Measuring Competitive Balance in Sports using Generalised Entropy with an Application to English Premier League Football

Vani Borooah, John Edward Mangan

Measuring Competitive Balance in Sports using Generalised Entropy with an Application to English Premier League Football

Journal:	*Applied Economics*
Manuscript ID:	APE-08-0697.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	30-Jul-2010
Complete List of Authors:	Borooah, Vani; University of Ulster, Economics and Politics Mangan, John; University of Queensland, Economics
JEL Code:	J44 - Professional Labor Markets and Occupations < J4 - Particular Labor Markets < J - Labor and Demographic Economics, L83 - Sports|Gambling|Recreation|Tourism < L8 - Industry Studies: Services < L - Industrial Organization, D63 - Equity, Justice, Inequality, etc. and Measurement < D6 - Welfare Economics < D - Microeconomics
Keywords:	Inequality, Entropy, Football, Premier League

.

Measuring Competitive Balance in Sports using Generalised Entropy with an Application to English Premier League Football[+]

Vani K. Borooah[*]
University of Ulster

John Mangan[**]
University of Queensland

July 2010

[+] We are grateful to an anonymous referee for valuable comments. Needless to say, the usual disclaimer applies.

[*] Corresponding author: School of Economics, University of Ulster, Newtownabbey, Northern Ireland BT37 0QB. (Email: vk.borooah@ulster.ac.uk).

[**] Department of Economics, University of Queensland, Brisbane, Australia (Email: J.Mangan@bel.uq.edu.au).

Abstract

A central issue in the economics of sport is the degree of competitive balance in sporting contests. The importance attached to competitive balance is predicated on the belief is that it is uncertainty about the outcomes of sporting contests that attracts spectators and sponsors. In a perfectly balanced competition, each team would have an equal chance of winning each match and, therefore, of winning the championship or the league. By contrast, the absence of competitive balance would mean that the results of sporting contests would become predictable and attendance at sporting contests would suffer. The general theme that underpins the issue of competitive balance is that of inequality. This paper proposes a general measure of competitive balance based on the based on the Generalised Entropy approach to measuring inequality and shows how this might be interpreted in terms of the league's welfare. The measures are applied to results from the 2006-07 season of the English Premier League.

1. Introduction

A widely held view in the economics of sport is that the long term sustainability of sporting leagues, in terms of both retaining the interest of their fans and generating income from the games, depends upon the degree of *competitive balance* in the league; yet, the continuing success of the English Premier League (EPL), and the intense interest that EPL games generate throughout the world, exists in spite of the fact that - as a number of studies have shown - the EPL is characterised by a low level of competitive balance.[1] This paper investigates this apparent paradox. In doing so, it investigates issues associated with the measurement of competitive balance and, in particular, the importance of *within-group* and *between-group* variation in team performance for generating interest in the league's matches. In summary, the paper shows that the high degree of balance *within* groups of teams in the EPL ensures a high level of fan interest, notwithstanding the fact that competitive balance in the EPL, *considered in its entirety,* is low.

The importance attached to competitive balance is predicated on the belief that it is uncertainty about the outcomes of sporting contests that attracts spectators. In a perfectly balanced competition, each team would have an equal chance of winning each match and, therefore, of winning the championship or the league. By contrast, the absence of competitive balance would mean that the results of sporting contests would become predictable and, in consequence, attendance at sporting contests would suffer (see *inter alia* Michie and Oughton, 2004; Szymanski, 2003 and 2007; and Zimbalist, 2003). In response to such concerns, several methods have been proposed for measuring competitive balance in sporting contests (Scully, 1989; Quirk and Fort, 1997, Humphreys, 2002; Michie and Oughton, 2004). The general theme that underpins these measures is that of *inequality analysis*. In the context of a league - which is the subject of this paper - the more equitable the inter-team distribution of the total number of points associated with the games' outcomes, the more balanced would be competition within the league.

However, the identification of competitive balance with inequality raises several ancillary questions. First, and foremost, how should inequality be measured? One of the drawbacks of existing analyses of competitive balance is that they do not fully mine the rich vein of methodology that the study of inequality provides. To use the language of Cowell (1995), many of the measures that are currently used came about more or less by accident, with concepts borrowed from statistics being pressed into service as tools of inequality measurement.

Second, if there is a lack of competitive balance in a league, can one identify its source? Is there competitive balance *within* subgroups of teams in the league but little balance *between* teams in the subgroups? Or, does the lack of competitive balance permeate the league in its entirety and affect all the teams.

Third, given that the teams in a league enjoy different degrees of success - as measured by their end of season points - what is the *effective* number of teams in a league? In essence, calculating the effective number of teams in a league involves assigning a scalar value to a vector of inter-team distribution of points. At one extreme, this scalar value will (should) equal the number of teams that play in the league. This will occur when points are distributed equally among the teams. However, when the points are distributed unequally between the teams, the effective number of teams will be less than the actual number of teams.[2]

While most studies of competitive balance are concerned with baseball, Michie and

---

[1] See Mehra et. al. (2006) and Brandes and Frank (2007).
[2] The concept of the effective number of parties is used in the analysis of electoral systems where votes are unequally distributed between parties: see Laasko and Taagepera (1979).

