The competitive balance of French football 1945-2002,

published in: *Économie Appliquée* LVI, 2003, nr. 4, 91-113.

Jan Groot and Loek Groot*

* SISWO/ Netherlands' Institute for the Social Sciences and Amsterdam School of Social Research (ASSR), University of Amsterdam

One of the most important issues in the scientific literature on sports is that of the competitive balance (CB for short). In this paper we present a new index, named the surprise index, with which to measure the CB. It shows that the French 'Championnat' football is quite well balanced. The surprise index has several comparative advantages vis- à-vis more conventional indices of CB. Firstly, it can easily be used to compare leagues over time, between different countries and even between different sports. Secondly, it enables us to say something about the optimal number of teams in a league. These advantages arise because of the way the CB is measured by the surprise index compared to more familiar CB-indices like the standard deviation or the concentration ratio. Whereas the latter only use the statistics of the final league table at the end of the season, the surprise index makes use of much more detailed information. The results of all matches within a competition enter into the formula. Moreover, it gives a relatively high weight to surprising outcomes within the league competition, whereas the saliency of these matches is lost when computing the standard deviation or the concentration ratio.

Un des sujets les plus importants dans la litérature des sports est celle de la balance compétitive (CB). Dans cet article nous présenterons un nouveau index que nous appellerons l'index de surprise. L'index de surprise se distingue des indicateurs du BC plus conventionnels par plusieurs avantages. D'abord, il permet a comparer des compétions dans le temps et dans des sports et pays différents. Ensuite il nous permet a calculer le nombre optimal des équipes participant dans une compétition. En comparaison avec les indicateurs du BC plus conventionnels, l'index de surprise se base sur plus d'informations, et par ailleurs fait compter des scores surprenants.

Matches between teams of very unequal strength are rarely interesting. However, sometimes a surprise happens, e.g. when against all odds the underdog wins against the towering favourite. For the index presented here, these surprising outcomes are salient: the more such outcomes, the higher the index and the higher the degree of CB. Many people do believe that due to the commercialization of football the CB is declining. To test assertions of this sort it is very important that the CB can be measured adequately. Well known indices are the standard deviation and the concentration ratio.[1] The more balanced a competition, the lower the standard deviation and the concentration ratio. These indices are very rough, to say the least. As will be shown in section 3, our index gives a much more precise estimate of the degree of superiority of the top teams. In a straightforward way, the index can also be used to measure the degree of inferiority of the worst teams within a league. Combining both statistics gives the opportunity to make statements about the optimal size of a league, where optimal refers to the number of teams which yields the highest balance of the competition. The highest (attainable) level of CB differs from the *optimal* level of CB. The latter depends on the distribution of the fan population over teams: teams from large cities with a large fan base must win more often than teams from thinly populated areas. However, even from the point of view from the optimal CB there is a limit to the extent of superiority of the more popular teams, because the outcome of the contests between the top teams and the worst teams within a league still must be uncertain. The study of CB is also used to investigate which measures (varying from revenue sharing, reverse draft to salary caps) are more conduceive to improve the CB. For a recent overview on how to measure CB and a discussion on the usefulnes of different measurements, see the symposium 'Competitive Balance in Sports Leagues' in the *Journal of Sports Economics* [2003, vol. 3, no. 2].

1. -The surprise index

For simplicity, consider a competition between three teams where team A is the all time champion, team B the eternal runner-up and C the team which traditionally brings up the rear. As is customary in most team sports, each team plays two games, home and away, against any other team. Table 1 summarizes the results, for four different situations, indicated by the bold Roman numbers in the top left corner.

Given the assumed rank order of A>B>C, situation I comes up to one's expectations. Team A wins all its games against B and C, and B wins both games against C. No outcome is surprising, and this might be because A is much better than B and C, and B much better than C. In situation IV only the results of the games between A and C are surprising, one draw and

---

[1] The concentration ratio C_j is defined as the actual number of match points, 2 for a win and 1 for a draw, collected by the top j teams divided by the maximum number of points they could have won.

a win for C. The games of B against A and C end as expected. In situation IV we cannot say anymore that A is unambiguously much better than B and C and therefore the competition as described by situation IV seems to be more in balance than that of situation I.

