Journal of Physical Education and Sport ® (JPES), 18(Supplement issue 3), Art 198, pp.1335 - 1340, 2018
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Measuring competitive balance in the major european soccer leagues

FRANCISCO TRIGUERO-RUIZ[1], ANTONIO AVILA-CANO[2]
[1]Department of Languages and Computer Science, University of Málaga, SPAIN
[2]Department of Economics and Economic History, University of Málaga, SPAIN

Published online: August 31, 2018
(Accepted for publication July 15, 2018)
DOI:10.7752/jpes.2018.s3198

Abstract:
Competitive balance of a sports competition can be measured by, among others, the standardized Herfindahl-Hirschman index. This index requires knowing its maximum value. So far, its maximum value has been identified with a distribution of competition scores, which we call *in cascade*. This distribution requires the scoring system to reward victory with twice as many points as ties. However, there are competitions such as the major soccer leagues that have a different remuneration pattern. Therefore, the *cascade distribution* does not generate the maximum value of the index. In this article, we show that: (i) solutions such as recalculating the results are not suitable; (ii) we redefine the distribution and (iii) we obtain the maximum index values for the major European soccer leagues. Thus, we can construct the standardized Herfindahl-Hirschman indexes for the 1997/98-2016/17 leagues and conclude that, overall, competitive balance has decreased.
Key words: competitive balance; european soccer leagues; herfindahl-hirschman index; cascade distribution; truncated-cascade distribution; perfectly unbalanced distribution.

Introduction

Competitive balance of a sports competition may be measured by different indexes (Borooah & Mangan, 2012; Eckard, 2001; Fort & Quirck, 1995; Gayant & Le Pape, 2012; Gayant & Le Pape, 2015; Horowitz, 1997; Humphreys, 2002; Pawlowski, Breuer & Hovemann, 2010; Schmidt, 2001; Schmidt & Berri, 2001; Utt & Fort, 2002). Andreff (2015); Larsen, Fenn & Spenner (2006); and Zimbalist (2002), among others, provided a broad panorama of the literature on the concept of competitive balance and its empirical application.

In a competition, there are aspects that influence the measurements of competitive balance. Therefore, the indexes should incorporate these characteristics, typical of sporting competitions. For example, the bilateral nature of the confrontations within the matches, prevents the monopoly case.

Therefore, the specific aspects of the sports competition can restrain the theoretical range of the indexes, leading to their standardization. Particularly, the Herfindahl-Hirschman index (HHI), which is a well-known concentration index and widely used in practice. Constructing the standardized Herfindahl-Hirschman index (HHI_{NORM}) requires knowing the minimum and maximum values of HHI (Owen, Ryan & Weatherston, 2007). The maximum value of HHI is generated by the distribution of results, which is called perfectly unbalanced distribution.

Measuring competitive balance is also affected by the scoring system used. Competitions such as the major European soccer leagues (Premier League in England, Primera División in Spain, Serie A in Italy, Bundesliga in Germany and Ligue 1 in France) have scoring systems that do not award twice as many points for winning as they do for ties. These leagues have a points award pattern $\{p_w, p_t, p_l\} = \{3, 1, 0\}$ (win, tie, loss).

In this case, the distribution of results used to obtain the minimum competitive balance is unstable because there are ties, and the total points at the end of the championship cannot be determined previously (Borooah & Mangan, 2012).

For this reason, it has been suggested that the results should be reconstructed with a *{2, 1, 0}* points award pattern (Gayant & Le Pape, 2015; Pawlowski, Breuer & Hovemann, 2010).

The aim of our work is to: (i) test the effects of reconstructing the league's results; (ii) find the maximum value of HHI for a competition, and (iii) determine the evolution of competitive balance in major European soccer leagues over the last two decades.

We show that reconstructing a league's results generates cardinal and ordinal negative effects that we discuss below. To test for such effects, we have reconstructed the scores for the major European soccer leagues over the period 1997/98-2016/17.

We prove that is necessary to redefine perfectly unbalanced distribution in order to construct a new one that generates the maximum level of concentration (minimum level of competitive balance): we call this

Corresponding Author: FRANCISCO TRIGUERO-RUIZ, E-mail: triguero@uma.es

Truncated-Cascade Distribution, which not involve recalculating the scoring based on the results, because it can be known previously.

We present a calculus and practical application of our proposal to the major European soccer leagues. This application constructs the competitive balance indexes for these leagues and over a considered time period. Thus, we can analyze the evolution of the competitive balance. This application will allow us to distinguish differences (i) between the leagues considered, and (ii) over the period under study. In addition, the results obtained in relation to the evolution of the competitive balance can be compared with other studies (Fort, 2007; Lenten, 2008; Michie & Oughton, 2004; Owen, 2012; Pawlowski, Breuer & Hovemann, 2010).