Oughton's (2004) work is one of the few studies of football. Some of these questions raised in the preceding section have been addressed by them in their comprehensive study of English football. Their paper is discussed in some detail in section 6 and their results are compared with those contained in this paper.

Against this background, this paper proposes a general measure of competitive balance based on the *Generalised Entropy (GE)* approach (due to Theil, 1967) to measuring inequality and shows how this might be interpreted in terms of the league's welfare via the *Social Welfare Function* (SWF) approach to measuring inequality (due to Atkinson, 1970).

2. Background

In a sports league consisting of *N* teams, each team plays every other team *twice* in the course of a season: home and away[3]. Suppose that, for every game that it plays, a team is awarded z_w points for a win, z_d points for a draw, and 0 points for a loss: $z_w \geq 2z_d$. Suppose, in a season, a proportion *p* of games have a winner (loser) and (*1-p*) end in a draw. Consequently, in the course of a season, the total number of wins (*W*) and draws (*D*) result in a total (*T*) of points awarded in the league:

$$\begin{aligned} T &= z_w W + 2z_d D = z_w pN(N-1) + 2z_d(1-p)N(N-1) \\ &= N(N-1)\left[z_w p + 2z_d(1-p)\right] = N(N-1)[(z_w - 2z_d)p + 2z_d] \end{aligned} \quad (1)$$

The important point is that when $z_w \neq 2z_d$, the total number of points depends not just on the number of games played but, also, on the number of wins and draws. For example, in the English Premier League, a winning side receives *three* points ($z_w = 3$) and, in the event of a

[3]There are $N \times (N-1)$ independent games: team 1 plays the other $N-1$ teams twice; team 2 plays the other $N-2$ teams twice (excluding team 1, whom it has already played); team 3 plays the other $N-3$ teams twice (excluding teams 1 and 2, whom it has already played); and so on.

draw, *each* side receives *one* point ($z_d = 1$): so, with *N=20* sides, if, at one extreme, every game was drawn, $T = 2N(N-1) = 760$; at the other extreme, if every game resulted in a win, $T = 3N(N-1) = 1140$.

Suppose team *i* wins W_i, and draws D_i, games in a season. Then team *i* 's end-of-season points are: $V_i = z_w W_i + z_d D_i$. Since each team plays $2(N-1)$ games, the maximum and minimum points a team can obtain in a season are, respectively, $2z_w(N-1)$ and 0. Consequently,

$0 \le V_i \le 2z_w(N-1)$

Under *perfect* competitive balance, each team would be expected win and to draw the same number of games: $W_1 = W_2 = ...W_N = pN(N-1)/N = p(N-1)$ and $D_1 = D_2 = ...D_N = (1-p)N(N-1)/N = (1-p)(N-1)$ games. So, under perfect competitive balance, each team would be expected to end the season with the same number of points:

$$V_1 = V_2 = ...V_N = z_w p(N-1) + 2z_d(1-p)(N-1) = (N-1)[(z_w - 2z_d)p + 2z_d] \qquad (2)$$

On the other hand, *competitive imbalance* implies that the total number of wins and draws would be unequally divided between the teams, the degree of imbalance increasing with the degree of inequality. Suppose the teams were ranked in descending order of the points they obtained (say team *1* with the maximum, and team *N* with the minimum, points). Then inequality would be greatest - competition would be most unbalanced - if all the results were perfectly predictable: team 1 wins all its *2(N-1)* games, $V_1 = 2z_w(N-1)$; team 2 wins all its *2(N-2)* games, $V_2 = 2z_w(N-2)$; and so on till, say, team *r,* $V_r = 2z_w(N-r)$; the N^{th} team loses all its games and finishes without any points, $V_N = 2z_w(N-N) = 0$.

Let the points share of the i^{th} team be denoted by $v_i \ge 0$, $v_i = V_i / T = V_i / N\overline{V}$, where $\overline{V}$

is the average number of points computed over all the teams and $\sum_{i=1}^{N} v_i = 1$. Then under perfect competitive balance, $v_i = 1/N$ and, when competition is most unbalanced, $v_i = [2z_w(N-i)]/[z_w N(N-1)] = 2(N-i)/N(N-1), i = 1...N$. So, when competition is most unbalanced, the points share of the teams would be: *2/N* for team 1 ; $[2(N-2)]/[N(N-1)]$ for team 2 ; and so on till $[2/[N(N-1)]$ for team *N-1* and 0 for team *N*.

3. Properties of Inequality Indices

The measurement of competitive balance, as argued above, is synonymous with the measurement of inequality and the literature on inequality informs us that a "good" measure of inequality should satisfy certain properties:

1. The *weak principle of transfers* (also known as the Pigou-Dalton property): an egalitarian transfer (i.e. from a richer to a poorer person) causes the value of the inequality index to fall and a regressive transfer (i.e. from a poorer to a richer person) causes it to rise.
2. *Scale independence:* if everyone's quantity (points, income etc.) increased by the same *proportion* inequality would remain unchanged.[4]
3. *Population homogeneity:* if the population of *N* persons is replicated, then inequality in this larger population of *2N* persons would remain unaltered.
4. *Decomposability*: if the population is divided into mutually exclusive groups, total inequality can be expressed as a function of inequality *within* subgroups and inequality *between* subgroups. More specifically, *additive* decomposability allows total inequality to be expressed as the *sum* of within subgroup and between subgroup inequalities.

