows:

$$V(y_i) = V\left(\frac{Y_i}{\alpha \times G}\right) = V(Y_i) \times \frac{1}{(\alpha \times G)^2}$$

$$V(y_i) = \left\{\left[\frac{(1-d)(d+9)}{4}\right] \times G\right\} \times \left\{\frac{1}{(3 \times G)^2}\right\}$$

$$V(y_i) = \left[\frac{(1-d)(d+9)}{4 \times 9 \times G}\right]$$

The ‘idealised’ standard deviation is:

$$\boldsymbol{\sigma_I}^{(\mathbf{3},\mathbf{1},\mathbf{0})} = \sqrt{\left[\frac{(\mathbf{1}-\boldsymbol{d})(\boldsymbol{d}+\mathbf{9})}{\mathbf{4}\times\mathbf{9}\times\boldsymbol{G}}\right]}$$

Then, for example, if $d = 0.25$ and each team plays 38 games in the season

$$\sigma_I{}^{(3,1,0)} \cong \sqrt{0.0051} \cong 0.0712$$

Importantly, Owen (2012) also points out that, unlike with the ('incorrect' for games that have a draw possibility) ISD, using these ('correct' for games that have a draw possibility) ISD it makes no difference to the calculated value of the ratio of observed standard deviation to 'idealised' standard deviation, ${\sigma_R}^{(C\&H)}$, for a given points system if the variable (X_i^*) is expressed as a share of the maximum potential value or in absolute values.

c) 'Home advantage corrected' ISD

Trandel and Maxcy (2011) challenge the assumption, made by Quirk and Fort (1992) for the binomial-based ISD, that when teams have "*equal playing strengths"* each outcome is *equally probable.* They argue that "… even two "perfectly-balanced" athletic teams are not equally likely to win any particular contest. Rather, the results of nearly all sporting events show that teams … playing in the location to which they are accustomed (i.e. playing at home) are more likely to win than are teams playing on the road … Since home advantage differs by sport, the errors also differ …". (ibid pp. 2-3). Consequently the ratio measure presented by Quirk and Fort (1992) does not provide the intended (as stated) comparative measure between sports.

Trandel and Maxcy (2011) provide a general expression for the home advantage corrected standard deviation of an evenly balanced league (with binomial outcomes) as:

$$\sigma^{HISD} = 2\sum_{k=1}^{g/2}\left\{\sum_{i=1}^{k}\frac{\left((g/2)!\right)^2 h^{\left((g/2)-k+2i-1\right)}1-h^{\left((g/2)+1+k-2i\right)}\left(.5-\frac{k-1}{g}\right)^2}{\left((g/2)-k+i\right)!\ (k-i)!\ \left((g/2)+1-i\right)!\,(i-1)!}\right\}$$

Where:
$\sigma^{HISD} =$ Home advantage corrected standard deviation of an evenly balanced league (with binomial outcomes)
$h =$ Probability of winning a game played at home
$k - i =$ (unordered) outcomes from $G/2$ possibilities
$g =$ Number of games played in the league by each team

Trandel and Maxcy (2011, p. 8) state that "[a]nalysts who have ignored this effect have therefore concluded that leagues have a greater degree of competitive balance than is truly justified." Without the correction, the binomial-based ISD will larger than the corrected measure if there is home advantage in the league. As a consequence leagues appear (with the binomial-based ISD) to be more (incorrectly) balanced for sports that have more home advantage.

However, more generally, the approach of producing a ratio of the standard deviation to an 'idealised' standard deviation based on the assumption that teams have "*equal playing strengths"* (σ_R) as a comparative measure of competitive balance has been criticised in the literature.

First, it does not address the fundamental criticism, as with any measure (including the standard deviation (σ_L) measure) with an upper bound that depends on the

number of teams in the league and/or the number of games played by each team (and the σ_R measure depends on both) that since the measure does not provide a value relative to the theoretical maximum imbalance, it also captures the scale effect of the league (which is unrelated to the competitive balance of the league). Owen (2010) shows, the case with the binomial-based ISD and a league in which each of the (N) teams play every other team the same (G) number of times, that the upper bound of the ratio measure $\left(\sigma^{(Q\&F)\ ub}{}_R\right)$ is given by:

$$\sigma^{(Q\&F)\ ub}{}_R = 2 \times \sqrt{\frac{G(N+1)}{12}}$$

Hence, although the ratio of observed standard deviation to the 'idealised' standard deviation (σ_R) explicitly incorporates the number of teams in the league and the number of games played by each team, as Owen (2010, p. 38) states "… it does not control for these variables in the sense of partialling out their effects." The problem is clear in the case where, because of the (positive) difference in the number of teams or games played per team in the respective leagues, a league that has less than maximum competitive balance produces a larger value for the measure than another league with maximum competitive balance. He notes, for example, that the upper bound (i.e. perfectly unbalanced) value for the First Division of New Zealand Rugby Union's Provincial Championship of 1.826 (in 1990) and 2.000 (in 1994). These are less than all the calculated values for the NBA from 1980 to 1992 reported by Vrooman (1995).

Furthermore, the interpretation of the lower limit of the measure is not clear. Quirk and Fort (1992, p. 246) note that "[t]he closer is the ratio of actual to idealized standard deviation to 1, the more competitive balance there is in the league." However, as the lower bound of the observed standard deviation (i.e. the numerator of the ratio measure) has a theoretical limit of zero, the ratio measure (σ_R) also has a lower limit of zero (regardless of the value of 'idealised' standard deviation). If the value of the observed standard deviation is less than the 'ideal' standard deviation (i.e. $\sigma_R < 1$), as Goosens (2006, p. 87) points out, "… in terms of the interpretation given by Quirk and Fort, we find a competition that is more equal than when the league is perfectly balanced." Goosens (2006, p. 87) notes, for example, that this occurred in the top football league in Germany in 1969 when the ratio (using the binomial-based ISD) had a value of 0.695.

Vrooman (1995, note 26, p.984) notes that a problem with this measure for inter-league comparisons could be that competition variances may reflect the greater importance of a few players for team production on smaller rosters.Owen (2010) also shows that the sensitivity of the metric to the number of teams in the league is greater than for the σ_L measure. Lenten (2009) notes that, since this measure (σ_R) is a relative standard deviation measure, it is highly sensitive to the occasional outlier observation.

Scully (1992) suggests that differences between the actual and ideal win percentages in a league could, to some extent, be explained by what is known as "momentum" in sports or serial correlation in time series analysis. He argues that the win percentage of a team can take more than one season to take full effect.

Consequently, shocks arising from exogenous changes to the playing rules, for example, may impart cycles in the win percentages and differences compared to the ideal win percentage.

Goosens (2006) observes that the fact that a league diverges from these 'ideal' league definitions does not mean that intervention is necessary. Whilst a certain level of competitive balance seems reasonable to hold the interest of spectators and sponsors neither definition of 'ideal' has been proven to be the 'optimal' level of competitive balance. Or, as Horowitz (1997, p. 374) noted, "… a league with teams of equal strength is not necessarily welfare enhancing … [and] … there may be an optimal level of team dominance."

(iii) *Normalised standard deviation* (σ_L^*)

As neither the standard deviation measure of competitive balance nor the ratio of the standard deviation to an 'idealised' measure themselves remove the scale effects of the league, an alternative approach proposed in the literature is to normalise the measure with respect to its upper limit.

Goosens (2006) proposed a normalised measure (for the standard deviation of win percentages) called the National Measure of Seasonal Imbalance $(NAMSI)$. This measure incorporates both the maximum and minimum possible standard deviation values to produce a measure with a maximum value of one (corresponding to a league with maximum competitive imbalance) and a minimum value of zero (corresponding to a league with perfect competitive balance), regardless of the number of teams in the league. The measure $(NAMSI)$ is calculated by:

$$NAMSI = \sqrt{\frac{\sum_{i=1}^{n}(w_i - 0.5)^2}{\sum_{i=1}^{n}(w_{i\,max} - 0.5)^2}}$$

Where:
$w_i =$ Win percentage of team i
$w_{i\,max} =$ Maximum value of win percentage of team i (with perfect imbalance)
$N =$ Number of teams (i) in the league

Owen (2010) provides an equivalent normalised measure but with a general expression for the upper bound. The measure (σ_L^*) is calculated by:

$$\sigma_L^* = \frac{\sigma_L'}{\sigma_L^{ub}}$$

Where:
$\sigma_L' =$ The standard deviation of win percentages statistic for a single season given by:

$$\sigma_L' = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(\frac{W_i}{G_i}\right) - 0.5\right]^2}{N}}$$

$\sigma_L^{ub} =$ The upper bound of σ_L given by:

$\sigma_L^{ub} = \sqrt{\frac{(N+1)}{12\,(N-1)}}$

$W_i =$ Number of wins of team i
$G_i =$ Number of games of team i
$N =$ Number of teams (i) in the league

Owen (2010) shows that the normalised standard deviation measure (applied to either 'win percentage' or 'absolute points' data) produces the identical result to the equivalent ratio of observed standard deviation to the 'idealised' standard deviation (σ_R) measure if that is also normalised with respect to its upper bound.

To simplify the terminology, let:
$ASD =$ Actual observed standard deviation
$ASD^* =$ Normalised actual observed standard deviation
$ASD^{ub} =$ Upper bound for the actual observed standard deviation
$ISD =$ 'Idealised' standard deviation
$RSD =$ Ratio of observed standard deviation to the 'idealised' standard deviation
$RSD^* =$ Normalised ratio of observed standard deviation to the 'idealised' standard deviation
$RSD^{ub} =$ Upper bound ratio of observed standard deviation to the 'idealised' standard deviation

Given that
$RSD^* \equiv \frac{RSD}{RSD^{ub}}$ and $ASD^* \equiv \frac{ASD}{ASD^{ub}}$

And
$RSD \equiv \left(\frac{ASD}{ISD}\right)$ and $RSD^{ub} \equiv \left(\frac{ASD^{ub}}{ISD}\right)$

It follows that

$$RSD^* = \frac{\left(\frac{ASD}{ISD}\right)}{\left(\frac{ASD^{ub}}{ISD}\right)}$$

Hence

$$RSD^* = \frac{ASD}{ASD^{ub}}$$

And

$$RSD^* = ASD^*$$

These normalised measures provide a measure of competitive balance which is applicable to compare 'open' (or 'closed') leagues of any size or number of games per team. Furthermore, for any given points scoring system it can be used to compare between 'points based' and 'win share based' measures. So a Major US sport league with a $(1,0)$ system can be compared with a FIFA league with a $(\alpha, 0.5\alpha, 0)$ system.

The Goosens (2006) approach can be preferred to the Quirk and Fort (1992) approach with the assumption of "equal playing strengths" because:

1. It provides a measure relative to the maximum value attainable and hence a league cannot indicate more competitive balance than another with the maximum for that league.
2. It also provides a more natural interpretation of both the upper (maximum imbalance) and lower (perfect balance) limits.
3. It does not require the use of an 'idealised' standard deviation (and associated issues)

Note, however, that the normalised measures are not invariant with respect to non-linear transformations of the point scoring system used by the leagues. So they cannot be used instead of the ratio of observed standard deviation to the 'idealised' standard deviation (σ_R) measure to compare, for example, a $(\alpha, 0.5\alpha, 0)$ system league with a $(3,1,0)$ system league.

(iv) *<u>Standard deviation of team strength</u>*

Koning (2000) proposes a measure of competitive balance that allows for win, loss and draw outcomes in games (which is not related to an 'idealised' standard deviation with the assumption of "equal playing strengths"). Fort (2007) refers to this approach as a significant extension of the literature (referenced by Koning, 2000) of the trinomial case in soccer. It is based on the standard deviation of a derived variable that represents the relative strength of teams in the league.

To calculate the strength of each team, Koning (2000) defines a latent variable model to transform the latent difference in strength between teams in a game into an observed outcome of a game (won, drawn or lost for the home team). He formulates the possibilities of the possible outcomes as functions of the team strength and home advantage parameters and uses an ordered probit model to derive the probability that a game is won, drawn or lost. He then uses maximum likelihood estimation to estimate the model parameter values. The measure is calculated as the standard deviation of the team quality parameters (i.e. deviation from the average quality) for each season. A low value for the measure indicates that there was little variation in quality between the teams and that the league was balanced.