Table 1. Home and away chart (italicized the results of the matches which enter the surprise index *S*, bold the number of realized surprised points *P*)

I	A	B	C	P
A	x	2-0	2-0	**0**
B	0-2	x	2-0	**0**
C	0-2	0-2	x	**0**

II	A	B	C	P
A	x	2-0	2-0	**0**
B	*1-0*	x	2-0	**2**
C	*1-0*	*1-0*	X	**6**

III	A	B	C	P
A	x	*2-2*	*2-2*	**3**
B	*2-2*	x	*1-1*	**2**
C	*1-1*	*1-1*	x	**3**

IV	A	B	C	P
A	x	2-0	*0-1*	**4**
B	0-2	x	2-0	**0**
C	*1-1*	0-2	X	**2**

Consider now situations II and III, in which the balance of the competition is at maximum. Teams B and C end up equal to team A, and to draw up the league table, taking into account that the goalbalance is zero for each team, one has to resort to the number of goals for or against. Apparently, all teams were equally strong. Since we cannot imagine a more exciting, thrilling league as described in situation II (or III), we can say that these competitions are in perfect balance. Now the rule applied to calculate the number of surprise points within a competition is simply determined by evaluating all matches with a surprising outcome with one point for a draw and two points for a win, multiplied by the rank order difference between the teams. Denoting i and j ($i<j$) as the rank number of the teams in the final league table at the end of the season, a win of team j against i is thus evaluated at (j-i)*2, and a draw at (j-i)*1. In other words, the weight attached to a surprising outcome of a match is simply the rank order difference (j-i). One team beating another just one place higher in the league table is far less surprising than when the team at the bottom of the list (with rank order N) beats the champion (with rank order 1). The former gives only two surprise points, the latter (N-1)*2 points.

The denominator of the index is defined by the weighted sum of the surprising outcomes in a counterfactual perfectly balanced competition, as described in situations II and III of Table 1. In situation II there are three surprising matches (see the italicized numbers). Applying the above mentioned rule gives 2 points for the game between B and A, and 4 and 2 points for the home wins of C against A and B, respectively. In situation III all 6 games end up in draws, and 4 are evaluated against 1 point because the rank order difference is just 1, and the other 2, with a rank order difference of 2, get 2 points each. In both perfect balanced competitions II and III, the number of surprise points, M, is at maximum and equal to 8. The surprise index (S) is simply the ratio of the number of the *realized* surprise points (P) and the *maximum* number of surprise points M. Clearly, in situations II and III the index S is equal to 1, whereas in situation I, S is equal to 0. It can easily be verified that for situation IV, with a draw and a win by C on A both weighted by a rank order difference of 2, the index equals (2+4)/8=0.75. Thus, the surprise index varies between 0 and 1. It is zero in a perfectly unbalanced competition, like in situation I: the champion wins all its games, the runner-up wins all its games except those against the champion, etcetera and the last team loses all its games. Seen from the final league table, no match was a surprise. The index is equal to one in a perfectly balanced competition, that is, if, like in situation II and III, it is really a competition *inter pares*. In between, the competition is more (un)balanced the closer the index approaches one (zero).

