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)	Mínimum	Máximum	Mean	Error margin (95%)
1997/98	0.1193	0.1058	0.2156	0.1235	0.0825	0.0825	0.2156	0.1292	0.0661
1998/99	0.1273	0.1395	0.1314	0.1586	0.1668	0.1275	0.1668	0.1447	0.0214
1999/00	0.0727	0.1757	0.1972	0.0559	0.1495	0.0559	0.1972	0.1302	0.0779
2000/01	0.1079	0.1393	0.1768	0.0996	0.0873	0.0873	0.1768	0.1194	0.0464
2001/02	0.0783	0.1929	0.1936	0.1049	0.1082	0.0783	0.1936	0.1196	0.0718
2002/03	0.1835	0.1593	0.1769	0.1016	0.1041	0.1016	0.1769	0.1311	0.0410
2003/04	0.1080	0.1608	0.2610	0.1351	0.1701	0.1080	0.2610	0.1672	0.0718
2004/05	0.1426	0.2086	0.1405	0.0788	0.1595	0.0781	0.2658	0.1455	0.0569
2005/06	0.1503	0.2193	0.2784	0.1327	0.1790	0.1327	0.2754	0.1943	0.0728
2006/07	0.1249	0.1731	0.2046	0.0793	0.1270	0.0755	0.2506	0.1418	0.0599
2007/08	0.1169	0.2493	0.1732	0.1170	0.1449	0.1170	0.2693	0.1683	0.0745
2008/09	0.1401	0.2487	0.1659	0.1538	0.1787	0.1411	0.2387	0.1706	0.0419
2009/10	0.2369	0.2159	0.1507	0.1568	0.1496	0.1507	0.2569	0.1850	0.0529
2010/11	0.1868	0.1154	0.1553	0.1011	0.1291	0.1011	0.1868	0.1575	0.0423
2011/12	0.1936	0.2075	0.1567	0.1470	0.1954	0.1470	0.2075	0.1797	0.0335
2012/13	0.2113	0.2130	0.2047	0.1275	0.2183	0.1275	0.2250	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	0.1982	0.1493	0.3009	0.1959	0.0276
2015/16	0.2282	0.1671	0.2067	0.1695	0.2046	0.1605	0.2652	0.1918	0.0333
2016/17	0.2881	0.2643	0.2879	0.2079	0.1694	0.1694	0.2681	0.2422	0.0663
Mean	0.1683	0.1904	0.1963	0.1376	0.1694	0.1186	0.2267	0.1677	-
Maximum	0.3009	0.2583	0.2879	0.2079	0.2477	0.1826	0.3009	0.2422	-
Minimum	0.0727	0.1058	0.1314	0.0559	0.0825	0.0659	0.1668	0.1194	-
Error margin (95%)	0.0807	0.0830	0.0809	0.0479	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 margin error, are bolded or highlighted. The highlighted values are used for the analysis of the seasons. We can

see that in 15 of 20 seasons there are some leagues outside the margin of error. The seasons in which there are no leagues out off the margin of error are: 1999/00, 2001/02, 2009/10, 2011/12, and 2015/16.

The bold values are used for the analysis of leagues. We can see that 64% cases have values outside the margin of error around the average.

Table 3 indicates that these results were shown to be significant under an ANOVA analysis for each factor for both seasons and leagues at 99% (p-value 0.00(***)).

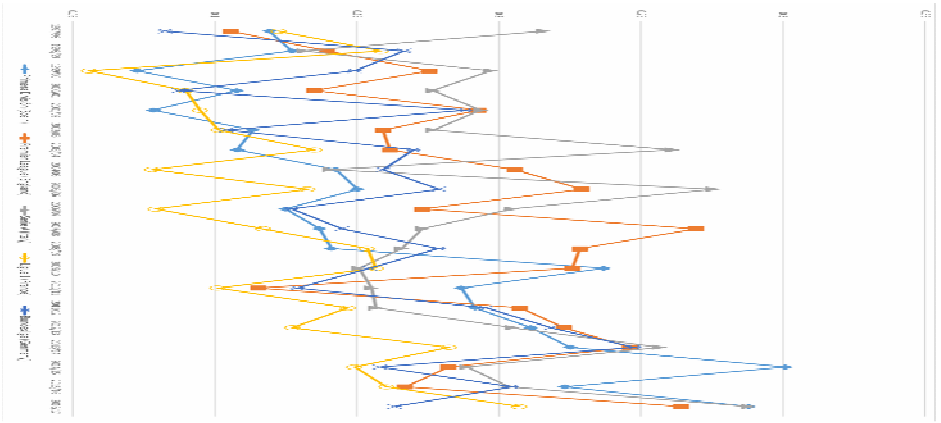
Table 3. Analysis of Variance (ANOVA): Two-Factor Without Replicaction