An advantage of this approach is that the measure is invariant to changes in the number of points awarded for a win or a draw. Koning (2000, p.426) also points out that the measure is invariant under changes in home advantage as the model separates home advantage from the team quality measure. However, as Goosens (2006, p. 91) notes, "[t]he use of this measure is not straightforward and an advanced knowledge of econometrics is necessary to apply it."

Koning (2000) estimates the model for the top league of football in the Netherlands (since the current competition was introduced in the 1955/56 season). He finds no systematic change in competitive balance until the mid-1960s. There was a marked decrease between 1965 and 1970 followed by an increase until 1976. There is no clear trend after that. Konng (2000, Fig. 2., p. 427) provides a plot of the standard deviation of the team strength parameter and of the standard deviation of absolute points each season from the mid 1950's to the mid 1990's which shows the closeness between the two measures.

3. Coefficient of variation

This measure is given by Goosens (2006) in a review of measures of competitive balance. It is calculated as follows:

$$Coefficient\ of\ variation = \frac{Standard\ deviation}{Mean}$$

Goosens (2006) notes correctly that the basic standard deviation measure cannot be used to make valid comparisons of competitive balance between leagues which operate with 2,1,0 points systems with leagues that operate with 3,1,0 point systems as the mean values of leagues with 3,1,0 point systems differ. More generally, this is true if the different point systems are not linear transformations of each other.

However, Goosens (2006) then states that in this case the use the coefficient of variation is necessary. Whilst this is correct if the games are played to a win-lose binary conclusion, it is not correct more generally. This can be seen by comparing the ratio of standard deviation and mean derived for the 2,1,0 system in the calculation of the 'idealised' standard deviation when there is a non-zero possibility of a draw outcome with the equivalent ratio for 3,1,0 systems (see 'Possible draw outcome ISD' above). The coefficient of variation is only identical when the probability of a draw is zero (i.e. $d = 0$). As the draw possibility is explicitly included in at least one of the leagues being compared, it can be concluded that this measure cannot be used as stated.

4. "Excess tail frequencies" for win percentages

This is a measure of the difference between two distributions - the actual distribution of win percentages for a league over its history and the idealised distribution based on the assumption of equal playing strength. The idealised distribution is closely approximated by the bell-shaped normal curve where, of the win percentages, approximately:

- two thirds lie within one standard deviation of the mean (0.5)
- 95% lie within two standard deviations of the mean (0.5)
- 99% lie within three standard deviations of the mean (0.5)

The "excess tail frequencies" measure of competitive balance is calculated as the difference between the percentage of cases that actually lie in the tails of the win distribution and the percentage of cases that would lie in the tails if all teams were of equal playing strength. The leagues are then ranked according to their degree of competitive balance.

The issues associated with the use of the 'idealised' measure relating to its use with the standard deviation measure also apply to this measure.

See, for example, Quirk and Fort (1992, Table 7.3 p. 256) for all the major US sports leagues, AL, NL, NBA, NFL and NHL. A related measure was adopted by Lee and Fort (2005) as a variable to identify structural change ("break points") in competitive balance for the MLB AL and NL from 1901 to 1999.

5. Relative entropy $(\boldsymbol{R})$

Horowitz (1997) proposes to measure the competitive balance of a sports league through the relative entropy measure of information theory (which is an adaptation of the Theil index statistic used to measure economic inequality). Horowitz (1997) proposes a measure of relative entropy for a sports league where all teams in the league play each other the same number of times. He describes this as a measure of "the degree of uncertainty about which team might have won a randomly-selected game relative to the maximum uncertainty possible." (ibid p. 376). The measure (R) is calculated as:

$$R = \frac{E}{E_{max}}$$

Where:
$E = -\sum_i p_i log_2 p_i$
$p_i =$ Proportion of the total of all wins in the league season of team i
$E_{max} = -\, log_2 \left(\frac{1}{N}\right)$
$N =$ Number of teams (i) in the league

The maximum value for this metric is one, corresponding to a perfectly balanced league, which occurs when the actual uncertainty equals the maximum possible uncertainty. The minimum value, corresponding to the least possible balance in the league, which occurs if the top team wins all of their games, the second team wins all of their other games etc., depends on the number of teams in the league. For example, a league with 20 teams has a minimum value of approximately 0.93. A decrease (increase) in the number of teams in the league does increase (reduces) the possible range for this measure. However, a league with only 10 teams has a minimum value of 0.89.

This measure can be applied to 'open' leagues, as it is a measure of the league given the teams in the league at that time, but is not suitable to compare leagues of different sizes as the scale of the measure is dependent on the number of teams.

Horowitz (1997) calculates the values of R for the American and National baseball leagues, from 1903 to 1995, and uses them to identify the trend in competitive advantage in each league over time. The regression method employed also enables Horowitz (1997) to test for the effect of factors that could be expected to affect the overall trend in competitive balance in a league either in the short or long run.

6. Herfindahl-Hirschman Index (HHI) – for concentration in a league season

The standard Herfindahl-Hirschman Index (HHI), used to measure concentration in an industry, is defined as:

$HHI = \sum_{i=1}^{N} s_i^2$
Where:
$s_i =$ Market share of firm i
$N =$ Number of firms in industry

This industry measure has a range from $\left(\frac{1}{N}\right)$ where all firms have an equal market share, to one, where the industry has a single firm monopolist.

An inherent problem with this measure, if used to compare between industries, is that it is correlated with the number of firms in an industry. This can be seen most clearly at the lower bound, when all firms have equal shares. An industry with more firms will produce a lower value for the measure than an industry with fewer firms (also with equal shares).

A normalised measure (HHI^*) with a minimum value of zero and a maximum value of one is given by:

$$HHI^* = \frac{HHI - \left(\frac{1}{N}\right)}{1 - \left(\frac{1}{N}\right)}$$

Where:
$HHI = \sum_{i=1}^{N} s_i^2$
$s_i =$ Market share of firm i
$N =$ Number of teams (i) in the league

As a measure the competitive balance of sports leagues in terms of the *concentration of the teams in a season*, the standard HHI has been applied by replacing the 'firm' with the 'team' and replacing 'market share of a firm' with 'league attainment *for a team over a season'*.

In this case, without the normalising adjustment, the inherent problem with the measure results in an overstated value of competitive balance in leagues with fewer teams compared to leagues with more teams.

A key difference with the application of the HHI to a sports league, compared to an industry, is that, the league structure imposes a restriction on the range of the measure which also depends on the number of teams in the league.

Three adapted measures of the HHI, proposed to measure the competitive balance of sports leagues in terms of the concentration of the teams in a season, are shown below. The first measure is the most direct application of the industry measure with the league structure limitation. The second measure recalibrates the metric to a lower bound of zero (but still with the league structure limitation). The third measure allows for differences in the number of teams (i.e. league structure restriction) and therefore provides a valid comparable measure between leagues of different sizes.

These measures are independent of the number of games played.

These 'single season' measures are applicable to 'open' leagues.

(i) <u>Share of all wins (games or points) in a season (HHI^C)</u>

League attainment for a team in a season has been defined in the literature in terms of games won (see, for example, Depken, 1999) and points won (see, for example,

Michie and Oughton, 2004). In these cases of measures of concentration between teams in a season the measure (HHI^C) is calculated as:

$$HHI^C = \sum_{i=1}^{N} \left(\frac{W_i}{\sum_{i=1}^{N} W_i} \right)^2$$

Where:
$W_i =$ Number of wins for team i's
$N =$ Number of teams (i) in the league

This measure has a minimum value of $\left(\frac{1}{N}\right)$ where each team wins the same number of games (or points). However, in a sporting context where each game comprises two (and only two) teams, if more than two teams are in the competition and all teams play the same number of games, the maximum value must be less than one. In a league structure the maximum value occurs if the top team wins all their games, the second team wins all their other games etc., is shown by Owen et al (2007, equation 3, p. 292) to be given by:

$$HHI^C_{max} = \frac{2(2N-1)}{3N(N-1)}$$

Where:
$N =$ Number of teams (i) in the league

For a league with twenty teams, for example, the minimum value is 0.05, with a perfectly competitive balanced league, and a maximum value of approximately 0.068. An increase in the number of teams in the league reduces both the upper and lower bounds of this measure. An increase from 20 to 22 teams makes an imperceptible (0.0018) decrease in the possible range for the measure.

Michie and Oughton (2004) call the version of this measure, based on shares of points won in a season, the H-Index and graphically show the values for the top league of professional football in England from 1947 to 2004 (ibid Figure 2.5, p. 13) and for the English Premier League from 1992 to 2004 (ibid Figure 2.6, p. 13). They find that the HHI measure "was roughly constant between 1947 and the 1980s, and that after that showed a rise. Over the period as a whole the H-index increased by around 13 per cent. However, part of this increase is due to the reduction in the number of teams (from 22 to 20)." (ibid pp. 13-14)

Michie and Oughton (2004) also convert this measure into the Herfindahl Index of Competitive Balance (HICB) noting that "the Herfindahl Index is sensitive to changes in the number of teams. This can be corrected for by dividing the index value of H that would be attained in a perfectly balanced league to give the H Index of Competitive Balance (HICB) as shown below:" (ibid p.14)

$$HICB = \left(\frac{\sum_{i=1}^{N} s_i^2}{1/N} \right) x\ 100$$

Where:
$s_i =$ Team i's share of points in a season
$N =$ Number of teams (i) in the league

They claim that this indexation improves the intuitive interpretation of the H measure as the index has a value of 100 for a perfectly balance league of any size. However, the upper limit still depends on the number of teams in the league (and not on the number of games played by teams in the league, provided that they all play the same number of games).

Michie and Oughton (2004) graphically represent the values for the H-index and HICB measures with data for the top league of professional football in England, 1947-2004 (ibid Figure 2.5, p. 13 and Figure 2.7, p. 15), Italy, 1956-2004 (ibid Figure 2.9, p. 19), Germany, 1964-2004 (ibid Figure 2.10, p. 19), France, 1954-2004 (ibid Figure 2.11, p. 20) and Spain, 1956-2004 (ibid Figure 2.12, p. 20).

(ii) <u>Deviated Herfindahl-Hirschman Index $(dHHI)$</u>

Depken (1999) proposes a measure of the deviation of the HHI^C measure from the 'ideal' distribution of wins. The measure $(dHHI)$ is calculated by:

$$dHHI = HHI^C - \frac{1}{N}$$

Where:
$HHI^C = \sum_{i=1}^{N} \left(\frac{W_i}{\sum_{i=1}^{N} W_i}\right)^2$
$W_i =$ Number of wins for team i's
$N =$ Number of teams (i) in the league

The interpretation of 'ideal' with this measure is again open to the criticisms cited with its use with the standard deviation measure. However, in this case, the adjustment has the effect of giving the measure a lower limit of zero, regardless of the number of teams in the league. Nevertheless, the upper limit is not fixed as it depends on the number of teams. The maximum value, which occurs if the top team wins all their games, the second team wins all their other games etc., is shown by Owen et al (2007, equation 4, p. 292) to be given by:

$dHHI_{max} = \frac{N+1}{3N\ (N-1)}$
Where:
$N =$ Number of teams (i) in the league

For a league with twenty teams, for example, the maximum value is approximately 0.018.

See Depken (1999, Figures 1 and 2, p. 210) for each league of MLB, 1920-1996.

Depken also shows that the relationship between this measure $(dHHI)$ and the measure of the standard deviation of win percentages (σ_L) is as follows:

$$\sigma_L = \sqrt{\frac{N}{4}\ x\ (dHHI)}$$

Depken states that this assumes:

a) All teams play the same number of games (which has not always been the case in MLB)
b) Each team plays an even number of games
c) All games have a positive result (i.e. there are no drawn games)

Depken (1999) uses his measure $(dHHI)$ as the dependent variable in a stochastic model, which he estimates using a Seemingly Unrelated Regression (SUR) approach, to identify the factors that have influenced the measure for the US MLB leagues. As a further development in the use of HHI measures, he also postulates that the distribution of talent between teams is a factor and introduces this by forming two independent variables from HHI calculated from the number of runs scored and number of runs allowed. He does not express these variables as deviations from the standard measure because he finds that this does not qualitatively alter the results. Standard deviation measures are also included as explanatory variables.