The rule allocating points to individual matches can be formally expressed as follows. Let R_{ij} denote the result of the match of team i at home against team j, where R_{ij} or R_{ji} is 2 if j wins, R_{ij} or R_{ji} is 1 if it is a draw, and R_{ij} or R_{ji} is 0 if i wins (of course, R_{ji} denotes the outcome of the match with team j at home against team i). Given that (j-i) is the rank order difference,

(1) $S = \frac{P}{M} = \frac{1}{M} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (R_{ij} + R_{ji}) \, (j-i)$

with the maximum score of surprise points (*M*) in a perfectly balanced competition of *N* teams given by:

(2) $M = 2\sum_{i=1}^{N-1} (N-i)i = (N-1)\, N(N+1)/3$

Intuitively, the expression for *M* can be understood as follows. We have seen that the competition as described by situation II in which each team both wins and loses a match against all other teams is perfectly in balance, with $S = 1$, so the numerator *P* of *S*, must be equal to *M*. The highest contribution to the numerator of *S* occurs when team *N* at the bottom of the list wins against the number 1, resulting in 2*(*N*-1) points, and in a situation II this happens only once. There are two matches with each a contribution of 2*(*N*-2), to wit team *N* wins from team 2, and team *N*-1 wins from team 1. There are three matches with each a contribution of 2*(*N*-3), namely a win of team *N* against team 3, a win of team *N*-1 against team 2 and a win of team *N*-2 against team 1. And so on down the line. The summation leads to the expression at the RHS of Equation (2). Situation III where all matches end in draws gives the same score for *M*: there are two matches with a rank order difference of *N*-1, there are 2 times 2 matches with a rank order difference of *N*-2, there are 2 times 3 matches with a rank order difference of *N*-3, etcetera, resulting in the same formula for *M*.

The index can be used to compare competitions over time (as Figure 1 illustrates, see below) or to compare competitions between countries. For instance, it would be interesting to investigate whether the French *Championnat* competition is more in balance than the Spanish *Primera Division*, the German *Bundesliga* or the Italian *Serie A*, or whether the competitions of great football countries are more balanced than those of football dwarfs like Scotland, Denmark, Sweden, Norway and Belgium. It goes beyond the scope of this paper, but in principle it is possible to use the surprise index to compare American football, baseball and basketball *as competitions*. If it appears that the index for one of those sports is significantly higher than for the others, then it becomes interesting to see whether this can be explained by special measures (e.g. salary or payroll caps, television and gate revenue sharing, reverse order draft in the rookie pick, etcetera) taken to enhance the balance of the competition. First we will make up the balance sheet of French football over the period 1945-2002. Next we concentrate on what we think is the most attractive property of our index, namely that it can be used to determine the optimal size of the competition.

2. -The competitive balance over the period 1945-2002

In interpreting the results two facts have to be taken into account. Firstly, the number of teams within the *Championnat* varies between 18 and 20 (see the fact sheet in Appendix C, which also contains in the last but one row the statistics for the last season 2002-2003, but not included in the figures below). Because both the numerator and the denominator of the index vary with the number of teams (see Eqs. (1) and (2) above), this has by itself no effect on the value of the index. Secondly, from the mid 1990s onwards, three instead of two match points are assigned to a victory. However, to avoid complications we stick to the rule that a victory counts for two points over the entire period under consideration.

Figure 1 about here

Figure 1 presents the CB index for the period 1945-2002. On average the index is 68%, with the seasons 1954/55, 1963/65, 1974/75, 1987/88 and 1999/00 as positive outliers with an index of 75% or more, and 1959/60 and 1979/80 as negative outliers with values below 60%. Given that the index can vary between 0 and 100%, an average value of 68% is quite good. To give an idea, the average value of the index for the Dutch football competition over the same time period is only 54%.[2] Figure 1 also shows only a very slight decrease in the CB over the last quarter of a century. Good news for those who feared that the ongoing commercialization of football has had a devastating effect on the balance of the competition. Although the decrease over the entire period is less than 3%-points, from 69.2 to 66.4%, basically it suggests that the CB of French football is steadily albeit slowly deteriorating.

Our index has a strong correlation with other indices commonly used to measure the CB.[3] The correlation coefficient between our index and the standard deviation is more than (minus) 90% and with the C_3 ratio it is more than (minus) 80%. The correlations are negative,

[2] This illustrates that the index can be used to compare different leagues, in this case the *Championnat* with the Dutch competition, even though the latter has less teams.