3.1 Generalised Entropy Measures

[4]For example, if under a new system, the points were awarded as 6 for a win, 2 for a draw, and 0 for a loss, then

Shorocks (1980) showed that the only inequality index satisfying properties 1-4 above were measures belonging to the *Generalised Entropy (GE)* family of measures, defined by the parameter θ, and written $GE(\theta)$:[5]

$$GE(\theta) = \frac{1}{N}\frac{1}{\theta^2 - \theta}\left[\sum_{i=1}^{N}\left(\left[\frac{V_i}{\overline{\overline{V}}}\right]^{\theta} - 1\right)\right]$$
$$= \frac{1}{\theta^2 - \theta}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i}{\overline{\overline{V}}}\right)^{\theta} - 1\right], \text{ if } \theta \neq 0,1 \tag{3}$$

$$GE(1) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i}{\overline{\overline{V}}}\right)\times \log\left(\frac{V_i}{\overline{\overline{V}}}\right) \tag{4}$$

$$GE(0) = \frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{V_i}{\overline{\overline{V}}}\right) \tag{5}$$

In particular, the Generalised Entropy indices, defined by equations (3), (4), and (5) are additively decomposable. By this is meant that if the N teams was divided into K groups (indexed, $k = 1...K$), with N_k teams in each subgroup, then the overall inequality index $GE(\theta; N)$ of equations (3), (4), and (5), could be written as:

$$GE(\theta; N) = \sum_{k=1}^{K} w_k GE(\theta; N_k) + \mathbf{B} \tag{6}$$

where the first ($\sum w_k GE(\theta; N_k)$), and second (**B**), terms in equation (6) represent, respectively, the contribution of *within* group inequality[6] and *between* group inequality to overall inequality.[7]

---

each team's points would double but inequality would remain unchanged.

[5]For a monotonic function $\Phi(.)$, this includes measures $\Phi[GE(\theta)]$ which are ordinally equivalent to $GE(\theta)$.

[6]Written, for weights w_k, as the weighted average of the subgroup inequality values, $GE(\theta; N_k)$.

[7]*GE(1)* and GE(0) of equations (4) and (5) are, respectively, the Theil and the Mean Logarithmic Deviation (*MLD*) indices (Theil, 1967, chapter 4).

3.2 Entropy and the Atkinson Index

An important inequality index is that due to Atkinson (1970). This index, referred to as the *Atkinson index (AT index)*, which pioneered the interpretation of inequality in terms of social welfare and, thereby, spawned the *Social Welfare Function* approach to inequality measurement, is written as:

$$I_\varepsilon = 1 - \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{V_i}{\overline{\overline{V}}} \right)^{1-\varepsilon} \right]^{1/(1-\varepsilon)} \quad \text{for } \varepsilon \neq 1 \tag{7}$$

$$I_\varepsilon = 1 - \frac{1}{\overline{\overline{V}}} \left[\prod_{i=1}^{N} V_i \right]^{1/N} \quad \text{for } \varepsilon = 1 \tag{8}$$

The AT index of equation (7) can be obtained as an ordinal transformation of the generalised entropy indices of equation (3) for $\theta < 1$ by setting $\theta = 1 - \varepsilon$ for $\varepsilon \neq 1$ and the AT index of equation (8) can be obtained from the GE index of equation (5) for $\theta = 0$ by setting $\theta = 1 - \varepsilon$ for $\varepsilon = 1$. Thus, the Atkinson index and the Generalised Entropy indices are *ordinally equivalent*.

3.3 Entropy and the Herfindahl Index

When $\theta = 2$ in equation (3), the generalised entropy index is:

$$GE(2) = \frac{1}{2} \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{V_i}{\overline{\overline{V}}} \right)^2 - 1 \right] = \frac{N}{2} \left[\sum v_i^2 - \frac{1}{N} \right]$$

which is cardinally equivalent to the Herfindahl index, $H = \sum_{i=1}^{N} v_i^2$, which is the sum of squares of the individual shares. Also, when $\theta = 2$:

$$GE(2) = \frac{1}{2}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i}{\bar{V}}\right)^2 - 1\right] = \frac{1}{2}\left[\frac{\sum(V_i - \bar{V})^2}{N\bar{V}^2}\right] = \frac{1}{2}(CV)^2$$

so that *GE(2)* is cardinally equivalent to the square of the coefficient of variation (*CV*).