Material & methods

Competitive balance may be measured with the standardized Herfindahl-Hirschman Index (HHI_{NORM}), proposed by Owen, Ryan & Weatherston (2007):

$$HHI_{NORM} = \frac{HHI - HHI_{min}}{HHI_{max} - HHI_{min}}$$

where if there are N teams in the championship, the Herfindahl-Hirschman Index is:

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

with $s_i = \frac{p_i}{\sum_1^N p_i}$, and p_i are the points obtained by *i*-team at the final of championship. Note that $HHI_{NORM} \in [0,1]$.

HHI_{min} is the HHI value for the perfect competitive balance distribution of shares (PCB). This is representative of the maximum competitive balance (Borooah & Mangan, 2012; Gayant & Le Pape, 2015). In this distribution, the teams obtain the same number of points, and so the share is equal for all them:

$$s^{PCB} = \left(\frac{1}{N}, \ldots, \frac{1}{N}\right) \Rightarrow HHI_{min} = \frac{1}{N}$$

In s^{PCB}, the standard deviation of the shares is zero, and the HHI is equal to the inverse of the number of teams (Depken, 1999) and reaches its lowest level (Owen, Ryan & Weatherston, 2007).

Therefore, it is necessary to know the distribution that generates HHI_{max}, that is, the HHI value for the most unequal distribution. It should be noted that the bilateral nature of the confrontations within the matches prevents a monopoly solution.

The perfectly unbalanced distribution (i.e., the distribution that represents the most imperfect competitive balance) has been characterized by Fort & Quirck (1997), Gayant & Le Pape (2012 and 2015), Horowitz (1997), Larsen, Fenn & Spenner (2006), Owen, Ryan & Weatherston (2007), and Utt & Fort (2002), among others. In this distribution, each team has defeated all those below them and they have lost against all the teams above them. Once the championship is over and the positions have been ordered, the results of distribution can be visualized as being *in cascade* in which each team has fewer points than the team in the preceding position. Competitive balance value is minimum in this distribution.

Nevertheless, given N, if $p_w = 2\,p_t$, the total points at the end of the championship is constant. But, if $p_w > 2\,p_t$, $\sum_{i=1}^{N} p_i$ depends on the number of wins and ties (Borooah & Mangan, 2012). Therefore, for a {3,1,0} pattern, the *Cascade distribution* does not represent the perfectly unbalanced distribution. This is the case in the major European soccer leagues and in the annual UEFA Champions League, in which the pattern is {3,1,0}.

The proposed solution has been that the results should be reconstructed using a {2,1,0} pattern as a solution to measure competitive balance levels in leagues that use the {3,1,0} pattern. We have use the data from BDFUTBOL (2018,03,14) for the calculus and practical application.

Results

We have reconstructed the scores for the major European soccer leagues over the 1997/98-2016/17 seasons. The analysis of results of each league and season shows changes in the ranking of position in 96% of the competitions. Some examples are as follows:

- Of the 100 seasons analysed, only four remain unchanged after the results have been reconstructed. There is no change in the classification in the 2001/02 season of the German league; in the 2009/10 season of the French league; in the 2010/11 season of the Italian league; and in the 2016/17 season of the Spanish league.

- Real Madrid won the Spanish league in the 2006/07 season and Barcelona would have won it with a pattern {2,1,0}. Analogously, Lens and Metz in the 1997/98 season of the French league.

- Valencia qualified for the UCL in the 1998/99 season of the Spanish league, but with a {2,1,0} pattern Celta de Vigo would have. Similarly, Leeds and Liverpool in the 1999/2000 English league season; Olympique

Marseille and Sochaux in the 2002/03 French league season; Werder Bremen and Hertha Berliner in the 2004/05 Bundesliga season; and Fiorentina and Genoa in the 2008/09 Italian league season.

- Zaragoza were relegated in the 2007/08 season of the Spanish league, but with a pattern {2,1,0} Osasuna would have been. Similarly, Bologna and Parma in the 2004/05 season of the Italian league; Energie Cottbus and Borussia Mönchengladbach in the 2008/09 season of the Bundesliga; or Birmingham and Wolverhampton in the 2010/11 season of the English league.

Therefore, reconstructing the scores with a {2, 1, 0} pattern is not the solution because, as seen above, it generates cardinal and ordinal negative effects. Besides, it is not necessary, because we can calculate the maximum value of HHI.