[3] Another index recently used to measure CB is the Gini coefficient in winning percentages. Utt and Fort [2002] adjust the Gini coefficient for the fact that in the most unequal outcome the best team cannot win all games within a league, but only its own games.

because a high value of the surprise index, indicating a balanced competition, corresponds to a low value for the standard deviation or the C_3 ratio. For example, in the extreme case of a perfectly balanced competition, the surprise index is 1, the standard deviation is zero and the C_3 index[4] tends (for large N) to its lower boundary of 50%. To make the latter two indices comparable with the surprise index a linear transformation is required, so that the transformed indices vary in the same way between 0 and 1 as our CB-index S. Denoting STD^u as the standard deviation of a completely unbalanced competition, the expression for the transformed standard deviation is:

$$STD^* = 1 - \frac{STD}{STD^u}$$

with

$$STD^u = 2\sqrt{N(N+1)/3}$$

It can easily be checked that STD^* varies between 0 and 1, corresponding to a completely unbalanced and a perfectly balanced competition respectively. In the same way C_j^* is defined as:

$$C_j^* = \frac{2(N-2)}{(N-3)}(1-C_j)$$

Strikingly, Figure 2 shows that the concentration ratio, based on only three numbers per year, gives almost the same information as the standard deviation, requiring information about N numbers, and the surprise-index, based on the results of all matches, so requiring $N(N\text{-}1)$ numbers. From the perspective of parsimony the concentration ratio is unambiguously the best index of CB. That our index conveys almost exactly the same pattern as when using the standard deviation or concentration ratio shows that the surprise index is as good as these more conventional measures of the CB.[5]

---

[4] The expression for the concentration ratio is $C_j = P/(2j(2N\text{-}j\text{-}1))$, where P denotes the number of match points collected by the top j teams at the end of the season, 2 points for a win and 1 for a draw. In a competition completely out of balance, C_j is equal to 1 for all j. In a perfectly balanced competition C_j is equal to $(N\text{-}1)/2(N\text{-}2)$ for all j, which for large N tends to its lower boundary of 0.5.

[5] For illustratory purposes Figure 2 only depicts the period 1947-1958.

As said, the standard deviation and the concentration ratio only make use of the global information given by the league table at the end of the season, whereas our index uses information of all matches which together make up the cross table of the competition. Since such detailed information, the so called grid results or cross tables, is easy obtainable, over a long period of time, it is a pity that it is never used to obtain a more accurate measure of CB. As a matter of fact, these cross tables are made for over 100 years, but nobody ever saw that these tables hide a very useful index of CB. Apart from that, we do not say that an index which uses more detailed information than another index using more aggregate data, is therefore always better. It is also important that the more detailed information is included in the index in an adequate way.[6] As we will show in the next section, the lack of parsimony of our index has a major advantage: the overall index of CB can be decomposed into an index of CB *per team*. Using that information, some assertions can be made about the optimal size of the competition.

3. -The optimal size of the competition

In principle, whether a competition is balanced or not has nothing to do with the number of teams. A competition of two teams can be completely out of balance if one team is far better than the other, while a competition with many teams, as the French *Championnat*, can have a high CB. However, a difference in quality between the best teams and the worst teams which is too big will impair the CB. If for instance the *Championnat* and the Second Division would be merged, then it is very likely that the CB-index will drop instantly and significantly. In the same vein, splitting the present *Championnat* into two competitions, the top half and the bottom half, might give two more balanced competitions, but not necessarily. The optimal size of the competition need not be exactly half of the number of teams now playing in the *Championnat*, given that the optimal size as defined here refers to the number of teams that maximizes the average CB over time. It need not be two teams either. This would only be the case if both teams were of equal strength, and would maintain that year after year. As soon as

[6] One might argue that the rule to use (the first power of) the rank order difference as weights to surprising outcomes is somewhat arbitrary. Indeed, one could also chose to use as weights the rank order difference raised to the square root, the square, or the third power. In general, the higher the power of the rank order difference used as weights, the more important the surprising outcomes of the games between teams with high rank order differences become. This can easily be seen by assuming that no weights at all are used, which is equivalent to using as weights rank order differences with power zero. In that case, all surprising outcomes are equally

one team becomes better than the other and wins both matches, the CB-index for that season falls to 0%.