3.4 Entropy and Information Theory

The GE indices, set out above in equations (3), (4), and (5), can be interpreted in terms of the information content of the data. If the function *h(.)* is defined, for $z \geq 0$, as:

$$h(z) = \frac{1 - z^{\beta}}{\beta} \text{ if } \beta \neq 0 \text{ and } h(z) = -\log(z) \text{ if } \beta = 0 \tag{9}$$

then the family of *Information-Theoretic* (*IT*) Measures, $H(\beta)$, is obtained as:

$$\begin{aligned} H(\beta) &= \frac{1}{1+\beta}\left[\sum_{i=1}^{N}\frac{1}{N}h\left(\frac{1}{N}\right) - \sum_{i=1}^{N}v_i h(v_i)\right] \\ &= \frac{1}{\beta+\beta^2}\sum_{i=1}^{N}v_i\left[v_i^{\beta} - N^{-\beta}\right] = \frac{1}{\beta+\beta^2}\left(\sum_{i=1}^{N}v_i^{1+\beta} - N^{-\beta}\right) \end{aligned} \tag{10}$$

The logic of this class of measures is as follows. Suppose a random variable *x* can take values $x_1 \ldots x_N$ with probabilities $p_1 \ldots p_N$, $0 \leq p_i \leq 1$, $\sum p_i = 1$. Hence the *information content* h_i of observing *x* take the value x_i can be regarded as a decreasing function of p_i : if p_i is large/small, then it would not/would be a surprise if $x = x_i$ and so the "information content", h_i of the observation would be small/large (Renyi, 1965). A measure of the "expected amount of information" or *entropy* conveyed by the observations, $x_1 \ldots x_N$ is $e = \sum p_i h(p_i)$ and equation (9), above, represents a formulation of the "information content" function, *h(.)*, in terms of a parameter β.

The measure of inequality $H(\beta)$ in equation (10), is obtained by subtracting the actual entropy of the distribution of point shares across the N teams, $v_1 ... v_N$, from the maximum possible value of this entropy which obtains when every team gets an equal share of points ($v_i = 1/N, \forall i$). The expression for $H(\beta)$ in equation (10) is derived from the definition of *h(.)* in equation (9).[8] The GE measure of equation (3) may be derived from the IT measure of equation (10) by setting $\theta = 1 + \beta$ in equation (10) and normalising for the population principle by multiplying equation (10) by N^{β}.[9]

3.4 The Interpretation of the θ Parameter in Generalised Entropy

All inequality indices should embody the *weak principle of transfers*. In the case under discussion, this principle (also known as the Pigou-Dalton property: Dalton, 1920) requires that a transfer of points from a "stronger" to a "weaker" team should cause the value of the inequality index to fall. But by *how much* the value of the inequality index will fall, following this "egalitarian" points transfer, will depend upon the value of the parameter, θ. The value of θ, therefore, measures the "transfer sensitivity" of the inequality index: the *larger* the value of θ, the *greater* will be the fall in inequality, following a transfer of points from a stronger to a weaker team. In the context of the "social welfare" approach to inequality measurement,

---

[8]In equation (10):

$$\sum v_i h(v_i) = \sum v_i \left[\frac{1 - v_i^{\beta}}{\beta}\right] = \frac{1}{\beta} - \frac{\sum v_i v_i^{\beta}}{\beta}, \text{ since } \sum v_i = 1$$

$$\text{and } \sum N^{-1} h(N^{-1}) = \sum N^{-1} \left[\frac{1 - N^{-\beta}}{\beta}\right] = \frac{1}{\beta} - \frac{N^{-\beta}}{\beta}$$

$$\Rightarrow H(\beta) = \frac{1}{\beta + \beta^2}\left[\sum v_i (v_i^{\beta} - N^{-\beta})\right] = \frac{1}{\beta + \beta^2}\left(\sum v_i^{1+\beta} - N^{-\beta}\right)$$

[9] The family of information-theoretic measures, $H(\beta)$ in equation (10), does not satisfy the principle of population homogeneity because of the presence of the term $N^{-\beta}$. However, dividing through by N^{β} ensures the property is satisfied.

pioneered by Atkinson (1970), discussed below, the value of θ represents league's degree of "inequality aversion".

4. Inequality and Welfare

Atkinson's (1970) pioneering approach to inequality measurement established a connection between inequality and welfare. Suppose that the *league's welfare*, *J*, is the sum, *over all the N teams*, of the utility attached to the points obtained by its teams:

$$J = \sum_{i=1}^{N} F(V_i) \tag{11}$$

The *league welfare function* (LWF), $F(.) \geq 0$ in equation (11), represents the *league's valuation* of the gain (to it) from team *i* ending the season with V_i points, with larger values of $F(.)$ representing higher levels of gain. The sum of the team-specific gains is the *league welfare* associated with a given total of points available for distribution between the *N* teams: $T = \sum_{i=1}^{N} V_i$.

The important point, made earlier, is that when the number of points for a win (z_w) was greater than *twice* the number of points for a draw (z_d) - as happens in the EPL with $z_w = 3$ and $z_d = 1$ - the total number of points depended not just on the number of games played but, also, on the number of wins and draws. Consequently, this total would only be known *at the conclusion* of the season's league games.[10]

The *change* in league welfare, following a change in the V_i, is:

[10] With *N* sides, if, at one extreme, every game was drawn, $T = z_d N(N-1)$; at the other extreme, if every game resulted in a win, $T = z_w N(N-1)$. However, if $z_w = 2 \times z_d$, the total number of points will be independent of the number of wins and draws: if, say, $z_w = 2$ and $z_d = 1$, $T = 2 \times N \times (N-1)$.

$$\Delta J = \sum_{i=1}^{N} a_i \Delta V_i \qquad (12)$$

where: $a_i = \partial F(V_i) / \partial V_i > 0$ is the *marginal change* in league welfare consequent upon a change in the points of team i. If it is assumed that the function *F(.)* is *strictly concave,* then marginal gain *decreases* for increases in V_i. Consequently, for a given total of available points, *T,* league welfare is maximised when all the teams are expected to finish the season with the same number of points: $V_1 = V_2 = .. = V_N$