We define a highly asymmetric final distribution of points when the champion has beaten the other teams, the runner-up has beaten all the teams below them in the final ranking, and so on up to a position (q) from which the teams have tied all their remaining matches. We will call it the *Truncated-cascade in (q) distribution*. If $q = N - 1$, we obtain the distribution that we call *Cascade distribution,* which constitutes a particular case of *Truncated-Cascade in (q) distribution*. Thus, the teams that occupied the first (q) positions won all matches except for those played against teams preceding them in the table, whereas the teams occupying $(N - q)$ positions tied all their matches (Avila-Cano & Triguero-Ruiz, 2018).

The values of HHI maximum obtained for N=18 (Bundesliga) and N=20 (Premier League, Primera División, Serie A and Ligue 1) with *Truncated-cascade in (q=7) distribution* for {3, 1, 0} pattern, are shown in Table 1. Note that these values are higher than the corresponding values for the {2, 1, 0} pattern. Therefore, when computing the competitive balance of a league HHI_{NORM}, it can lead to lower concentrations than those corresponding to the *Cascade distribution*.

Table 1. Values of HHI_{max}

	Truncated-Cascade	**Cascade**
	{3,1,0}	{2,1,0}
	q*=7	q*=19
N=18	0.0839378	0.0762527
N=20	0.0754015	0.0684211

We applied these issues to the major European soccer leagues in order to calculate the competitive balance levels. For each of these major leagues, we calculated the standardized HHI on the basis of the shares defined according to the current scoring system, {3, 1, 0}, and the number of teams in each season. So, Table 2 shows the levels of competitive balance, measured with HHI_{NORM}, in the leagues and the seasons considered.

Table 2. Evolution of competitive balance in the major European soccer leagues (Standardized HHI)

Season	Primera División (Spain)	Premier League (England)	Serie A (Italy)	Ligue 1 (France)	Bundesliga (Germany)	Minimum	Maximum	Mean	Error range (95%)
1997/98	0.1193	0.1058	0.2156	0.1275	0.0825	0.0625	0.2156	0.1292	0.0651
1998/99	0.1273	0.1395	0.1314	0.1586	0.1668	0.1273	0.1668	0.1447	0.0214
1999/00	0.0727	0.1357	0.1972	0.0559	0.1493	0.0559	0.1972	0.1302	0.0779
2000/01	0.1079	0.1353	0.1768	0.0906	0.0873	0.0873	0.1768	0.1194	0.0464
2001/02	0.0783	0.1929	0.1936	0.0949	0.1882	0.0783	0.1936	0.1496	0.0718
2002/03	0.1135	0.1593	0.1769	0.1016	0.1041	0.1016	0.1769	0.1311	0.0430
2003/04	0.1080	0.1618	0.2610	0.1351	0.1701	0.1080	0.2610	0.1672	0.0718
2004/05	0.1426	0.2068	0.1405	0.0781	0.1593	0.0781	0.2068	0.1453	0.0569
2005/06	0.1503	0.2293	0.2754	0.1327	0.1790	0.1327	0.2754	0.1933	0.0728
2006/07	0.1249	0.1751	0.2046	0.0793	0.1270	0.0793	0.2046	0.1418	0.0599
2007/08	0.1369	0.2693	0.1732	0.1170	0.1449	0.1170	0.2693	0.1683	0.0745
2008/09	0.1411	0.2287	0.1659	0.1538	0.1787	0.1411	0.2287	0.1736	0.0419
2009/10	0.2369	0.2259	0.1607	0.1568	0.1546	0.1507	0.2369	0.1850	0.0529
2010/11	0.1868	0.1654	0.1553	0.1011	0.1291	0.1011	0.1868	0.1575	0.0423
2011/12	0.1916	0.2075	0.1567	0.1470	0.1954	0.1470	0.2075	0.1797	0.0326
2012/13	0.2113	0.2230	0.2047	0.1275	0.2183	0.1275	0.2230	0.1970	0.0490
2013/14	0.2252	0.2464	0.2566	0.1826	0.2477	0.1826	0.2566	0.2317	0.0370
2014/15	0.3009	0.1824	0.1888	0.1493	0.1582	0.1493	0.3009	0.1959	0.0756
2015/16	0.2232	0.1671	0.2037	0.1605	0.2046	0.1605	0.2232	0.1918	0.0333
2016/17	0.2881	0.2643	0.2879	0.2073	0.1634	0.1634	0.2881	0.2422	0.0683
Minimum	0.0727	0.1068	0.1514	0.0559	0.0825				
Mean	0.1643	0.1904	0.1958	0.1276	0.1604	0.1186	0.2247	0.1677	-
Maximum	0.3009	0.2693	0.2879	0.2073	0.2477	0.1826	0.3009	0.2422	-
Error range (95%)	0.0307	0.0220	0.0209	0.0179	0.0194	-	-	-	-

Besides, Table 2 shows that the differences in the levels of competitive balance are significant both between leagues and over the seasons. The HHI_{NORM} values out of the range from the average value and the