Summarizing, *a priori* not much can be said about the optimal size of a league. Therefore, it cannot be ruled out that a higher number of teams than at present in the *Championnat* is the optimal size. However, as we will argue later on, it is more likely that the optimal number of teams is less than that it is higher than the present number of teams in the *Championnat*. Note that optimal is taken here in a special sense, related to achieving a higher level of the CB: one could also argue that the optimal number of teams is the one that maximizes welfare (see e.g. Koning [2000])). According to this line of argument the optimal CB need not be a perfectly balanced competition or the highest attainable level of CB, e.g. when clubs differ in their fan support, welfare maximation requires a higher chance to become champion for strongly supported teams. Fort and Maxcy [2002, p. 155] label this line the "uncertainty of outcome hypothesis (UOH) literature", analysing the effect of CB on fan interest, whereas the literature on measuring CB proper and what has happened over time with the CB is labelled as "the analysis of competitive balance (ABC) literature itself".

As one can see from the fact sheet in Appendix C, the number of teams in the *Championnat* varies over time. Can something be said about the optimal number of teams, seen from the perspective of maximizing the CB? A first clue can be obtained by comparing the CB-index in the seasons with 20 teams with the CB-index in the seasons with only 18 teams, which is 67% against 69%. The smaller competitions have on average a slightly higher value for the CB-index, but the difference is not big. A second clue can be obtained by focusing on the matches between the top 3 teams and the bottom 3 of the list in each season. This is illustrated in Figure 3. The trend is sloping slightly downwards. Over in total 57 seasons, the CB-index for the top 3 against the bottom 3 has decreased with only 3.4%-points from 32.3% to 28.9%. To fix ideas, in a competition with 18 teams a value of 31% is already achieved if each of the top 3 teams just loses one game against the worst three teams and all other 15 games are won (the index is then at least equal to (15*2+14*2+13*2)/270 = 0.31 where 270 corresponds to the value for M in the slimmed down competition of only these 6 teams whilst maintaining their rank numbers 1, 2, 3, 16, 17 and 18). The downward trend

important, irrespective of whether the worst team wins a game against the number 1 or wins a game against the last but worst team.

suggests that the matches between the top and the bottom of the list in a league of 22 teams become slightly more and more just walkovers.[7]

Figure 3

The low figure for the index of the top teams against the worst teams can either be due to the top teams being too good, or the worst teams too bad. Under the former, a higher CB would result if the top teams would promote to, say, a European superleague; under the latter a higher CB would result if the worst teams would be eliminated from the competition. First we investigate the hypothesis that the number of teams in the league is too high, because the worst teams contribute too little to the balance of the competition. This possibility can easily be illustrated by following the thought experiment of what would happen if we allow a junior team to participate in the *Championnat*. For certain, this team would lose all matches and the overall CB-index would drop instantly by more than 25%-points.[8] Therefore we can say that for the optimal CB it would be better if the junior team participates in another competition. In the same vein we can check whether the worst teams in the *Championnat* contribute sufficiently to a balanced competition, that is, realize sufficient surprising outcomes. If that is not the case, a higher CB can be achieved by reducing the number of teams. Although there will by definition be worst teams in a league, the question here is whether they are 'worthy' number lasts.

The CB-index *per team* can be calculated by dividing the number of realized surprise points of team k by the maximum number M_k in a perfectly balanced competition. It can easily be checked that for a team with rank k, M_k is equal to k*(k-1). The CB-index for team k, S_k, is:[9]

$$(3)\ \ S_k = \frac{1}{k(k-1)} \sum_{i=1}^{k-1} (R_{ik} + R_{ki})\,(k-i)$$

with the overall index S equal to:

[7] The average value of 30% for the index of the top 3 against the last 3 is somewhat higher than in The Netherlands, with an average value of 22%.