(iii) *Normalised Herfindahl-Hirschman Index* $(nHHI)$

To provide a comparable measure between leagues with different number of teams, Owen et al (2007) proposes a normalised measure of competitive balance, with a range of zero, corresponding to perfect balance, to one, in the case of the most imbalanced league. The measure $(nHHI)$ is related to the HHI^C and $dHHI$ measures and calculated as follows:

$$nHHI = \frac{HHI^C - HHI^C_{min}}{HHI^C_{max} - HHI^C_{min}} = \frac{dHHI}{dHHI_{max}}$$

Where:

$HHI^C = \sum_{i=1}^{N} p_i^2$, $HHI^C_{min} = \frac{1}{N}$, $HHI^C_{max} = \frac{2(2N-1)}{3N(N-1)}$,

$dHHI = HHI^C - \frac{1}{N}$, and $dHHI_{max} = \frac{N+1}{3N\,(N-1)}$

$p_i =$ Team i's share of all games (or points) won in the league season

$N =$ Number of teams (i) in the league

7. Concentration ratios

The concentration ratio measures of the competitive balance of a sports league are based on an assessment of the concentration of a subset of the teams in the league (i.e. a number of top teams) relative to a comparative statistic. Two applications in the literature with different comparative statistics are presented below. In the first case the concentration of the top teams is compared to the maximum *attainable* points *they* could theoretically achieve. I call this measure, proposed by Koning (2000), the 'attainable' concentration ratio'. In the second case it is compared to the total number of points *attained* by *all teams in the league*. This version presented, for five teams, was proposed by Michie and Oughton (2004) who called it the 'C5 ratio'.Both measures could be applied to 'open' (or 'closed') leagues. Issues related to comparisons between leagues and over time are highlighted below.

(i) 'Attainable' concentration ratio

Koning (2000) defined the ratio for a team as the number of points achieved by a number of top teams relative to the maximum number of points they could attain. The ratio can be expressed as:

$$C_K = \frac{\text{Total points won by the top K teams}}{\text{Maximum possible number of points for the top K teams}}$$

Formally, this is given by:

$$C_K = \frac{\sum_{k=1}^{K} P_k}{KW(2N - K - 1)}$$

Where:
$K =$ The number of top teams
$W =$ The number of points awarded for a game won
$P_k =$ The number of points won by the kth from top team
$N =$ Number of teams (i) in the league

This measure has a maximum value of one, corresponding to the least possible balance in the league, which occurs if the top team wins all of their games, the second team wins all of their other games etc., to the kth from top team (i.e. all top teams secure the maximum possible points available given the constraint of the league structure). It has a minimum value, which occurs if all games are drawn, given by:

$$C_K^{min} = \frac{D(2N - 2)}{W(2N - 1 - K)}$$

Where:
$K =$ The number of top teams
$W =$ The number of points awarded for a game won
$D =$ The number of points awarded for a game drawn
$N =$ Number of teams (i) in the league

It should be added that a necessary condition of the measure is that $W \geq 2D$ (or a perfectly unbalanced league will produce fewer points than a perfectly balance league and the measure can become perverse). The limits are independent of the number of games played by each team, provided that they all play the same number of games against each other. Although the minimum value does depend on the number of teams in the league.

This measure takes no account of the concentration between the teams that are not included in the measure as top teams. Only their results against the top teams influence the measure. Koning (2000) calculates the values for this measure both with the top team only (C_1) and with the top four (C_4) teams in the top league of football in the Netherlands since the 1955/56 season. He finds that the top team captured around 75% of the points it could achieve each season until the mid-1960s. It then increased to a peak of 94% in the 1971/72 season. This was followed by an irregular period with no clear trend. He also found that, "[c]ontrary to common

opinion and despite the slight upward trend, the value of [(C_4)] is not high when compared with typical values encountered during the 1960s." (ibid p. 429)

(ii) <u>Five club concentration ratio</u> (C5 ratio)

A five club concentration ratio (C5 ratio) is proposed by Michie and Oughton (2004). The C5 ratio can be expressed as:

$$\text{C5 Ratio} = \frac{\text{Total points won by the top five clubs}}{\text{Total number of points won by all clubs}}$$

This measure has a minimum value with a perfectly competitive balanced league of 20 teams of 0.25 (i.e. 5/N, where N is the number of clubs). The maximum value occurs if the top 5 teams win all their games and the other 15 teams draw all their other games. In a league with 20 teams, where 3 points are awarded for a win, 1 point for a draw and 0 points for a loss, the C5 ratio has a maximum value of approximately 0.55 (i.e. M/(M+T) where M is the maximum number of points the top 5 clubs can achieve in a season and T is the minimum number of points possible for the remaining teams).

With reference to the C5 ratio, Michie and Oughton (2004, Table 2.1, p.16) "[n]ote that this maximum value [given by M/(M+T)] exceeds the value for a perfectly unbalanced league, given by:

$$\frac{\sum_{i=1}^{5}(N - X_i)}{\sum_{i=1}^{N}(N - X_i)}$$

Where:
$N =$ Number of teams (i) in the league
$X_i =$ League position of team i with the top club ranked 1 and the bottom club ranked N."

However, this is true if, *and only if*, the league awards *more than* 2 points for a win (with 1 point for a draw and 0 points for a loss). If the league awards 2 points for a win (and 1 point for a draw and 0 points for a loss), as was the case for the top league of professional football in England for the period covered by Michie and Oughton (2004) up to the 1981/82 season, the maximum values are identical. In this case, a league with 20 teams has a maximum value for the C5 ratio of less than 0.45. Hence the applicable range for this measure is further reduced.

An increase in the number of teams in the league also reduces the maximum value of the C5 ratio. However, the values for this measure are independent of the number of games played by teams in the league (provided that they all play the same number of games against each other).

Michie and Oughton (2004) graphically represent the values for the C5 ratio with data for the top league of professional football in England, 1947-2004 (ibid Figure 2.2, p. 10) and for the English Premier League (for football), 1992-2004 (ibid Figure 2.3, p.10).

They find that “between 1947 and 1987 the C5 ratio was unchanged: however, between 1989 and 2004 the index rose by 6.4 per cent.” (ibid p. 9) They conclude that “[t]hese changes are significant given the constraints imposed on the value of the index by the points scoring system.” (ibid p. 9)

Michie and Oughton (2004) also present the C5 ratio as an index based on the perfectly competitive “ideal” minimum value of the measure. They call this the C5 Index of Competitive Balance (C5ICB). The index value is calculated as follows:

$$C5ICB = \left(\frac{C5\,ratio}{5/_N}\right) x\ 100$$

Where:
$N =$ Number of teams (i) in the league

This indexation improves the intuitive interpretation of the C5 ratio measure. The index has a value of 100 for a perfectly balance league of any size and the increment in value equates to the percentage reduction in competitive balance. However, the upper limit still depends on the points system and the number of teams in the league (but not on the number of games played by teams in the league (provided that they all play the same number of games against each other).

Michie and Oughton (2004, Figure 2.4, p. 11) graphically represent the values for the C5ICB measures with data for the top league of professional football in England, 1947-2004.

8. Index of dissimilarity $(\boldsymbol{ID})$

The index of dissimilarity is a measure of difference in relative proportions of two groups. It is used, for example, in human geography to assess the degree of segregation between ethnic populations. Mizak and Stair (2004) propose an application of the index of dissimilarity as a measure of competitive balance for sports leagues. In this case the measure (ID) is given by:

$$ID = 0.5 \sum_{i=1}^{N} |X_i - Y_i|$$

Where:
$X_i = \frac{1}{N}$
$Y_i =$ Team i’s share of total league wins
$N =$ Number of teams (i) in the league

The interpretation of the index is that it gives the number of wins that would need to be ‘reallocated’ to produce equality of result. A higher value indicates less competitive balance.

The measure has a lower limit of zero, corresponding to a perfectly balance league. The upper limit depends on the number of teams in the league. In a league with 20 teams, for example, where all teams play each other the same number of times, the

maximum value is approximately 0.26. This indicates that, for a league with 20 teams, the difference between and maximum imbalance and perfect balance could be achieved (theoretically) by changing the result of 26% of the games from won to lost (or vice versa).

Mizak et al (2005) point out that a deficiency of this measure is that the same result could be produced, for example, by a league with a single dominant team and a league with a larger number of less dominant teams. This is because it does not include the distribution of wins in the league. This measure is applicable to both 'open' and 'closed' leagues. Mizak and Stair (2004) apply this measure to both leagues of the MLB for each year from 1986 to 2004 and for selected prior years back to 1929. Mizak and Stair (2004) also apply the index of dissimilarity to payroll data for the MLB teams and calculate the correlation coefficient with the index of dissimilarity for wins in the same year. Furthermore, they regress the index of dissimilarity for payroll data, together with dummy variables for the pre-1976 and subsequent seasons and the league (AL or NL) on the index of dissimilarity for wins. From this, they estimate the competitive balance in each league in the forthcoming season.

9. Lorenz curve and Gini coefficient – for teams within season

The Lorenz curve provides a graphical representation of inequality which can be measured by the Gini coefficient. When measuring, for example, income inequality, the Gini coefficient has a minimum value of zero, when the Lorenz curve follows the equality line and a maximum value of the reciprocal of the population size. This tends towards a value of one in the case of maximum inequality, corresponding to the situation where a single person has all the income and the population size tends towards infinity.

Michie and Oughton (2004) propose a version of the Lorenz curve to show the level of competitive balance in a league season which they call a Lorenz Seasonal Balance Curve (LSBC). They state that "… it measures seasonal competitive balance (rather than dominance) based on each team's share of points in a season, so that instead of calculating one curve for a 50 or 100 year time period we calculate the Lorenz curve for each season with each curve reflecting seasonal inequality." (ibid p.17)

To plot the LSBC they rank the teams in the league from the team with the lowest number of points up to the team with the highest number of points (i.e. title winner) and calculate the share of points won by each team. From this they calculate the cumulative share of points with the addition of each team in order of increasing points. The cumulative percentage of points is then plotted (on the y-axis) against the cumulative percentage of clubs (on the x-axis).

The Gini coefficient as, for example, adopted by Schmidt (2001, equation 1, p.22), for a season (t) of a sports league, if every team plays the same number of games in a season, is calculated by ranking each team relative to its win percentage (w_i), such that $w_N \geq w_{N-1} \geq \cdots \geq w_1$ and then the measure of concentration (G_t^C) is approximated by:

$$G_t^C = 1 + \frac{1}{N_t} - \frac{2}{N_t^2 \overline{w}_t}\ x\left(w_{i,t} + 2xw_{i-1,t} + 3xw_{i-2,t} + \cdots + Nxw_{1,t}\right)$$

Where:

$\overline{w}_t =$ Average value of all teams' win percentage in the league in season t

$N_t =$ Number of teams in the league in season t

See, for example, Schmidt (2001, Table 2, p. 23) for average values of Gini coefficients for decades of the two MLB leagues from 1901 to 1990 and 1991 to 1998.

However, Utt and Fort (2002) challenged the Gini coefficient calculated by Schmidt (2001), and also by Schmidt and Berri (2001), as a measure of competitive balance within seasons for two reasons. First, they note that the most unequal outcome cannot have one team winning all the games played in the entire league so the maximum value is less than one and consequently the calculated Gini coefficient understates the level of inequality.

Second, in the case of Major League Baseball (MLB), the measure "… ignores a host of other complexities, such as unbalanced schedules (teams do not play the same number of games against all opponents), league expansion, and interleague play and presents additional challenges." (Utt and Fort (2002, p.368)

Consequently, Utt and Fort (2002) conclude that "… until a remedy for these complications is found, we will stick with the tried and true standard deviation of winning percentages (and their idealized values) for within-season competitive analysis of winning percentages." (ibid p.373). Berri, Schmidt and Brook (2007) argue that the first point raised by Utt and Fort (2002) is not really a problem as "… we would also not likely see a Gini coefficient of one in any study of income inequality in a society. Although it is perhaps physically possible for a society to give all of its income to one person, such a society is not likely to be observed." (ibid p.242, note 32) In fact, in this example, a value of one would also require an infinite population and consequently it is not even possible.

Although Michie and Oughton (2004) do not provide Gini coefficients, the Lorenz curves they present (ibid Figure 2.8, p.18) would also be subject to the criticism of Utt and Fort (2002) in that they do not recognise that the maximum area of inequality below Lorenz curve is smaller than the area shown and that consequently the representation understates the impression of the level of inequality. As an example of the limitation imposed by a league structure, the Gini coefficient for the most unbalanced league (which occurs if the top team wins all of their games, the second team wins all of their other games etc) with twenty teams is 0.35 (rather than 0.95 without the constraint of the league playing structure).