Source of Variation	*Sum of Squares*	*Degrees of freedom*	*Mean of squares*	*F*	*P-value*	*F crit.*
Seasons	0.112590804	19	0.005925832	4.112978174	5.0296E-06	1.725029095
Leagues	0.059650806	4	0.014912702	10.35054963	9.58532E-07	2.492049297
Error	0.109498081	76	0.001440764			
Total	0.281739692	99				

Therefore, we can see that the major European soccer leagues have, on average, a level of concentration of 0.1677. They have a maximum value (0.3009) in the 2014/15 season of the Spanish Primera Division and a minimum value (0.0559) in 1999/00 season of the French Ligue 1.

The Italian and English leagues have globally the lowest degree of competitive balance. Their average values of HHI_{NORM} are greater than 0.19. For its part, the French league presents, on average, the highest level of competitive balance (HHI_{NORM}=0.13). The Spanish and German leagues have similar and intermediate levels (HHI_{NORM} ≈0.16).

Figure 1 shows the evolution of competitive balance in the major European soccer leagues using *{3,1,0}* and *Truncated-cascade distribution*. Competitive balance is measured with HHI_{NORM}. Figure 1 allows us to appreciate the tendency to increase concentration and decrease competitive balance.



Fig, 1. Evolution of competitive balance in the major European soccer leagues (HHI_{NORM}).

Table 4 includes the main estimated parameters of the temporary regressions made for each league.

Table 4. Regression data (Figure 1).

		Primera División (Spain)	Premier League (England)	Serie A (Italy)	Ligue 1 (France)	Bundesliga (Germany)
Time	Coefficient	0.009744	0.004254	0.001810	0.003965	0.003953
	Standard error	0.001239	0.001582	0.001729	0.001204	0.001369
	p-value	0.000***	0.015**	0.309	0.004***	0.010***
Constant	Coefficient	0.062025	0.145751	0.176844	0.085946	0.118924
	Standard error	0.014841	0.018947	0.020714	0.014425	0.016395
	p-value	0.001***	0.000***	0.000***	0.000***	0.000***

The p-values are significant at the 99% for the Spanish Primera Division, French Ligue 1 and the German Bundesliga, and at the 95% for the English Premier League. In the Italian Serie A, the temporary decline of competitive balance is not significant.

In summary, we can affirm that, with the Italian exception, the levels of competitive balance have decreased in the major European soccer leagues over the last two decades, while there are different levels between leagues.

Discussion & Conclusions

Competitive balance can be measured by standardized indexes, such as the HHI, which require knowing their minimum and maximum values. In this case, the minimum value is $HHI_{min}=1/N$, but the maximum is not unambiguously identified.

The distribution of results generated by the HHI_{max} is unstable if there are ties and the scoring system does not comply with the condition of rewarding wins with twice as many points as ties. The solution proposed so far has been to reconstruct the results with a pattern that complies with this condition. In particular, the reconstruction of results can be found in Fort (2007), Gayant & Le Pape (2015), Owen (2012) and Pawlowski, Breuer & Hovemann (2010).

We have verified that the recalculation of HHI using a scoring system other than the real one, causes severe changes, both in cardinal and ordinal terms. Therefore, this solution is not desirable.

We have constructed the distribution of results that allows us to achieve the maximum value of HHI. It is a highly asymmetric distribution that we call *Truncated-cascade distribution*.

We obtain the highest HHI values for the major European soccer leagues and in the 1997/98-2016/17 seasons. We can then calculate the corresponding standardized HHI values.

In this way, we can analyze the evolution of competitive balance, and conclude that the differences between leagues and seasons are statistically significant. The analysis of the evolution of competitive balance in the major European leagues has been also carried out by Michie & Oughton (2004). These authors used indexes other than the standardized HHI and concluded that, between 1954 and 2004, there were different trends among the leagues considered. However, these authors do not use a statistical test for it.

In addition, these authors have found increased trends in the evolution of the English, Italian and German leagues. In the French league there is no clear trend for the whole period, although it has been increasing since the 1990s. For the Spanish league, they conclude that it has a cyclical behavior with a moderate decline in competitive balance since the 1970s.

On the other hand, Pawlowski, Breuer & Hovemann (2010) also obtained an increase in concentration in the English, Italian and German leagues between the 1992/93 and 2007/2008 seasons. By contrast, the competitive balance of the French league would increase, while in the Spanish league the result depends on the index used.