[8] Suppose the value of the index before the junior team enters the competition is S. Because of the entrance of that team the number of teams in the league increases by one from 18 to 19 and the denominator of the CB-index increases from 1938 to 2660 (see Eq. 2). Since the numerator does not change (the new team does not realize any surprising result), the new index is equal to 1938*S / 2660 = 0.73*S.

[9] Applying Equation (3) to situation IV of Table 1 gives S_2 = 0 and S_3 = 1.

(4) $S = \frac{1}{M}\sum_{k=2}^{N} M_k S_k$

Suppose that the overall index S is 10%, then the worst team N is a 'worthy' worst team if the CB-index for that team, S_N, is about as high as S. For instance, if the worst team would win a match against the number 1, realizing in one stroke 2*(N-1) points which is exactly one tenth of M_N if N=20 and lose all its other matches, it is a worthy worst team because S_N is also equal to 10%. Of course, if the overall index S is higher, then S_N also needs to be higher in order for the worst team to qualify as a 'worthy' worst team in that competition. In the limit, if S approaches one indicating an extremely balanced competition, then the worst team must be as good as all its rivals so that S_N also approaches one. Equation (4) expresses that the overall index S is a weighted sum of the CB-indices per team, with M_k / M as weights.

The first row of Table 2 shows the average CB-indices per team for the worst three teams over the years in which there were 20 teams in the league, pitched against the average of the overall index S for these same years. The CB index for the team with rank 20, S_{20}, is on average 51%, which is 18%-points below the overall average of 67%. S_{19} is equal to 59%, which is also only 9%-points less than the overall average. S_{18} is equal to 68%, which is even higher than the overall average. The balance in the competition could therefore be higher if the number of teams would be restricted to only 19 teams. A further reduction of the number of teams would not lead to a higher CB. However, there is a disturbing result that goes strongly against this conclusion. One would expect that in the seasons with only 18 teams in the league the average value of the overall index would be considerably higher than in the seasons with 20 teams, but the difference is only 2%-points (see the second column of Table 2). The CB-indices per team of the worst three teams (S_{18}, S_{17} and S_{16}) in the slimmed down competitions of 18 teams are somewhat higher than the corresponding indices S_{20}, S_{19} and S_{18} in the bigger competitions of 20 teams. If we would focus on the competition of 18 teams (see the second row of Table 2), we would conclude that a competition restricted to 17 teams could be justified. This is puzzling.

Table 2. The average of the overall index versus the indices per team (in %)

N	S	S_{20}	S_{19}	S_{18}	S_{17}	S_{16}	S_{15}	S'_1	S'_2	S'_3	S'_4
20	67	51	59	68				49	57	63	67

18	69			54	66	67	71				

The same kind of reasoning can be applied to the best teams within a competition. The idea here is the mirror image of the exercise above. The champion cannot, by definition, realize any surprise points,[10] it can only give them away. If the competition is balanced, then this will occur very often. If the best team is clearly superior to the rest, then it will give very few opportunities to other teams to score surprise points, which is not conducive for the overall CB. In a perfectly balanced competition, as in situation II of Table 1, where each team wins and loses one match against all other teams, the (maximum) number of points given away (M_i') by a team of rank i to teams with rank $j>i$ is equal to $(N-i)(N-i+1)$. To obtain an index S_i' for the degree of superiority of the best teams, the number of points actually given away, call this P_i', must be divided by M_i', so:

$$(5)\quad S_i\text{'} = \frac{P_i\text{'}}{M_i\text{'}} = \frac{1}{M_i\text{'}} \sum_{j=i+1}^{N} (R_{ij} + R_{ji})\,(j-i)$$

with, analogous to (4), the overall index S equal to:

$$(6)\quad S = \frac{1}{M} \sum_{i=1}^{N-1} M_i\text{'} S_i\text{'}$$

If S_1' is very low, this indicates that the champion has won most of its games, especially against the teams with the lowest rank. If S_1' is very low compared to the overall CB-index S, this indicates that the champion was too good for that competition. A competition without the truly superior champion would have been more balanced.