In response to the second criticism by Utt and Fort (2002), that this measure "… ignores a host of other complexities", Mikak, Stair and Rossi (2005) note that the problems are due to the denominator of the measure not being fixed. As a solution, they suggest that the denominator of the Gini coefficient is disregarded altogether and the area of inequality (i.e. the numerator of the Gini coefficient) is made the measure.

10. Surprise index for leagues (S_L)

Groot and Groot (2003) take a different approach to measure the competitive balance of a league. They propose a measure, which they call the Surprise index. It is based on the actual result (i.e. win, draw, loose) of individual matches for each team which they weight with an assigned value for each outcome and the final position of those teams in the league for that season.

A 'surprise' is deemed to have occurred for a team that achieves a win or a draw against a team that finished the season in a higher position. For those games, the lower positioned 'surprise' team is allocated a 'score' of two points for a win and one point for a draw. The score for each game is weighted, to represent the degree of 'surprise', by a factor given by the difference in final league position (i.e. rank order) of the two teams involved in the game. The index is formed by expressing the total number of surprise points relative to the number with a perfectly balanced league. Groot and Groot (2003) rather misleadingly call the number with a perfectly balanced league the 'maximum'. In fact the theoretical maximum is the number with a perfectly balanced league multiplied by the number of times the teams play each other. Formally, the Surprise index (S_L) is calculated as:

$$S_L = \frac{P}{B_L}$$

Where:
$P =$ Total surprise points for the league
$P = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(R_{ij} + R_{ji}\right) (j - i)$
$N =$ Number of teams (i) in the league

$B_L =$ Number of surprise points with a perfectly balanced league
$B_L = \frac{N\,(N+1)\,(N-1)}{3}$

And
Team i ends the season in a higher position in the league table (i.e. with a lower rank number) than team $(i < j)$
$R_{ij} =$ The result of a game with team i at home against team j (and vice versa)
The 'surprise' outcomes give 'scores' of:
$R_{ij},\ R_{ji} = 2$ if the game results in a win for team j
$R_{ij},\ R_{ji} = 1$ if the game results in a draw
$R_{ij},\ R_{ji} = 0$ if the game results in a win for team i

A larger value for S_L denotes a greater level of competitive balance. The maximum value for the index is one, corresponding to a perfectly balance league. This would occur, for example, if every team played every other team twice and they each won one of those games. The minimum value is zero, corresponding to the case of maximum imbalance (i.e. with no surprise results).

An advantage of this measure, compared to other measures of concentration, is that it takes account of more information. It could be argued that requiring more information it is a disadvantage. However, Groot and Groot (2003) point out that the additional information is readily obtainable.

They also note that a further advantage of this measure is that it can be used to compare leagues of different sizes. This facilitates comparison of leagues over time, comparison between leagues in different countries and even comparison between leagues for different sports. However, the correspondence between the surprise scores and the points actually awarded is not maintained if the scoring system of the league changes in the comparative period.

Groot and Groot (2003) applied the index to the results of games in the top league of professional football in France from 1945 to 2002 (ignoring the change in points system used by the league from the mid 1990's onwards). They conclude that the level of balance is slowly, but unmistakably, decreasing.Groot and Groot (2003) also compared the results given by the Surprise index with results given by the normalised standard deviation measure and a normalised three club concentration ratio (where the normalisation produces a measure with a range of one, corresponding to perfect balance, to zero, corresponding to maximum imbalance). They found that all three measures convey almost exactly the same pattern and note that "from the perspective of parsimony the concentration ratio is unambiguously the best indicator of CB [competitive balance]." (ibid p. 7)

B) Measures of dominance

These measures focus on the aspect of competitive balance related to the relative performance of particular teams or groups of teams over a number of seasons rather than on closeness of a league competition. Vrooman (2006, p. 351) states that "[a]s it relates to the effects of free-agent talent acquisition, the most important aspect of competitive balance or dominance concerns season-to-season continuity. If the league is competitively balanced in this respect, then dynasties and doormats are the exception, and the rarely experienced season-to-season reversals are the rule."

Five types of measure that have been proposed in the literature are presented below.

1. Descriptive statistics (counts, percentages, frequency, averages etc.)

There is a wide range of possible descriptive statistics for the dominance of a sports league. Six are presented below.

(i) Number (percentage etc.) of league titles per team

The number of times the same team has won the title is a simple measure of dominance. Rottenberg (1956), in one of the earliest papers in the sports economics literature, stated that "[a] simple test [of whether the reserve rule has been successful] is one which counts the number of times each team has won its league pennant." (p. 247) Others to use this measure include Scully (1989, Table 4.1, p. 86), for the number of league championships ("pennants") won by each team in each league of MLB from 1901 to 1987, and Syzmanski and Kuypers (1999, Table 7.1, p. 258) for English football league championships, 1946/47 to 1997/98.

Noll (1991) considers professional basketball in the US and observes that two teams (the Celtics and the Lakers) have dominated the forty-year history of the league, winning two-thirds of the league's championships.

Szymanski and Kuypers (1999) consider the top football (soccer) leagues in five European countries over the fifty-two seasons from 1946/47 to 1997/98. They find that two clubs have dominated the championships in England (Liverpool and Manchester United, 45 per cent of championships), Scotland (Rangers and Celtic, 76 per cent), Spain (Real Madrid and Barcelona, 61 per cent), and the Netherlands (Ajax and PSV, 68 per cent) whilst in Italy three clubs (Juventus, A C Milan and Inter Milan, 69 per cent) dominate all other clubs.

Michie and Oughton (2004) consider the English Premier League for football and note that a simple indicator of domination is reflected in the fact that since the formation of the Premier League 12 years ago, the title has been won by one club (Manchester United) in 8 out of 12 seasons and by just two clubs (Arsenal and Manchester United) in 11 of the 12 seasons. They also note that there have been other periods when the top flight has been dominated by one club, such as Liverpool's run in the 1970s and 1980s when the club finished 1st or 2nd in 14 out of 15 seasons between 1973 and 1987, winning the league in 10 of those seasons.

Scully (1989) points out an issue with titles as a measure of quality (or dominance) in the specific case of the MLB since 1969 when divisional play-offs were established to determine the champion which is more widely applicable to competitions with a league format followed by a knock-out format to decide the title winner. The team that wins the title may not have had the best (i.e. most dominant) record in the league stage.

(ii) *Consecutive title wins*

Syzmanski and Kuypers (1999) identify the number of consecutive years that the same team has won the title as an indication of dominance. They note that "Scotland's top division has been characterised by long successive periods of domination, with both Glasgow clubs having won the title for nine consecutive years (Rangers 1988-89 to 1996-97 and Celtic 1965-66 to 1973-74)." (ibid p. 258)

A more informative version of this measure is used by Lenten (2009) who notes the length and frequency of consecutive title wins of teams. He calculates a weighted sum (given by the number of consecutive title wins per length multiplied by their individual lengths) for a league history. A higher number indicates more team dominance and less competitive balance.

Lenten (2009) also presents an analogous measure based on the length and frequency of non-consecutive title wins by teams. He calculates a weighted sum (given by the number of non-consecutive title wins per length of non-consecutive winners multiplied by the individual lengths) for a league history. A higher number indicates less team dominance and more competitive balance. He applies both of these measures to the Australian Football League and National Rugby League in Australia from their inception to 2006.

(iii) <u>'Lifetime' achievement of teams (A)</u>

This is a measure of the average win percentage of a team over its 'lifetime' in the league which has the useful advantage that it can be compared with other teams in the league or teams from other leagues. The measure is the number of 'idealised' standard deviations (for the average number of games played per season and on the assumption of equal playing strengths) that the 'lifetime' win percentage deviates from the league average win percentage (0.5). The measure (A) for a team (i) is given by:

$$A_i = \frac{\bar{w}_i - 0.5}{\bar{\sigma}_{Ii}}$$

Where:
$\bar{w}_i =$ 'Lifetime' win percentage of team i, given by:
$l_i = \frac{\sum_{t=1}^{T} w_{i,t}}{T_i}$
$w_{i,t} =$ Win percentage of team i for year t in the league
$T_i =$ Total number of years in the league for team (i)

It is assumed that T_i is sufficiently large that the sample mean will be distributed approximately as the normal distribution with the standard deviation of:
$\bar{\sigma}_{Ii} =$ The (average over T_i years) 'idealised' standard deviation for team (i), on the assumption of equal playing strengths, given by:
$\bar{\sigma}_{Ii} = \frac{0.5}{\sqrt{\bar{G}_i}\, x \sqrt{T_i}}$
$\bar{G}_i =$ Average number of games of team (i) per season in league
$T_i =$ Total number of years in the league for team (i)

A higher (lower) value for A_i suggests that team i has exhibited a higher (lower) level of dominance, i.e. "over achieved" ("under achieved") relative to a team with a lower value.

This measure is not dependent on values for other teams, and hence not dependent on league size or composition) and allows for differences in number of games per season. It is therefore directly applicable to 'open' leagues.

See Quirk and Fort (1992, Tables 7.6 through to 7.11, pp. 264-269) for "over achievers" and "under achievers" in all the major US sports leagues, AL, NL, NBA, NFL and NHL for teams with at least 10 years in the league from 1901 to 1990.

(iv) <u>Number of different title winners</u>

A league with fewer different title winners over a period suggests that they have a higher level of dominance within the league. Buzzacchi, Szymanski and Valletti (2003, p. 174) "… conjecture that fans care about balance in the sense that they want a reasonable prospect that the identity of the winners will change from time to time (although they may also care about the variance of success among the teams within a season)."

Buzzacchi, Szymanski and Valletti (2003) reports the number of teams that had the highest win percentages, in the regular season of the MLB, NFL and NHL, and the number of teams that won the league championships in soccer in England, Italy and Belgium over periods from 1950 to 1999.

Syzmanski and Kuypers (1999, Table 7.2, p. 258) also reports this measure for English football, from 1946/47 to 1997/98. They show that England had sixteen different champions in this period whilst Scotland had eight in the same period.

(v) Number of 'top' teams

A broader measure of dominance than title winners includes teams that are 'close' to the top of the league. There are many versions of this measure in the literature. Examples include:

- Borland (1987) uses the number of different teams in the finals of Australian Rules football in the past three seasons, divided by the number of finals berths available.
- Eckard (2001a) includes teams that have finished in the top four of the league over five year periods and express the number as a percentage of the number of teams in the league. He applies the measure to the US MLB Leagues from 1975 to 1999).
- Ross and Lucke (1997) considers the number of times a team finished within five or fewer places of the title and compared this to the number of times this would be expected if the league were perfectly competitive and every team had the same incidence.
- Buzzacchi, Szymanski and Valletti (2003) reports the number of teams that entered the top 5 ranks of the MLB, NFL and NHL, (all on the basis of the regular season, 1951 to 2000) and in soccer in England (1951 to 2000), Italy (1951 to 2000) and Belgium (1953 to 2000).

Scully (1989) relates the actual number of league titles a team has won over a period to the number they should have won if all teams had equal strength and the winning of games was determined by chance. The measure is calculated for a specific period by taking the difference between the actual number of title wins for a team and the expected number of titles on the assumption that teams were of equal playing strength during the period. The expected number for each team for each season is the reciprocal of the total number of teams in the league. This is summed for each year that the team is in the league. A positive difference suggests a level of dominance by a team in the period. See, for example, Scully (1989, Table 4.1, p. 86) for MLB teams from 1901 to 1987.

Eckard (1988) addresses the hypothesis that regulations for the major football conferences of the National Collegiate Athletic Association (NCCA) reduce competitive balance. Although the sporting structure in this case does not produce a single overall title winner, national rankings (published sports writers' Top 10 and Top 20 lists) provide competition for the mythical national championship and provide data for two measures as dominant teams will recur on these lists.

- Appearances on list of top teams
  This measure is a simple count of the number of different schools making appearances in the sports writers Top 10, or Top 20, lists of teams in a given period. A larger (smaller) number of teams implies more (less) competitive balance in the conferences.

- Re-entry on list of top teams
  This measure is the average, over a number of years, of the number of teams that appear on the annual sports writers' Top 10, or Top 20, list which do *not* appear in the list for the previous five years. A larger (smaller) average number of teams implies more (less) re-entry and hence more (less) competitive balance.