The LWF has *constant elasticity* if, for $\varepsilon \geq 0$, *F(.)* can be written as:

$$F(V_i) = \frac{V_i^{1-\varepsilon} - 1}{1-\varepsilon}, \varepsilon \neq 1; F(V_i) = \log(V_i), \varepsilon = 1 \qquad (13)$$

since then: $a_i = \partial F(V_i) / \partial V_i = V_i^{-\varepsilon} > 0 \Rightarrow (\partial a_i / \partial V_i) / (V_i / a_i) = -\varepsilon < 0$. The *percentage* change in the welfare weights, following a *percentage* change in points is, therefore, *both negative and constant*. The greater the value of ε, the greater the proportional *decrease* in the welfare weight associated with a team in response to a proportional *increase* in its points. The parameter ε represents, as shown below, the league's *aversion to (inter-team) inequality*.

Consider two teams in the English Premier League: a "weak" team, Charlton ($i = C$) and a "strong" team Liverpool ($i = L$). Suppose one was evaluating the welfare consequences of increasing Charlton's points by ΔV_C and reducing Liverpool's points by $\Delta V_L = \beta \Delta V_C$. If $\Delta V_i = 0$ for $i \neq C, L$, then the change in the value of J, following this redistribution between Charlton and Liverpool is:

$$\Delta J = a_C \Delta V_C + a_L \Delta V_L = V_C^{-\varepsilon} \Delta V_C - V_L^{-\varepsilon} \beta \Delta V_C \qquad (14)$$

In the 2006-07 Premier League, Liverpool and Charlton finished on 68 and 34 points, respectively, so that $V_L / V_C = 2$. Setting $\Delta J = 0$ in equation (14) implies:

$$\beta = (V_C / V_L)^{-\varepsilon} = 2^{\varepsilon} \tag{15}$$

Suppose that Charlton's points were increased by one. If $\varepsilon = 0$, then, from equation (15) $\beta = 1$ and the league would be prepared to reduce Liverpool's points by one leaving the total of points, T - and hence, the number of teams in the league N - unchanged. However, if $\varepsilon > 0$, the league would be prepared to reduce Liverpool's points by more than one : if $\varepsilon = 1$ then $\beta = 2$, and, in order to leave its welfare unchanged ($\Delta J = 0$) the league would be prepared to reduce Liverpool's points by two in order to secure a point increase for Charlton.

In this sense, the value of ε represents the degree to which the league is averse to competitive imbalance: higher values of ε represent a willingness to sacrifice larger amounts of the total number of available points in order to obtain greater equality in its distribution. Since the total number of points available for distribution is $T = \sum_{i=1}^{N} V_i$, where $z_d N(N-1) \leq T \leq z_w N(N-1),\ z_w > 2z_d$, a reduction in the total number of available points would be effected by reducing N, the number of teams in the league.

Let $V^* \leq \overline{V} = \Omega / N$ represent the number of points which, *if received by all the teams,* would yield the *same* level of league welfare as the *existing* distribution of points: $V_1, V_2 .... V_N$. Then V^* may be termed the *equally distributed equivalent points*. In consequence:

$$N \times F(V^*) = \sum_{i=1}^{N} F(V_i) \Rightarrow V^* = F^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} F(V_i)\right)$$

By the definition of *F(.)* in equation (13), for $\varepsilon \neq 1$:

$$\frac{(V^*)^{1-\varepsilon}-1}{1-\varepsilon}=\sum_{i=1}^{N}\frac{1}{N}\left(\frac{(V_i)^{1-\varepsilon}-1}{1-\varepsilon}\right)\Rightarrow (V^*)^{1-\varepsilon}=\frac{1}{N}\sum_{i=1}^{N}(V_i)^{1-\varepsilon}\Rightarrow V^*=\left[\frac{1}{N}\sum_{i=1}^{N}(V_i)^{1-\varepsilon}\right]^{1/1+\varepsilon} \qquad (16)$$

while, for $\varepsilon = 1$:

$$\log(V^*)=\frac{1}{N}\sum_{i=1}^{N}\log(V_i)\Rightarrow V^*=\left[\prod_{i=1}^{N}V_i\right]^{1/N} \qquad (17)$$

Atkinson's (1970) inequality index applied to the inter-team distribution of points yields:

$$I_\varepsilon = 1-(V^*/\overline{V}) = 1-\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i}{\overline{V}}\right)^{1-\varepsilon}\right]^{1/(1-\varepsilon)} \quad \text{for } \varepsilon \neq 1 \qquad (18)$$

$$I_\varepsilon = 1-(V^*/\overline{V}) = 1-\frac{1}{\overline{V}}\left[\prod_{i=1}^{N}V_i\right]^{1/N} \quad \text{for } \varepsilon = 1 \qquad (19)$$

When $\varepsilon = 1$, the (log of) the Atkinson index is expressed as the logarithm of the ratio of the arithmetic mean to the geometric mean:

$$\log(I_\varepsilon)=\log(\overline{V}/V^*)=\log\left[\frac{\overline{V}}{\prod_{i=1}^{N}(V_i)^{1/N}}\right]=\log(\overline{V})-\frac{1}{N}\sum_{i=1}^{N}\log(V_i) \qquad (20)$$

Bourguigon (1979) suggested an interpretation to equation (20): if the function, *F(.)* in equation (11) is of the logarithmic form (i.e. $F(V_i)=\log(V_i)$), then $N \times \log(I_\varepsilon)$ represents the distance between the maximum amount of league welfare and the amount of welfare actually attained.