Figure 4 about here

Figure 4 shows the development of the S_1'-index over the last 60 years. The linear trend line runs slightly downwards. This suggest that the champion becomes slightly more and more superior compared to the competition as a whole. A comparable downward tendency, a little sharper, is present in the development over time of the S_2' index.

---

[10] For that reason S_1 in Eqs. (3) and (4) cannot be calculated.

In the last four columns of Table 2 the *S'*-indices are given for the top 4 teams, again pitched against the CB-index overall. The top two seems to be a class above the rest, but the numbers 3 and 4 fit in nicely with the overall CB (63 and 67% versus an overall index of 68%).[11] Strictly, the procedure to determine whether the top teams belong to another, even stronger competition, should be stepwise: After deciding that the number 1 is superior because S_1' is far below S, a new overall index S should be calculated, based on the remaining teams, which in turn should be compared with the new S_1'. Is the latter far below the new S, then also the new number 1 should be eliminated from the competition, and so on until the two converge. Note that by hypothetically removing the number 1 from the competition the entire rank order of the remaining teams may change. In the calculations underlying Table 2, it is assumed that the rank order does not change because of the elimination of a team.

Summarizing, if one would have the choice to raise the balance of the competition, either by promoting the best two teams to a European superleague or by relegating the worst two teams to the Second Division,[12] then the former would probably have a slightly higher effect on the CB than the latter. The top teams are more superior to the rest than the worst teams are inferior to the rest (compare the sum of S_{20}, S_{19} and S_{18} with the sum of S_1', S_2' and S_3'). Admittedly, to separate the best teams from the competition would have a major drawback on the attractiveness of the First Division in France. It is not altogether imaginary that a league without say Paris Saint Germain, AS Monaco and Olympique Marseille would be seen as a league playing for fun, while elsewhere, in a closed European superleague, the real champions are playing.[13]

4.- Conclusion

In this paper we presented a new index to measure the CB. A minor disadvantage of the index compared to other indices of the CB is that it requires much more information. Making use of the results of the games in the French *Championnat*, our index shows only a slight downward

---

[11] The *S*'-indices for the top 4 teams in the Dutch competition are 26, 38, 46 and 57% respectively, against an overall index *S* of 54%.

[12] Of course not in the usual sense where the relegated teams are immediately replaced by teams promoted from the Second Division. What is meant here is a once and for all reduction of the number of teams in the league by a once and for all extra relegation of worst teams to a lower level.

[13] Hoehn and Szymanski [1999, p. 230-1] argue for such American-style European superleagues. Their highly hypothetical league structure is composed of four regional European leagues, among which a South West League composed of five teams from France (Paris Saint Germain, AS Monaco, Olympique Marseiile, Girondins Bordeaux and FC Nantes), five from Spain, three from Portugal and two from Belgium.

tendency over the period 1945-2002. Over the entire period, more than half of a century, the overall CB as measured by the surprise index has decreased with only 3% points. Perhaps the tendency and pattern of our index over time is comparable to the greenhouse effect of higher temperatures and the melting of ice caps because of carbon dioxide: the process goes slowly, but unmistakably in the wrong direction. Further, it seems that the CB can be improved upon, either by allowing the top teams to participate in a kind of European superleague or by relegating more teams to a lower level. Finally, we want to point out that the usefulness of our index is not limited to professional football leagues. It can also be used by the organizing football associations to investigate whether competitions at all kinds of levels, varying from the juniors to the highest levels of non-professional leagues, are reasonably balanced. Admittedly, a Sisyphus work, but it might show which competitions can be improved upon by adjusting the number of teams within a league. For the thousands of active players this may preclude that too many matches end up in boring 10-0 walkovers and too few in exciting 2-2 games.