Goosens (2006) proposes a measure of the number of teams finishing in the top three in the league in consecutive five year periods. She recognises that the choice of the number of top positions is arbitrary but argues that three positions is suited to European national football leagues because in most European countries there are two or three teams that are commonly considered to be dominant. The period is based on an assumption that spectators would have this timeframe in mind when they consider the dominance of teams. She also recognises that more research is necessary to validate this assumption. Nevertheless, she concludes that "[f]or measuring dominance of teams in European football we believe this measure is one of the best." (ibid p. 94)

(vi) *Identity of 'top' teams*

Curren, Jennings and Sedgwick (2009) are interested in the identity of the dominant teams as this "could help explain why changes in competitive balance occur". (ibid p. 1738). They formulate a "Top 4 Index" by counting the number of occasions that each team finished a league season in the top four places, summing the incidence of the four teams with the most occurrences and expressing the total as a proportion of the total number of available places over the period of the measure. They calculate values or the top league of professional football in England from the 1948/49 to 2007/08 seasons (inclusive) and for ten year intervals.

(vii) *Frequency of failure to win a league title*

Noll (1991) includes an approach the issue of dominance based on "the extent to which any team remains a doormat for a long period." (ibid p.40) He presents a measure which is the difference between the theoretical probability of a team, in an equally balance league, failing to win a championship and the actual frequency over a number of seasons.

The actual frequency is calculated by first taking the difference between the average number of teams in the league during the period and the number of teams that had won at least one championship and then dividing by the average number of teams during the period. See Noll (1991, Table 1.7, p. 42) for the NBA, 1951-1989.

2. Time series association

These measures all require comparable data for two (usually consecutive) periods. The first consequence of this is that it is problematic to apply the measures to leagues with differences in the specific teams between periods, such as with 'open' leagues or leagues that change in size. The second consequence is that with data for T season, the measure only produces a value for T-1 seasons. Four approaches to measure dominance through time series association are presented below.

(i) *<u>Correlation coefficient</u>* (r)

The correlation coefficient is a standard statistical measure. Applied to a sports league, it measures the extent to which the teams fare relative to either other teams (by rank) or to their own performance (by win percentage) in different seasons. In this case, the measure (r) is calculated as:

$$r = \frac{\sum_{i=1}^{N}(x_i\ y_i) - \bar{X}\sum_{i=1}^{N} y_i}{\sqrt{[\sum_{i=1}^{N}(x_i{}^2) - \bar{X}\sum_{i=1}^{N} x_i]\,[\sum_{i=1}^{N}(y_i{}^2) - \bar{Y}\sum_{i=1}^{N} y_i]}}$$

Where:
$N =$ Number of teams (i) in the league
$\bar{X}, \bar{Y}$ denote the average values for the periods x and y respectively

For the rank correlation coefficient
$x_i =$ Position in league of team i in period x
$y_i =$ Position in league of team i in period y

For the correlation of win percentages
$x_i =$ Win percentage of team i in period x
$y_i =$ Win percentage of team i in period y

The measure has a range of +1 to -1 in the limiting cases with +1 corresponding to perfect imbalance. A value of -1 would indicate a perfect reversal of balance. Note that a zero value does not necessarily imply 'no relation'. It implies 'no linear relation'.

This is a strong measure of dominance in the sense that it captures the extent to which all of the teams in the league replicate performance. The dominance of a subset of teams could be masked by the differing performances of the other teams.

See, for example, Butler (1995), who applies the measure to team's win percentage between consecutive years for both the MLB leagues from 1947 to 1991.

Daly and Moore (1981) apply the measure to the ranking of teams. In this case they ranked the eight original franchises in each of the MLB leagues and calculated the correlation coefficient for pairs of periods from 1955 to 1973. As this measure requires the same teams to be in the league in both periods it is not directly applicable to leagues that change in composition (such as 'open' leagues). However, the number of teams in the MLB changed during this period and so Daly

and Moore (1981) used the measure with ranking based only on the win percentages in inter-team play for the eight original franchises over a number of years.

(ii) *Adjusted churn* (AC)

Mizak, Neral and Stair (2007) propose a simple measure of time series association which they call the Adjusted Churn. This is a measure of competitive balance based on the absolute difference in standings (league rank) for teams in consecutive seasons. The measure (AC) for a season (t) is calculated as follows:

$$AC_t = \frac{C_t}{C_{t\,max}}$$

Where:
$C_t =$ Churn
$C_t = \frac{\sum_{i=1}^{N} |k_{i,t} - k_{i,t-1}|}{N}$
$C_{t\,max} =$ The maximum possible value of C_t (for the value of N)
$k_{i,t} =$ Rank of team i in season t
$N =$ Number of teams (i) in the league

Since the calculated value of churn is divided by the maximum value, the measure produces an 'adjusted' churn maximum value of one, when there is the maximum possible change in league standings, implying a balanced league. It has a lower limit of zero, with an ossified league when there is no change in league standings, implying minimal competitive balance.

Mizak, Neral and Stair (2007) calculate values for the Adjusted Churn measure for the US MLB American League East Division (1995-2007) and average values for decades of both the American League and National League from 1910-19 to 1990-99 and 2000-07.

(iii) *Autoregressive win percentage*

Vrooman (1996) estimates the season-to-season continuity of team quality with an autoregressive model of team win percentages with a one season lag (and binary variables for large-market clubs and teams playing in new stadiums). He applies the model to periods of the MLB between 1970 and 1993 and compares autoregressive coefficients for winning percentages (as measures of continuity) and the constant terms (as measures of competitive balance).

(iv) *'Top 4' recurrence*

Curren, Jennings and Sedgwick (2009) propose a measure of dominance based on the percentage of teams finishing in the top four positions in the league repeating this achievement the following season. They call the measure 'PROB'.

The maximum value (100%) is attained if the top four teams remain the top four teams throughout the measured period. The minimum value (0%) is attained if no team remains in the top four in the season immediately following a 'top four' finish.

The average value (50%) indicates that teams have an even chance of being in the top four, given that they finished in the top four in the previous season.

They calculate the measure for each season of the top league of professional football in England from 1949/50 to 2007/08 (inclusive) and show a plot of the moving average for five, ten, fifteen and twenty year periods.

3. Herfindahl-Hirschman Index (HHI) – for teams over a number of seasons ($\boldsymbol{HHI^D}$)

The Herfindahl-Hirschman Index (HHI) used to measure concentration in an industry is presented with the measures of concentration (above). An adapted HHI measure, proposed to measure the competitive balance of sports leagues in terms of the *dominance of teams over a number of seasons* has the general form given by:

$$HHI^D = \sum_{i=1}^{N} x_i^2$$

Where:
$x_i =$ Share of { …. } of team i over a selected number of seasons
$N =$ Number of teams (i) in the league

It has limits corresponding to the industry measure.

However, the restrictions to the HHI measure imposed by a league structure do not apply to this multi-season measure. Furthermore, the values are independent of the number of games played.

The general form has been applied to produce different specific measures. Three of the HHI, proposed to measure the competitive balance of sports leagues in terms of the dominance of teams over a number of seasons, are shown below.

(i) <u>Titles won</u>

This measure is a function of the number of teams that have won a title in the period and the relative share of titles between them. In this case:

$x_i =$ Share of titles won by team i over a selected number of seasons

An increase in the number of title winners will ('ceteris paribus') reduce the value given by this measure. The minimum value, 1/N, corresponds to perfect competitive balance where, in a league with N teams, each team would win the title, on average, every N seasons. The maximum value, one, would occur if the same team wins the title every season.

Humphreys (2002) uses this measure for a comparison of measures with the standard deviation of win percentages measure and the Competitive Balance Ratio (CBR). He calculates values for the American League and National League of MLB in decades from the 1900s to the 1990s and for the sub-periods 1990-1994 and 1995-1999. He concludes that "variation in the CBR measure over time does a better job explaining observed variation in attendance in MLB than the other two

alternative measures." (ibid p. 147). This measure is applicable to 'open' leagues. It could also be applied to measure dominance in competitions such as the UEFA Champions League.

(ii) *Number of top (or bottom) positions*

Including a larger number of 'top' positions (rather than only the title winners) provides a broader measure of dominance. Eckard (2001a) proposes to extend the measure to include the top four positions. In this case:

$x_i =$ Share of appearances in top four positions by team i over a selected number of seasons

An increase in the number of positions list will ('ceteris paribus') reduce the value given by this measure. The measure has a minimum value of $\left(\frac{1}{N}\right)$ which corresponds to perfect competitive balance where, each team would appear on a list, on average, every $\left(\frac{N}{4}\right)$ seasons. The maximum value $\left(\frac{1}{4}\right)$ would occur if the same teams appeared on the list every season.

This measure is applicable to 'open' leagues. It could also be applied to measure dominance in competitions such as the UEFA Champions League.

Eckard (2001a) also proposes a converse measure of dominance by replacing the top four positions with the bottom four positions. In this case:

$x_i =$ Share of appearances in bottom four positions by team i over a selected number of seasons

The limits for this measure are the same as for the top four measure. Eckard (2001a) calculates values for the measure for both the top four and bottom four position for periods of five seasons of both the American League and National League of MLB (1975-1999). His data excludes new expansion teams from the metric for the five year period in which they joined the league. Milwaukee switched from the AL to the NL in 1998 and, as an established team, is counted in both leagues for the five year period in which they switched.

(iii) *'Virtual' league appearances*

Eckard (1998) shows that this measure can be used to measure the dominance of teams even when they are not involved in direct sporting competition! He considers the sport of college football in the US where teams compete in regional leagues. In lieu of an integrated sporting competition, sports writers assess the performance of all the teams on an annual basis and produce a list that ranks the 'top' teams. This forms a 'virtual' league. Citation on a list indicates a dominant team. Recurring citations indicate greater dominance. In this case:

$x_i =$ Share of appearances on the list by team i over a selected number of seasons

An increase in the number of teams appearing on the list will ('ceteris paribus') reduce the value given by this measure. The list includes a specific number of teams (S) and $S < N$. The measure has a minimum value of $\left(\frac{1}{N}\right)$ which corresponds to perfect competitive balance where, each team would appear on a list, on average, every $\left(\frac{N}{S}\right)$ seasons. The maximum value $\left(\frac{1}{S}\right)$ would occur if the same teams appeared on the list every season.

Eckard (1998) uses this measure for a test of the hypothesis that NCCA regulations (effectively introduced in 1952) reduced competitive balance. He calculates values for the measure for the Top 10 and Top 20 lists for teams in the major football conferences of the NCCA. He compares the metric for the pre-enforcement years (1924-1951 and 1936-1952 for the Top 10 and Top 20 lists respectively) with the respective lists for the period 1957-1984 when the regulations were in effect. He finds that appearances in national rankings are more concentrated among fewer teams after enforcement began and concludes that competitive balance did decline after NCAA enforcement began.

The measure is independent of number of teams and could be applied to 'open' leagues and competitions such as the UEFA Champions League, where 'qualification' could correspond to the sports writers' list.

4. Lorenz curve and Gini coefficient – for leagues over a number of seasons $(\boldsymbol{G^D})$

The Lorenz curve and Gini coefficient were introduced as measures of concentration. They have also been used in the sports economics literature to measure of the extent of inequality in the winning of league championships. In this case, the convention in the literature has been to show the Lorenz curve above the line of perfect equality and so the teams are shown in decreasing order of success (along the x-axis).

To calculate the Gini coefficient (as a positive number) the teams are ranked in increasing order of success (k). The Gini-coefficient measure of domination (G^D) over a given number of seasons, is then approximated by the statistical formula:

$$G^D = \sum_{i=1}^{n} (x_i y_{i+1} - x_{i+1} y_i)$$

Where:
$y_i =$ Cumulative percentage titles won by teams to team i, in rank order with $k_1 \leq k_2 \leq \cdots \leq k_n$
$x_i =$ Cumulative percentage number of teams to team i
$N =$ Number of teams included in the period

This measure has a minimum value of zero, when the Lorenz curve follows the equality line and a maximum value of one minus the reciprocal of the number of teams included in the measure.

With the application of this measure to the dominance of teams over a number of seasons the limitation of the league structure, which applied to the statistic as a measure of concentration in a league season (discussed above), does not apply. However, in this case the measure is dependent on the number of teams included in the period. This dependency compromises the use of the measure to compare between periods or between leagues with different numbers of teams included.