From equation (2), the *equally distributed equivalent teams, N^** may be defined as:

$$N^{*} \; 1 \; \frac{V^{*}}{z_w \; 2z_d \; P \; 2z_d}$$

where N^* is the number of teams *in a perfectly balanced league* (i.e. all teams obtain V^* points) which would yield the same level of league welfare as N teams in an unbalanced league (i.e. the N teams obtain points: $V_1, V_2, \ldots V_N$).

From equation (16), V^*, is what Anand and Sen (1997) term the "1-ε" average of the V_i ($i = 1 \ldots N$). When $\varepsilon = 0$, the league is indifferent as to how a given overall total of points is distributed between the teams: $V^* = \overline{V}, N^* = N$ and $I = 0$. For $\varepsilon > 0$, $V^* < \overline{V}$, $N^* < N$, and $I > 0$. The higher the value of the inequality aversion parameter, ε, the smaller will be the values of V^* and N^* and, therefore, higher will be the value of the inequality index, I. When $\varepsilon = 1$, V^* is the geometric mean (equation (17)), and when $\varepsilon = 2$, V^* is the harmonic mean (equation (16)) - of the teams' points $V_1, V_2 \ldots V_N$.

The reverse of the Atkinson transformation yields the league welfare as (Sen, 1973):

$$J = \overline{V}(1 - I) \tag{21}$$

The league welfare function in equation (21) has a natural interpretation: the welfare from an overall number of average points $\overline{V}$ - implied by having N teams in the league - *is reduced* by the extent of competitive imbalance. Given that there are N teams in the league, equation (21) says that the value of the Atkinson index (I) is a measure of the league welfare (J) associated with N teams. This welfare depends upon the degree of inter-team inequality (i.e. competitive imbalance) and this, in turn, is determined by the inter-team distribution of the total points *and* upon the degree to which the league is averse to inter-team inequality (i.e. competitive imbalance).

5. Empirical Results: English Premier League

In the 2006-07 season, there were $N=20$ teams in the EPL and $N(N-1) = 380$ games were played, of which 282 , or 74.2 percent, were won (or lost) and 98 , or 25.8 percent were drawn. Consequently, the total number of points distributed in 2006-07 between the 20 teams was $3\times 282 + 2\times 98 = 1024$. Under the Premier League's system of awarding points, the total number of points available for distribution between the teams (*T*) cannot be known at the start of the season but must await the end of the season when the number of wins and draws has been determined. The total depends, therefore, not just on the number of games played but, also, on the number of wins and draws: at one extreme, if every game in 2006-07 had been drawn, $T = 2\times N(N-1) = 760$; at the other extreme, if every game in 2006-07 had resulted in a win, $T = 3\times N(N-1) = 1140$.

The arithmetic, geometric, and harmonic means of the total points obtained by each team at the end of the 2006-07 season were, respectively: $\overline{V} = 52.1$, $V^*_{\varepsilon=1} = 49.96$, and $V^*_{\varepsilon=2} = 47.96$. Consequently, the values of N^* when $\varepsilon=1$ and $\varepsilon=2$ were, respectively, $N^*_{\varepsilon=1} = 1 + \{49.96/[(3-2)\times 0.7421+2]\} = 19.2$ and $N^*_{\varepsilon=2} = 1 + 47.96/2.7421 = 18.5$. In other words, in order to ensure perfect competitive balance, the league would have been prepared to play a smaller number of teams than the 20 which actually played: 19 teams when the degree of aversion to competitive imbalance was $\varepsilon=1$, and 18 teams when the higher degree of aversion to competitive imbalance was higher, with $\varepsilon=2$. When, as in equation (20), *log(.)* was used as the cardinal representation of *F(.)*, the gap between the maximum and the attained levels of league welfare ($\log(\overline{V} - N^{-1}\sum_{i=1}^{N}\log(V_i))$) was 1.06 percent.

Using the decomposability properties of GE indices (see equation (6)), the teams can be subdivided into groups with a view to examining how much of the league's competitive imbalance (as measured by inequality in the distribution of points between its teams) was due to competitive imbalance *within* the groups (*within group inequality*) and how much was the result of competitive imbalance *between* the groups (*between group inequality*).

If the 20 teams were subdivided into *two* groups - the first group consisting of the top five teams[11] and the second group comprising the remaining teams - then, with $\theta=2$, 65 percent of the EPL's competitive imbalance in 2006-07 could be attributed to competitive imbalance *between* the league's top five teams and the rest of the teams and only 35 percent could be attributed to competitive imbalance *within* the two groups of teams.[12] When the 20 teams were subdivided into four groups - the top five, the next five, the penultimate five, and the bottom five teams - then 84 percent of the Premier League's competitive imbalance in 2006-07 could be attributed to competitive imbalance *between* the four groups and only 16 percent could be attributed to competitive imbalance *within* the four groups.[13]

5.1 Inequality by Source of Points

Each team has two sources of points: points earned from away games (V_i^A) and points earned from home games (V_i^H) such that $V_i = V_i^A + V_i^H$. One may, therefore, ask how much of the overall inequality in the distribution of points between the league's teams stems from inequality in the distribution of points from away games and how much stems from inequality in

[11]In 2006-07, these were: Manchester United, Chelsea, Liverpool, Arsenal, and Tottenham.