In any case, it should be noted that, without prejudice to the calculations made by these authors, given that these values have not been standardized, their comparability is not feasible. Of particular relevance is the fact that the maximum value that the index can reach is not taken into consideration. This problem is solved, in our case, by applying HHI_{NORM} with the maximum effective value of each competition. This standardization makes its values vary between 0 and 1. Besides, unlike the previous authors, we use statistical tests to prove the significance of the regression trends.

In this article we can see in the1997/98-2016/17 seasons, the tendency to increase concentration and decrease competitive balance. Particularly, we can affirm that, with the Italian exception, the levels of competitive balance have decreased in the major European soccer leagues over the last two decades.

Conflicts of interest

The authors do not have any conflicts of interest to declare.

References

Andreff, W. (2015). *Disequilibrium sports economics. Competitive imbalance and budget constraints*. Cheltenham, UK: Edward Elgar Publishing Ltd.

Avila-Cano, A., & Triguero-Ruiz, F. (2018). The distribution of soccer leagues scores that generates the minimum of competitive balance: Truncated-Cascade Distribution. Málaga Economic Theory Research Center. *Working Papers*. WP 2018-4. ISSN 1989-6808.

BDFUTBOL (2018,03,14). Retrieved from https://www.bdfutbol.com/

Borooah, V.K., & Mangan, J. (2012). Measuring competitive balance in sports using generalized entropy with an application to English premier league football. *Applied Economics*, 44(9), 1093-1102, doi: 10.1080.00036846.2010.537638

Depken, C.A. (1999). Free-agency and the competitiveness of major league baseball. *Review of Industrial Organization*, 14, 205-217.

Eckard, E.W. (2001). Free agency, competitive balance and diminishing returns to pennant competition. *Economic Inquiry*, 39(3), 430-443.

Fort, R., & Quirck, J. (1995). Cross-subsidization, incentives and outcomes in professional team sports leagues. *Journal of Economic Literature*, 33, 1265-1299.

Fort, R. (2007). Comments on measuring parity. *Journal of Sports Economics*, 8(6). 642-651. doi: 10.1177/1527002506298122

Gayant, J.P., & Le Pape, N. (2012). How to account for changes in the size of Sports Leagues: The Iso Competitive Balanced Curves. *HAL archives-ouvertes*, halshs-00856122.

Gayant, J.P., & Le Pape, N. (2015). The metrics of competitive imbalance. In A. Andreff (ed.), *Disequilibrium sports economics. Competitive imbalance and budget constraints*. (pp. 104-130). Edward Elgar Publishing Ltd., Cheltenham, UK.

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

Humphreys, B.R. (2002). Alternative measures of competitive balance in sports leagues. *Journal of Sports Economics*, 9, 191-210.

Larsen, A., Fenn, A.J., & Spenner, E.L. (2006). The impact of free agency and the salary cap on competitive balance in the National Football League. *Journal of Sports Economics*, 7(4), 374-390, doi: 10.1177/1527002505279345

Lenten, L.J.A. (2008). Unbalanced schedules and the estimation of competitive balance in the Scottish Premier League. *Scottish Journal of Political Economy*, 55(4), 488-508.

Michie, J., & Oughton, C. (2004). Competitive balance in football: Trends and effects. *Research Paper*, 2. Football Governance Research Centre, Birkbeck College, London.

Owen, P.D., Ryan, M., & Weatherston, C.R. (2007). Measuring competitive balance in professional team sports using Herfindahl-Hirschman Index. *Review of Industrial Organization*, 31, 289-302, doi: 10.1007/s11151-008-9157-0

Owen, P.D. (2012). Measuring parity in sports leagues with draws: Further comments. *Journal of Sports Economics*. 13(1), 85-95. DOI: 10.1177/1527002510393738

Pawlowski, T., Breuer, C., & Hovemann, A. (2010). Top clubs´ performance and the competitive situation in European domestic football competitions. *Journal of Sports Economics*, 11(2), 186-202, doi: 10.1177/1527002510363100

Schmidt, M.B. (2001). Competition in Major League Baseball: The impact of expansion. *Applied Economic Letters*, 8, 21-26.

Schmidt, M.B., & Berri, D.J. (2001). Competitive balance: The case of Major League Baseball. *Journal of Sports Economics*, 2(2), 145-167.

Utt, J., & Fort, R. (2002). Pitfalls to measuring competitive balance with Gini coefficients. *Journal of Sports Economics*, 3(4), 367-373, doi: 101177/152700202237502

Zimbalist, A. (2002). Competitive balance in sport leagues: an introduction. *Journal of Sports Economics*, 32(2), 111-121.