A further issue with this measure arises in the situation when it is applied to periods in which the contestants for the title changed over time (as, for example, with 'open' leagues or the addition or departure of franchises) as all teams will not have the same number of opportunities to win the title over the period of the measure. Both Quirk and Fort (1992) and Goosens (2006) only include all teams that have been in the league for at least ten years during the period covered by the measure. They also weight the teams included with the number of seasons they are included in the competition.

Szymanski and Kuypers (1999) make the simplifying assumption that twenty two (non-specific) teams have equal chances of winning the title each season although they recognise that, in the case of European football leagues, with the system of promotion and relegation, this is only an approximation. Quirk and Fort (1992, Figure 7.6 p. 260) show Lorenz curves for teams in all the major US sports leagues, AL, NL, NBA, NFL and NHL with at least ten years in the league from 1901 to 1990 and report the Gini coefficients for each of the leagues.

Szymanski and Kuypers (1999, Figure 7.1, p. 259) show the Lorenz curve to compare the long-run domination between the top leagues of (association) football in England, Spain, Italy, Netherlands and Scotland over the seasons from 1946/47 to 1997/98. They find that the Netherlands exhibits the greatest concentration (i.e. most domination and least competitive balance) followed by Spain, Italy and Scotland with England exhibiting the least concentration of championships in clubs.

Goosens (2006) applies this measure to teams in eleven European national football leagues with at least ten years in the league over the forty two year period from the 1963/64 season (when the Bundersliga was formed in Germany) to the 2004/05 season. She shows the Lorenz curves for each country (ibid Figure 5, pp. 115-116) and makes direct comparisons between the leagues with the same number of (eligible) teams. She also reports the Gini (weighted) coefficients for each national league (ibid Figure 6, p. 117).

5. Surprise index for teams $(\boldsymbol{D_i})$

Groot and Groot (2003) argue that the most attractive property of the Surprise index (measure of concentration presented above) is that it can be decomposed into an index of competitive balance *per team* and used to assess the league size which corresponds to the most balanced league. The merits (or otherwise) of this approach are beyond the scope of this review.

Nevertheless, the team measure proposed by Groot and Groot (2003) to assess dominance at (or near) the top of a league could provide the basis for a measure of the extent to which the success of the teams at (or near) the top of the league is the

result of their ability to ‘surprise’ (or dominate) their close rivals relative to their ability not to ‘drop’ points (and consistently dominate) against the less successful teams.

To identify dominant teams at the top of the league, a ‘surprise’ is deemed to have occurred for a team that is only able to draw or loses to a team that finished the season in a lower position. For those games, the higher positioned ‘surprise’ team is allocated a score of one for a draw and two for a loss. The score for each game is weighted, to represent the degree of ‘surprise’, by a factor given by the difference in final league position (i.e. rank order) of the two teams involved in the game to produce a number of ‘surprise points’ (‘dropped’) per team per game. This accentuates the difference between teams.

The index for a *team* is formed by expressing the total number of surprise points (‘dropped’) for the team relative to the number that finished the season in that position, with a perfectly balanced league (i.e. if the team wins one game and loses the other game against every other team in the league). The index measure of points ‘dropped’ by team i (D_i) is given by:

$$D_i = \frac{P_i^D}{B_i}$$

Where:
P_i^D = Number of surprise points ‘dropped’ by team i
$P_i^D = \sum_{j=i+1}^{N} (R_{ij} + R_{ji})(j - i)$
B_i = Number of surprise points that would be ‘dropped’ by team i with a perfectly balanced league
$B_i = (N - i)(N - i + 1)$
N = Number of teams (i) in the league

And
Team i ends the season in a higher position in the league table (i.e. with a lower rank number) than team $(i < j)$
R_{ij} = The result of a game with team i at home against team j (and vice versa)
The ‘surprise’ outcomes give ‘scores’ of:
$R_{ij},\ R_{ji}$ = 2 if the game results in a win for team j
$R_{ij},\ R_{ji}$ = 1 if the game results in a draw
$R_{ij},\ R_{ji} = 0$ if the game results in a win for team i

The value of this statistic as a measure of team dominance within a season would be, for example, in comparisons between teams that finished the season with a comparable number of points awarded by the league competition. A higher value for the team measure (D_i) would indicate that the team had achieved its success with relatively greater dominance over the higher ranked teams. A lower value for the team measure (D_i) would indicate that the team had achieved its success with relatively greater dominance over the lower ranked teams. This approach was not developed by Groot and Groot (2003).

C) Measures combining concentration and dominance

A single measure of competitive balance has the advantages that:

- It provides a 'high level' quantified assessment of the level of competitive balance
- It can provide a single explanatory variable for competitive balance in, for example, estimates of demand.

Four approaches to providing a combined measure are presented below.

1. Distribution of 'lifetime' win percentages in a league $(\boldsymbol{L})$

This is a measure of the competitive balance of a league over its 'lifetime'. It is a measure of concentration in a league over its 'lifetime' derived from a measure of the dominance of its teams over their 'lifetime' in the league. The measure of the dominance of a team is its 'Lifetime' Achievement statistic, A_i, (see Measures of dominance, number 12 above). The measure of distribution is the percentage of teams with a 'lifetime' win percentage within a specified number of 'idealised' standard deviations from the population mean (0.5). The league measure (L) is given by:

$$L\ (\%) = \frac{N^L}{N}\ x\ 100$$

Where:

$N^L =$ Number of teams with a value for $\bar{w}_i$ such that $(0.5 - Zx\bar{\sigma}_{Ii}) < \bar{w}_i < (0.5 + Zx\bar{\sigma}_{Ii})$

$N =$ Number of teams (i) in the league

$Z =$ Specified number of 'idealised' standard deviations

$\bar{w}_i =$ 'Lifetime' win percentage of team i, given by:

$$l_i = \frac{\sum_{t=1}^{T} w_{i,t}}{T_i}$$

$w_{i,t} =$ Win percentage of team i for year t in the league

$T_i =$ Total number of years in the league for team (i)

It is assumed that T_i is sufficiently large that the sample mean will be distributed approximately as the normal distribution with the standard deviation of:

$\bar{\sigma}_{Ii} =$ The (average over T_i years) 'idealised' standard deviation for team (i), on the assumption of equal playing strengths, given by:

$$\bar{\sigma}_{Ii} = \frac{0.5}{\sqrt{\bar{G}_i}\ x\ \sqrt{T_i}}$$

$\bar{G}_i =$ Average number of games of team (i) per season in league

$T_i =$ Total number of years in the league for team (i)

A larger (smaller) proportion of teams close to the league mean value corresponds to a more (less) balanced league.

See, for example, Quirk and Fort (1992, Table 7.12 p. 269) with three standard deviation limits for all the major US sports leagues, AL, NL, NBA, NFL and NHL for teams with at least 10 years in the league from 1901 to 1990.

2. ANOVA-type measures

These measures are based on the same teams competing over a number of seasons. This makes them suitable for 'closed' leagues with constant size but not directly applicable when the number of teams changes or for 'open' leagues. Two measures proposed in the literature, and the relationship between them, are presented below.

(i) Variance decomposition (% Time)

Eckard (1998) notes that a seller's cartel, where the firms agree to collude to stop attracting customers from each other, has a main objective of increasing market share stability. He argues that the regulation of college sport in the US (by the National Collegiate Athletic Association, NCCA), that restricts player recruiting, creates a buyer's cartel for players and that the equivalent measure is of the stability of performance of the teams.

He tests the hypothesis that (the effective enforcement of) this regulation has had the effect of reducing competitive balance as it stabilises the league positions of teams and reinforces an advantage for already successful schools that offer players greater "legal" non-pecuniary benefits. Or, as expressed in Eckard (2001a, p. 214): "Less balance within a league implies that teams with good, middling or poor records tend to repeat them year after year."

He assesses competitive balance by separating the pooled variance of win percentages into the variation in a team's performance between seasons and the variation between teams. This is shown as:

$\sigma^2{}_L = \sigma^2{}_{time} + \sigma^2{}_{cum}$

Where:

$\sigma^2{}_L =$ Total variance of pooled win percentage data for all teams in a league over a defined period

$$\sigma^2{}_L = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(w_{i,t} - 0.5\right)^2}{T\ x\ N}$$

$\sigma^2{}_{time} =$ Sum of the variances of each team over a defined period

$$\sigma^2{}_{time} = \frac{\sum_{t=1}^{T}\left(w_{i,t} - \overline{w}_i\right)^2}{T}$$

$\sigma^2{}_{cum} =$ Variance of the sum of the win percentages for each team over a defined period

$$\sigma^2{}_{cum} = \frac{\sum_{i=1}^{N}(\overline{w}_i - 0.5)^2}{N}$$

And

$w_{i,t} =$ Win percentage of team i in season t

$\overline{w}_i =$ Average win percentage of team i over T seasons

$N =$ Number of teams (i) in the league

$T =$ Number of seasons (t)

His hypothesis is that the regulation has the effect of reducing ${\sigma^2}_{time}$, as league positions ossify, and increasing ${\sigma^2}_{cum}$, with the advantage for already successful schools that offer players greater "legal" non-pecuniary benefits.

Eckard (1998) compares data (where possible) for NCCA conference members for the period from 1927 to 1951, before "effective NCCA enforcement began in 1952" (ibid p. 356), with the period from 1957 to 1981, following "a five year transition period" (ibid p. 356) from 1952 to 1956. He finds that:
"All five conference time variances decline after 1952. … Similarly, all five cumulative variances increase in the post-enforcement period." (ibid pp. 360-361) The mean differences between pre- and post-enforcement periods are statistically significant at the one per cent level. This supports the hypothesis that the (effectively enforced) regulation had the effect of reducing competitive balance in NCCA leagues.

Eckard (2001a) recognises that whilst for a given total variation (${\sigma^2}_L$), a decrease (increase) in the cumulative variance (${\sigma^2}_{cum}$) and increase (decrease) in the time variance (${\sigma^2}_{time}$) indicate that competitive balance has increased (decreased), if ${\sigma^2}_L$ changes, the interpretation is more complex. Consequently he proposes an additional measure ($\%\,Time$) which is calculated as follows:

$$\%\,Time = \frac{{\sigma^2}_{time}}{{\sigma^2}_L}\ x\ 100$$

As $\%\,Time$ increases, the time variance is a larger proportion of the total and the cumulative variance is smaller, indicating a greater competitive balance. This measure has a maximum value of one and a minimum value of zero.

Eckard (2001a) points out that the applicability of these measures is restricted to leagues with the same teams throughout the period which also play an equal number of games against each other team.

Humphreys (2002) criticises the variance decomposition by Eckard (1998) and argues that, since a fundamental rule of variances is that $var(X+Y) = var(X) + var(Y) + 2cov(X,Y)$ [Note: I have corrected the sign on the covariance term which is incorrectly shown as negative], the measure proposed by Eckard only holds when: $cov(X,Y) = 0$ and this is "… a situation that rarely occurs in actual sports leagues." (ibid p. 135) Eckard (2003) shows that this criticism by Humphreys (2002) is invalid and that the cross-product term in the ANOVA sum-of-squares expansion, which in other contexts corresponds to the covariance, is zero. Humphreys (2003) accepts that his criticism was wrong.

A small additional point to note with Humphreys (2002) is that he erroneously reports the values of the Eckard decomposition of two hypothetical leagues as ${\sigma^2}_{cum} = 0.35$ for League 1 and ${\sigma^2}_{time} = 0.35$ for League 2 (p.135). In fact these are the corresponding values for the standard deviations. The variance values are both 0.125.