[12]The use of θ=2 implies that the competitive imbalance (inequality) was measured using the coefficient of variation. The coefficient of variation embodies the property of being *transfer neutral*: three points transferred from the top team to the team ranked second - Chelsea beats Man United - would improve competitive balance (reduce inequality) by as much as a transfer of three points from the second lowest ranked team to the lowest ranked team (Watford beats Charlton).

[13]The value of *GE(2)* , computed over the distribution of points across the 20 teams, was 0.044. When the teams were divided into two groups - the top five and the rest - the between group contribution - the value of the term **B** in equation (6) - was 0.028 or 65 percent of total inequality; when the division was into four groups, the value of **B** was 0.037 or 84 percent of the total.

the distribution of points from home games? This is analogous to identifying the sources of income (wages, dividends, benefits etc.) and enquiring about the proportions of overall income inequality that could be explained by inequality in the distribution of its different components.

Shorrocks (1982) showed that the proportionate contribution, s^j of income component j to overall inequality was given by: $s^j = \frac{\text{covariance}(\mathbf{y}^j, \mathbf{y})}{\text{variance}(\mathbf{y})}$, where: $\mathbf{y}^j$ and $\mathbf{y}$ were, respectively, the vector of values of income component j (say, wages) and total income across the N income earners and $\sum s^j = 1$. The decomposition rule embodied in the above equation is unique and invariant in that it "avoided one of the major problems encountered in applied work on distributions: that of having constantly to qualify results by stating that they hold only for the particular index selected" (Shorrocks, 1982, p. 205).

The average number of points per team in the 2006-07 season was 51.2 (a total of 1,024 points across 20 teams) of which home and away games resulted in averages of, respectively, 32.2 and 19.9 with a correlation of 0.67 between home and away game points. Using he decomposition by source equation above, inequality in the distribution of home game points contributed 49 percent, with inequality in the distribution of away game points contributing 51 percent, to inequality in the distribution of total points.

6. Comparison with Other Studies

This paper's contribution is that, within the over-arching theme of measuring competitive balance using the methods of inequality analysis, it paid special attention to: the *decomposition of inequality*, both by subgroups of teams and by the source of team points (home and away matches); the *loss of league welfare* associated with competitive imbalance and, by corollary, the number of teams in the league *which with perfect competitive balance* would yield the same amount of welfare as the current number of teams and the existing level of balance.

Michie and Oughton (2004) in their magisterial study of competitive balance in English football employed several inequality measures to quantify competitive balance. The first measure which they considered was the "standard deviation of win percentages" (SDW) defined as the ratio of the standard deviation of the percentage of matches each club wins in a season to an "ideal" standard deviation based on the distribution of win percentages in a perfectly balanced league where each club had a probability of 0.5 of winning each match. The GE measures employed in this paper relax the constraint of only considering wins to considering wins *and* draws. Moreover, by studying the inter-team distribution of the total number points in the league and, by examining the inequality in this distribution, the GE measures implicitly compare the actual distribution of total points with an idealised distribution in which the total of points is equally distributed between the teams.

Another measure of competitive balance considered by Michie and Oughton (2004) was the Herfindahl index of concentration applied to the share of each club in the league's total points. As shown in section 3.3, the *GE (2)* is cardinally equivalent to the Herfindahl index. A third measure of balance considered by Mitchie and Oughton (2004) was based on the Lorenz curve: the greater the distance of the curve from the line of "perfect balance", the higher the level of competitive imbalance. The GE measures of competitive balance, proposed in this paper, are also Lorenz-based in that higher values of the GE measure are associated with the Lorenz curve being further away from the perfect balance line.[14]

A final measure employed by Michie and Oughton (2004) is the "five club concentration

[14] This is because the GE measures, by virtue of satisfying the scale independence property, belong to the family of relative inequality indices.

ratio" (C5), defined as the ratio of points won by the top five clubs to the total number of points, and the related "five club index of competitive balance" (C5CIB) obtained, for *N* teams, as the ratio: $\frac{C5}{5/N}$. As this paper showed, the GE measures are capable of decomposing inequality between groups (in this case, of teams) with a view to determining the proportions of total inequality emanating from within, and from between, group inequality.

This allowed this study to ask the Mitchie and Oughton (2004) question in a more general form: if one divided the teams into groups, how much of the league's competitive imbalance stemmed from imbalance within the teams in a group and how much stemmed from imbalance between the groups? When the teams were subdivided into the top five teams and the rest, the answer was equivalent to that provided by the C5 and the C5CIB: using data from the 2006-07 season, 65 percent of the league's imbalance was due to imbalance *between* the group of comprising top five teams and the group comprising the other teams and 35 percent was due to imbalance *within* the two groups of teams. Michie and Oughton (2004), by calculating the share of points claimed by the top five teams (C5 ratio) - which ranged from 29 percent in 1993, to a high of 37 percent in 2003, falling back to 35 percent in 2004 - made a *factual* point. This paper's finding draws out the implications of this by using the methods of decomposition analysis to make an *analytical* point: 65 and 35 percent of the league's imbalance came from, respectively, between and within group inequality.