(ii) <u>Competitive Balance Ratio</u> (CBR)

Humphreys (2002) proposes a single measure of competitive balance that incorporates both concentration and dominance to use as a variable in an empirical demand function to assess the effect of competitive balance on attendance at major league baseball. His measure, the Competitive Balance Ratio (CBR), is given by:

$$CBR = \frac{\bar{\sigma}_T}{\bar{\sigma}_N}$$

Where:
For the numerator
$\bar{\sigma}_T =$ The "average time variation" given by:
$\bar{\sigma}_T = \frac{\sum_i \sigma_{T,i}}{N}$ The "league wide measure" from:
$\sigma_{T,i} = \sqrt{\frac{\sum_t (w_{i,t} - \bar{w}_i)^2}{T}}$ The "within team variation"

For the denominator
$\bar{\sigma}_N =$ The "average variation across seasons"
$\bar{\sigma}_N = \frac{\sum_t \sigma_{N,t}}{T}$
$\sigma_{N,t} = \sqrt{\frac{\sum_t (w_{i,t} - 0.5)^2}{N}}$ The "within season variation"
(assuming all teams play the same number of games)

And
$\sigma_{T,i} =$ Standard deviation of the win percentage of team i over T seasons
$\sigma_{N,t} =$ Standard deviation of the win percentages of the N teams in the league in season t
$w_{i,t} =$ Win percentage of team i in season t
$\bar{w}_i =$ Average win percentage of team i over T seasons
$N =$ Number of teams (i) in the league
$T =$ Number of seasons (t)

Humphreys (2002) argues that the appealing properties of this measure are that it is easier to compare different time periods because it does not have to be compared to an idealised value that depends on the number of games played in each season and it has intuitively appealing upper and lower bounds of zero and one.

To compare the CBR measure with the standard deviation of win percentages (measure of concentration) and the HHI measure of titles won (measure of dominance), Humphreys (2002) calculates values for each measure for the American League and National League of MLB in decades from the 1900s to the 1990s, and for the sub-periods 1990-1994 and 1995-1999. He finds that "the CBR uncovers important distinctions between several periods that have similar σ_Ls and HHIs." (ibid p. 139) He notes that the CBR for the National League is much larger for the decade of the 1910s than for the 1920s whilst both of the other measures have similar values for the two decades. During the 1920's the same set of teams (Pittsburgh, St. Louis and New York) were consistently in the upper division and the

same set of teams (Boston, Brooklyn and Philadelphia) were consistently in the lower division. "The CBR captures this relative stratification in standings but the other two measures do not." (ibid pp. 141-142) He also finds that the "CBR reveals a change in the level of competitive balance in the second half of the 1990s relative to the first half of that decade that is not reflected by the standard deviation of winning percentage." (ibid p. 142)

Eckard (2003) shows that Humphreys' CBR measure is actually nothing more than a standard deviation version of Eckard's $\%\,Time$ measure (expressed as a decimal fraction).

3. Mobility gain function $(\boldsymbol{MGF_t})$

Lenten (2009) proposes a measure of competitive balance which takes the form of a mobility gain function. This measure is based on the categorisation of possible changes in team performance relative to the team's performance the previous season and to the average performance of the league. These categories are each weighted (by assumption) to model the impact on competitive balance applicable to each team for each season. The total 'gain' of all the teams in the league is averaged to produce the measure. Formally, the measure is given by:

$$MGF_t = \frac{\sum_{i-1}^{N} y_{i,t}}{N}$$

Where:
$y_{i,t} =$ Competitive gain function for team i in season t which has three possible values determined as follows:

Possible values for $y_{i,t}$		
(1) $y_{i,t} = 0$	(2) $y_{i,t} = \alpha - \frac{\left(w_{i,t} - 0.5\right)^2}{\alpha}$	(3) $y_{i,t} = \alpha$
if either: $w_{i,t-1} = 0.5$ or: $w_{i,t} \leq w_{i,t-1} < 0.5$ or: $0.5 < w_{i,t-1} \leq w_{i,t}$	if either: $w_{i,t-1} < w_{i,t} < 0.5$ or: $0.5 < w_{i,t} < w_{i,t-1}$	if either: $w_{i,t} \leq 0.5 < w_{i,t-1}$ or: $w_{i,t-1} < 0.5 \leq w_{i,t}$

And
$w_{i,t} =$ Ratio of games won to games played by team i in season t
$w_{i,t-1} =$ Ratio of games won to games played by team i in season $t-1$
$\alpha = \left|w_{i,t-1} - 0.5\right|$
$N =$ Number of teams (i) in the league

The gain function increases in value with a movement from season to season towards the average performance in the win percentages of teams in the league. This corresponds to an increase in competitive balance.

In the situation (1) scenarios the team has not got closer to competitive balance (i.e. it has not 'gained' competitive balance in the season t relative to season $t-1$). In the situation (2) scenarios the team got closer to competitive balance but remained either above or below the average. In the situation (3) scenarios the team moved from being either above or below average to being equal to or the other side of average. In these cases the 'gain' is limited to the difference between the team's performance in season $t-1$ and the average (i.e. its maximum possible gain).

The gain function specified includes a quadratic range for the situation (2) scenarios. Lenten (2009) also considers a linear gain function in this range and refers to the linearised mobility gain function as MGFL. This removes the assumption that the largest marginal increase in the gain function comes when a dominant team in one season becomes slightly less dominant in the following season.

Lenten (2009) applies the mobility gain function (and the MGFL measure) to the Australian Football League and National Rugby League in Australia from their inception to 2006. Lenten (2009) also uses the data set to compare the results for this measure (and the MGFL) with that of other measures. He selects the following measures for comparison:

- Ratio of standard deviation to 'idealised' standard deviation of win percentages (σ_R)
- Two concentration indexes (C3ICB and C5ICB)
- Index of dissimilarity
- Herfindahl-Hirschman Index of competitive balance (HICB)
- Gini coefficient
- Range

He finds that:

- The results of three measures, the standard deviation ratio, the index of dissimilarity and the HICB are strongly correlated.
- The standard deviation ratio is the best 'all purpose' measure (as it has the weakest correlation with any other measure)
- The MGF (and MGFL) measures pick up a very different set of competitive balance effects compared with all of the other measures. This should be expected as it is the only measure in the comparison which includes both a measure of concentration and a measure of dominance.

However, this measure requires comparable data for two consecutive periods. The first consequence of this is that it is problematic to apply the measures to leagues with differences in the specific teams between periods, such as with 'open' leagues or leagues that change in size. The second consequence is that with data for T season, the measure only produces a value for T-1 seasons.

4. Markov models

A Markov process is a stochastic state dependent model whereby the state in one period affects the outcome in the following period. Markov transition models are frequently used to model, for example, disease progression. This approach can be applied to the probability that a team's performance in one season depends on its performance in the previous season to measure the competitive balance of a sports league. Buzzacchi, Szymanski and Valletti (2003) call measures of competitive balance that take into account the mobility of teams in the ranks of the leagues measures of "dynamic competitive balance".

Two applications of Markov models to provide measures of competitive balance of sports leagues are presented below. The first involves statistical test of theoretical and actual transitional probabilities which allows for the testing of a wide range of hypothesis regarding competitive balance relating to strata of a league structure. The second provides a 'Gini type' single statistic measure of the competitive balance of a league system.

(i) *Transitional probability tests*

A simple Markov model for a sports league could distinguish between three possible states in a league, 'winners' (W), 'contenders' (C) and 'losers (L). The 'transitional probability' for a team is the probability that it will make the transition from its initial to a state in the next period. In this case the conditional probabilities for the specified states are denoted as follows:

Winners (W)	Contenders (C)	Losers (L)
P_{WW}	P_{CW}	P_{LW}
P_{WC}	P_{CC}	P_{LC}
P_{WL}	P_{CL}	P_{LL}

Where, for example, P_{CW} represents the probability that a 'contender' in one season will be a 'winner' in the next season.

In each season, teams in each state have the possibility of remaining in that state or of moving to either one of the other states. Hence, for example, $P_{WW} + P_{WC} + P_{WL} = 1$, as each winner must transition to become either a winner again or to become a contender or a loser.

In a perfectly balanced league, it can be argued that the transitional probabilities for a team are independent of the initial state and then every team in the league has the same transitional probability of becoming a winner or contender or loser. The probability of each transition would equal the number of positions in the state divided by the number of teams in the league. Over a number of seasons in which both the number of positions in each state and the number of teams in the league could vary, the 'balanced' transitional probabilities are given by:

$$P_{ij}{}^{B} = \frac{\sum_{t=1}^{T} \frac{S_j(t)}{N_t}}{T}$$

Where:

$P_{ij}{}^{B}$ = 'Balanced' transitional probability of a team moving from state i to state j

$S_j(t)$ = Number of positions in state j in season t

N_t = Number of teams in the league in season (t)

T = Number of seasons (t)

The actual transitional values are calculated, over a number of seasons, as the proportion of transitions from any one state to any other state, summed over each season and divided by the number of seasons. For example, a league with four positions for teams deemed to be 'winners' in which, in the first season one of the winners remained a winner and in the following season, three of the winners remained winners would have a value as follows:

$$P_{WW} = \frac{\frac{1}{4} + \frac{3}{4}}{2} = 0.5$$

More generally, over a number of seasons in which both the number of positions in each state and the number of teams in the league could vary the transitional probabilities are calculated as:

$$P_{ij}{}^{A} = \frac{\sum_{t=1}^{T} \frac{X_{ij}(t)}{S_j(t)}}{T}$$

Where:

$P_{ij}{}^{A}$ = Actual transitional probability of a team moving from state i to state j

$X_{ij}(t)$ = Number of teams moving from state i to state j in season t

$S_j(t)$ = Number of positions in state j in season t

T = Number of seasons (t)

Competitive balance within a season can be assessed by comparing actual values with 'balanced' transitional probabilities and tested statistically for difference. The actual values can also be used to assess competitive balance between periods.

An interesting facet of this measure is that it can be applied to sporting structures that do not produce a single overall title winner.

See, for example, Hadley, Ciecka and Krautmann (2005, Table 1, p. 383) for the transitional probabilities of MLB teams qualifying, and not qualifying, for postseason play for 1982-1993 and 1995-2003, which correspond to pre- and post-strike periods.

Hadley, Ciecka and Krautmann (2005) also use a similar process to calculate the probability mass function, conditional on current states, for the number of winning seasons during the following decade to infer how a team's current performance affects its future performances over some time period.

A significant limitation of the model employed by Hadley, Ciecka and Krautmann (2005) is the assumption that the transition probabilities are invariant over time. Koop (2004) relaxes this assumption and includes an ordered probit specification to allow for the possibility that the transition probabilities vary over time and over teams. This allows testing of additional hypothesis related to competitive advantage. He gives, as an example, questions like 'Is it difficult for a small market team to build a champion?'. He applies this model to the MLB from 1901 to 2000.

Buzzacchi, Szymanski and Valletti (2003, note 4, p. 169) point out that the method proposed by Koop (2004) is not naturally adapted to inter-league comparisons. Furthermore, the measure requires comparable data for two consecutive periods. The first consequence of this is that it is problematic to apply the measures to leagues with differences in the specific teams between periods, such as with 'open' leagues or leagues that change in size. The second consequence is that with data for T season, the measure only produces a value for T-1 seasons.

(ii) <u>Gini-type measure</u>

Buzzacchi, Szymanski and Valletti (2003) also use a stochastic Markov model but propose a single statistic measure of the competitive balance of a league system based on the number of teams ending a season at, or near, the top of a league compared to the theoretical number that would be expected if all teams had equal probability of winning each of their games. They propose a "Gini-type" measure $G(T^*)$ calculated by:

$$G(T^*) = \frac{\sum_{T=1}^{T^*} y^L(k,t) - \sum_{T=1}^{T^*} y_a^L(k,t)}{\sum_{T=1}^{T^*} y^L(k,t)}$$

Where:
$T^* =$ Number of seasons included in the metric $(t = 0, \dots, T)$
$y^L(k,t) =$ Theoretical number of teams appearing in rank k or higher in a given league $L = \{CL, OL\}$ over a period of T seasons
$y_a^L(k,t) =$ Actual number of teams appearing in rank k or higher in a given league $L = \{CL, OL\}$ over a period of T seasons

This measure has a lower limit of zero, corresponding to a perfectly balanced league (when the theoretical number is equal to the actual number). The maximum value occurs when the same teams finish every season in the top k positions and depends on the number of seasons included in the calculation. Buzzacchi, Szymanski and Valletti (2003) note that this measure depends on the starting year and that the longer the time series the less informative is the more recent data.