The further advantage that the GE measures offered in terms of decomposability was that the groups could be defined in any desired manner and not necessarily restricted to the top five teams and the rest. Indeed, this study considered a four-fold division of the league's teams and found that that 84 percent of the league's competitive imbalance stemmed from a lack of balance *between* the four groups implying that there was a high degree of competitive balance *within* each of the groups.

7. Conclusions

This paper examined measurement issues associated with competitive balance against the backdrop of the EPL: a significant characteristic of the EPL is that a high level of popularity co-exists with a low level of competitive balance. It was argued that a closer alliance between the measurement of competitive balance and the insights provided by the methods of inequality analysis - the particular case here being those of generalised entropy - deepens our understanding of competitive balance and provides valuable insights into the implications of a lack of balance. By identifying the sources of imbalance and the potential impact of structural change in sports leagues (such as changes in the size of the league), it also provides a means of extending the analytical compass of competitive balance measures. In this context, the paper developed the concept of the effective numbers of teams and also reinterpreted well known general entropy indices in terms of the league's aversion to inter-team inequality.

It was found that when the 20 teams, that currently comprise the EPL, were divided by league position, starting with the top five teams, the bulk (84 percent) of the EPL's competitive imbalance could be attributed to between-group imbalance and only a small amount to within-group imbalance. This confirms the widely held conjecture that, in its present form, the EPL is segmented into several (effectively) non-competing groups: the top five teams who realistically aspire to league and European Championship honours; those competing for admission to lesser European competitions; and the lower groups whose sole aim is simply to be part of the EPL.

The competitive aspects of the league take place *within* groups of teams and competitive imbalance results from large disparities in personnel and resources between the teams. When within group competition is high then, despite the lack of competition between teams in the different groups, a sporting league may still be successful. Within the specific context of the EPL, the apparent lack of desire to change the current EPL format suggests that it does not see a lack of competitive balance in the League, taken its entirety, as a problem. Perhaps intuitively, the EPL is aware of the validity of this paper's analysis: that league success does not require effective competition between *all* the teams in the league. However, should the EPL seek to move towards more balanced competition between *all* its teams, this study's analysis of the effective number of teams suggests that a possible solution might be a reduction in the number of teams in the league.

References

Anand, S. and Sen, A.K. (1997), Concepts of Human Development and Poverty: A Multidimensional Perspective, Human Development Report 1997 Papers, New York: UNDP.

Atkinson, A.B. (1970), "On the Measurement of Inequality", Journal of Economic Theory vol. 2, pp. 244-63.

Bourguignon, F. (1979), "Decomposable Income Inequality Measures", Econometrica, vol. 47, pp. 901-20.

Brandes, L and Franck, E (2007), "Who made Who? An Empirical Analysis of Competitive Balance and the Attractiveness of Team Sports", Eastern Economic Journal, vol. 33, no. 3, pp. 379-403.

Cowell, F.A. (1995), Measuring Inequality, Prentice Hall-Harvester Wheatsheaf: London, Second Edition.

Dalton, H. (1920), "The Measurement of the Inequality of Incomes", Economic Journal, vol. 30, pp. 361-84.

Humphreys, B.T. (2002), "Alternative Measures of Competitive Balance in Sports Leagues", Journal of Sports Economics, vol. 3, pp. 133-47.

Michie, J. and Oughton, C. (2004), Competitive Balance in Football: Trends and Effects, Research Paper 2004 No. 2, Football Governance Research Centre, Birkbeck College: London.

Laasko, M. and R. Taagepera, 1979, "Effective Number of Parties: A Measure with Application to Western Europe", Comparative Political Studies, vol. 12, pp. 3-27.

Mehra, S. and Zuercher, T. J., "Striking Out 'Competitive Balance' in Sports, Antitrust and Intellectual Property" . Berkeley Technology Law Journal, Fall 2006 Available at SSRN: http://ssrn.com/abstract=927771.

Quirk, J. and Fort, R. (1997), Pay Dirt: The Business of Professional Teams Sports, Princeton NJ: Princeton University Press.

Renyi, A. (1965), "On the Foundations of Information Theory", Review of the International Statistical Institute, vol. 33, pp. 1-14.

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

Shorrocks, A.F. (1980), "A Class of Additively Decomposable Inequality Measures", Econometrica, vol. 48, pp. 613-25.

Shorrocks, A.F. (1982), "Inequality Decomposition by Factor Components", Econometrica, vol. 52, pp. 1369-85.

Sen, A.K. (1973), On Income Inequality , Oxford University Press: Oxford.

Szymanski, J. (2003), "The Economic Design of Sporting Contests", Journal of Economic Literature, vol. XLI, pp. 1137-87.

Szymanski, J. (2007), "The Champions League and the Coase Theorem", Scottish Journal of Political Economy, vol. 54, pp. 355-73.

Theil, H., 1967, *Economics and Information Theory*, North-Holland: Amsterdam.

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