The measure is designed to measure the competitive balance of either a closed or open league.The calculation of the theoretical number depends on whether the same teams compete each season (a closed league) or the league has a system of relegation and promotion (an open league). For a closed league, the expected number of teams (y^{CL}) placed in the top k positions after T seasons is given by:

$$y^{CL}(k,T) = N - \frac{(N-k)^T}{N^{T-1}}$$

Where:
$N =$ Number of teams in the league

For an open league, the expected number of teams (y^{OL}) placed in the top k positions after T seasons is given by:

$y^{OL}(k,T) = \sum_{l=1}^{L} n_l w_l\,(k,T)$
Where:
$L =$ Number of leagues (l) ordered from $l = 1, \ldots, L$ such that 1 is the highest and L is the lowest
$n_l =$ Number of teams in league l
$w_l(k,T) =$ The probability that after T seasons, a team that started in league l in the initial period $(t = 0)$ has been placed, at least once, in the top k places of the top league and is given by:
$w_l(k,T) = 1 - \sum_{t=0}^{T} 1 - \frac{d_l(l,t)k}{n_l}$

And:
$d(l,t) =$ The probability that a team is in league l at time t (if the outcome of each league is random) given by:
$d(l,t) = d(l,t-1)\frac{n_l - R(l) - P(l)}{n_l} + d(l-1,t-1)\frac{R(l-1)}{n_{l-1}} + d(l+1,t-1)\frac{P(l+1)}{n_{l+1}}$
$P(l) =$ The number of promotions to adjacent league above league l
$R(l) =$ The number of relegations to adjacent league below league l

And it is assumed that:
$P(1) = 0$ i.e. No team is promoted from the top league
$R(L) = 0$ i.e. No team is relegated from the lowest league
$d(0,t) = d(L+1,t) = 0$

Buzzacchi, Szymanski and Valletti (2003) also propose an alternative to the measure $G(T^*)$. This measure $G(T^*)'$ is calculated by:

$$G(T^*)' = \frac{n_e(k,t)}{n(k,t)}$$

Where:
$n(k,t) =$ The equivalent dimension of a closed league with a constant structure that would have generated the same number of teams appearing in rank k over the same period as $y^L(k,t)$
$n_e(k,t) =$ The equivalent dimension of a closed league with a constant structure that would have generated the same number of teams appearing in rank k over the same period as $y_a^L(k,t)$

They note that this measure indicates competitive balance at the end of a period without concentrating on how a particular configuration is reached over time – contrary to the measure $G(T^*)$.

Conclusion

A wide range of statistical measures have been proposed in the literature to quantify the competitive balance of professional sports leagues. This reflects the types of summary statistics that may be useful (in general), differences in the underlying data sets due to differences in the league structures (in particular) and specific differences in context in which the term 'competitive balance' is being used.

This review has categorised the measures in the literature according to whether they are:

a) Measures of concentration
b) Measures of dominance
c) Measures combining concentration and dominance

It has presented details of the measures to facilitate their appropriate application and interpretation.

Bibliography

Berri, D. J., Schmidt, M. B. and Brook, S. L. (2007). The wages of wins: Taking measure of the many myths in modern sport. Stanford Business Books.

Borland, J. (1987). The demand for Australian Rules football. The Economic Record, September, pp. 220-230.

Butler, M. R. (1995). Competitive balance in major league baseball. The American Economist, Vol. 39, No. 2, pp. 46-52.

Buzzacchi, L., Szymanski, S., and Valletti, T. (2003). Equality of opportunity and equality of outcome: Open leagues, closed leagues and competitive balance. Journal of Industry, Competition and Trade, Vol. 3 (3), pp. 167-86.

Cain, L. P. and Haddock, D. D. (2006). Measuring parity: Tying into the idealized standard deviation. Journal of Sports Economics, Vol. 7, No. 3, pp. 330–338.

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, pp. 3-80.

Curran, J., Jennings, I. and Sedgwick, J. (2009). Competitive balance in the top level of English football, 1948-2008: an absent principle and a forgotten idea. International Journal of the History of Sport, 26, 1668-1680.

Daly, G. and Moore, W. J. (1981). Externalities, property rights and the allocation of resources in major league baseball. Economic Inquiry, 19: 1, pp. 77-95.

Depken, C. A. (1999). Free-Agency and the competitiveness of Major League Baseball. Review of Industrial Organization, Vol. 14 (3), pp. 205-217.

Dobson, S. M. and Goddard, J. A. (2011). The Economics of Football, 2nd edition, Cambridge University Press.

Dobson, S., Goddard, J. and Ramlogan, C. (2001). Revenue convergence in the English soccer league. Journal of Sports Economics, Vol. 2, No. 3, pp. 257-274.

Downward, P., Dawson, A. and Dejonghe, T. (2010). Sports Economics; theory, evidence and policy. Routledge.

Eckard, E. W. (1998). The NCAA cartel and competitive balance in college football. Review of Industrial Organization, Vol. 13 (3), pp. 347–369.

Eckard, E. W. (2001a). Baseball's Blue Ribbon Economic Report: Solutions in search of a problem. Journal of Sports Economics, Vol. 2, Issue 3, pp. 213-227.
Eckard, E. W. (2001b). The origin of the reserve clause: Owner collusion vs. public interest. Journal of Sports Economics, Vol. 2, Issue 2, pp. 113-130.

Eckard, E. W. (2003). The ANOVA-based competitive balance measure. Journal of Sports Economics, Vol. 4, No. 1, pp. 74–80.

El-Hodiri, M. and Quirk, J. (1971). An economic model of a professional sports league. Journal of Political Economy, Vol. 79, No. 6, pp. 1302-1319.

Goosens, J. (2006). Competitive balance in European football: comparison by adapting measures: national measure of seasonal imbalance and top 3. Revista Di Dirrito Ed Economica dello Sport, Vol. 2, pp. 77-122.

Hadley, L., Ciecka, J. and Krautmann, A. (2005). Competitive Balance in the aftermath of the 1994 players' strike. Journal of Sports Economics, Vol. 6, No. 4, pp 379-389.

Horowitz, I. (1997). The increasing competitive balance in Major League Baseball. Review of Industrial Organization, Vol. 12, pp. 373–387.

Humphries, B. (2002). Alternative measures of competitive balance. Journal of Sports Economics, Vol. 3, No. 2, pp. 133-148.

Késenne, S. (2007a). The Economic Theory of Professional Sports Leagues. Cheltenham: Edward Elgar.

Késenne, S. (2007b). The peculiar international economics of professional football in Europe. Scottish Journal of Political Economy, Vol. 54, No. 3, pp. 388-399.

Késenne, S., and Szymanski, S. (2004). Competitive balance and gate revenue sharing in team sports, Journal of Industrial Economics, LII, 1, pp. 165-177.

Knowles, G., Sherony, K. and Haupert, M. (1992). The demand for Major League Baseball: A test of the uncertainty of outcome hypothesis. The American Economist, Vol. 36, No.2, pp. 72-80.

Koning, R. H. (2000). Balance in competition in Dutch soccer. Statistician, Vol. 49, No. 3, pp. 419-431.

Koop, G. (2004). Modelling the evolution of distributions: An application to Major League baseball. Journal of the Royal Statistical Society. Series A (Statistics in Society), Vol. 167, No. 4, pp. 639-655.

Kuypers, T. (1997). The beautiful game? An econometric study of audiences, gambling and efficiency in English football. Unpublished PhD Thesis, University College London.

Lee, Y. H. and Fort, R. (2005). Structural change in MLB competitive balance: The depression, team location and integration. Economic Inquiry, Vol. 43, No. 1, pp. 158-169.

Leeds, M. and von Allmen, P. (2008). The Economics of Sports. Third edition. Pearson Education Inc.

Lenten, L. J. (2009). Towards a new dynamic measure of competitive balance: A study applied to Australia's two major professional 'football' leagues. Economic Analysis and Policy, Vol. 39, No. 3, pp. 407-428.

Levin, R., Mitchell, G., Volcker, P. and Will, G. (2000). The Report of the Independent Members of the Commissioner's Blue Ribbon Panel on Baseball Economics. NY: Major League Baseball.

Michie, J., and Oughton, C. (2004). Competitive balance in football: Tends and effects. Football Governance Research Centre, Birkbeck College, University of London.

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. Economic Bulletin, Vol. 26, No. 3, pp. 1-7.

Mizak, D. and Stair, A. (2004). The relationship between payroll and performance disparity in major league baseball: an alternative measure. Economics Bulletin, Vol. 12, No. 9, pp. 1-14.

Mizak, D., Stair, A. and Rossi, A. (2005). Assessing alternative competitive balance measures for sports leagues: a theoretical examination of standard deviations, gini coefficients, the index of dissimilarity. Economic Bulletin, Vol. 12, No. 5, pp. 1-11.

Noll, R. (1989). The economics of sports leagues. In: Law of Professional and Amateur Sports, Ed by Gary A Uberstine, K R Stratos and R J Grad, Deerfield Ill, The West Group, 1989, 17-2.

Noll, R. (1991). Professional basketball: Economic and business perspectives. In: Staudohar, P. D. and Mangan, J. A. (eds). The Business of Professional Sports. University of Illinois Press, Urbana and Chicago. pp. 18-47.

Owen (2010). Limitations of the relative standard deviation of win percentages for measuring competitive balance in sports leagues. Economics Letters, 109, p38-41.

Owen (2012). Measuring parity in sports leagues with draws: Further comments. Journal of Sports Economics, Vol. 13, No. 1, pp. 85-95.

Owen, P. D., Ryan, M, and Weatherspoon, C. R. (2007). Measuring competitive balance in professional team sports using the Herfindahl-Hirschman index. Review of Industrial Organisation, Vol. 31, No. 4, pp. 289-302.

Quirk, J. and El-Hodiri, M. (1974). The economic theory of professional sports leagues, in Government and the sports business, edited by Noll, R., Washington: Brookings, 1974

Quirk, J. and Fort, R. D. (1992). Pay Dirt, The Business of Professional Team Sports. Princeton, New Jersey: Princeton University Press.

Rascher, D. (1997). A model of a professional sports league. In Hendricks, W. (ed.), Advances in the Economics of Sport, Vol. 2, Greenwich and London: JAI Press, pp. 27-76.

Rottenberg, S. (1956). The baseball players' labor market. Journal of Political Economy, Vol. 64, No. 3, pp. 242-58.

Schmidt, M. B. (2001). Competition in Major League Baseball: The impact of expansion. Applied Economics Letters, Vol. 8, No. 1, pp.21-26.

Schmidt, M. B. and Berri, D. J. (2001). Competitive balance and attendance: The case of Major League Baseball. Journal of Sports Economics, Vol. 2, No. 2, pp. 145-167.

Scully, G. W. (1989). The Business of Major League Baseball. Chicago: University of Chicago Press.

Scully, G. W. (1992). Momentum in sports. In: Scully, G. W. (ed). Advances in the economics of sport. Vol. 1, JAI Press, Greenwich, Connecticut. pp.135-149.

Skaperdas, S. (1996). Contest success functions. Economic Theory, 7, pp.283-90.

Staudohar, P. (1999). Salary caps in professional team sports. In: Jeanrenaud, C. and Késenne, S. (eds.). Competition policy in professional sports: Europe after the Bosman case. Antwerpen: Standaard Edition.

Szymanski, S. (2001). Income inequality, competitive balance and the attractiveness of team sports: Some evidence and a natural experiment from English soccer. Economic Journal, 111, F69-F84.

Szymanski, S. (2004). Professional team sports are only a game: The Walrasian fixed-supply conjecture model, contest-Nash equilibrium, and the invariance principle. Journal of Sports Economics, Vol. 5, No. 2, pp. 111-126.

Szymanski, S. & Kuypers, T. (1999). Winners and Losers: The Business Strategy of Football. Viking.

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. Vol. 7, Issue 1. ISSN (Online) 1559-0410, DOI: 10.2202/1559-0410.1297, January.

Utt, J., and Fort, R. (2002). Pitfalls to measuring competitive balance with gini coefficients. Journal of Sports Economics, Vol. 3, No. 4, pp. 367-373.

Vrooman, J. (1995). A general theory of professional sports leagues. Southern Economic Journal, Vol. 61, No. 4, pp. 971-990.

Vrooman, J. (1996). The baseball player's labour market reconsidered. Southern Economic Journal, Vol. 63, No. 2, pp. 339-360.

Winfree, J. and Fort, R. (2012). Nash conjectures and talent supply in sports league modeling: A comment on current modeling disagreements. Journal of Sports Economics, Vol. 13, No. 3, pp. 306-313.

Zimbalist, A. (2002). Competitive balance in sports leagues. Journal of Sports Economics, Vol. 3, No. 2, pp. 111–121.

Zimbalist, A. (2003). Mat the Best Team Win: Baseball Economics and Public Policy. Washington D. C.: Brookings Institution Press.