References

Fort, R. and J. Maxcy [2003], Competitive Balance in Sports Leagues: An Introduction, *Journal of Sports Economics*, Vol. 4, no. 2, p. 154-160.

Hoehn, T. and S. Szymanski [1999], "The Americanization of European football", *Economic Policy*, 28, p. 205-33.

Koning, R.H. [2000], "Balance in competition in Dutch soccer", *The Statistician* 49 (3), p. 419-31.

Utt, J. and R. Fort [2002], Pitfalls to Measuring Competitive Balance With Gini Coefficients, *Journal of Sports Economics*, Vol. 3, no. 4, p. 367-73.

Appendix A

Aside from the rank order difference, there is another *natural* candidate to use as a weight, the ratio of the number of match points between two teams. All teams within a league compete for match points in the same pool, and the one who obtains the most is the champion, the one who obtains the least is relegated. If one team obtains twice as much match points as another team, then this seems to be a good measure for the relative strength of both teams. However, to apply this rule leads to contradictory results. Consider situations IV and V in Table A.

The only difference is that team A loses its home game against C in situation IV but wins it in situation V. Note that situation IV is identical to that of Table 1, except that the last column shows the number of match points, 2 for a win and 1 for a draw. If the surprising outcomes are weighted by the ratio in the number of match points between two teams, then the draw of C against A in situation V is evaluated as 1*7/1, whereas the win plus the draw of C against A in situation IV is evaluated as (2+1)*5/3 = 5. Since the denominator of the index is the same in both situations it should be the case that situation V describes a more balanced competition than situation IV, which is definitely not true. The problem with the rule to use the ratio of match points between two teams as weights is that there is no guarantee that the numerator of the index will increase monotonically as the balance in the competition increases. The fewer match points the worst teams obtain, the more they are worth if weighted by the ratios.

Table A. Cross table (surprising outcomes italicized; MP = match points)

IV	A	B	C	MP
A	x	2-0	*0-1*	5
B	0-2	x	2-0	4
C	*1-1*	0-2	X	3

V	A	B	C	MP
A	x	2-0	2-0	7
B	0-2	x	2-0	4
C	*1-1*	0-2	x	1

Appendix B

The simplifying assumption that rank order of the teams within a competition does not change when a team from the top or the bottom of the list is removed is not entirely innocuous. To see this consider situation VI in Table B.

Table B. Cross table (surprising outcomes italicized; WP = match points)

VI	A	B	C	P
A	x	*1-1*	2-0	**1**
B	*2-0*	x	*1-1*	**3**
C	0-2	*2-0*	x	**2**

The overall index S is equal to 75%, whereas $S_3 = 50\%$ (see Eq. (3)), which might be a reason to eliminate C from the league. If we focus only on teams A and B, and disregard all results against team C, then the number of surprise points is 3, and given that the maximum surprise points in a balanced competition of only two teams is 2, it follows that the new overall index S is equal to 150%. As is clear, this is entirely due to the (non-justifiable) assumption that the rank order does not change (situation VI illustrates that the rank order A-B does change by eliminating C). Taking into account that the rank order between A and B reverses gives only 1 surprise point, and therefore an index of 50%. In general, the assumption of non-changing rank orders leads to an overestimation of the overall index of the hypothetical slimmed down competition.

The issue of the rank order brings us to the question whether it is appropriate to use the final league table at the end of the (current) seasons to determine the rank number of the teams. Why not the rank number of the (incomplete) league table at the moment the games are played, or the rank numbers of the final league table of the preceding season? The choice made in this paper for the first alternative means that it can only be determined *ex post* whether a match was surprising or not. For the main theme of this paper, to measure the development of the CB over a long period of time, we think this choice is perfectly right. Matches played in a particular season are evaluated according to the final rank numbers of the teams in the same season. Doing otherwise, e.g. using the rank numbers of the previous season, would make the S-statistic for a particular year an amalgam of data of two seasons (with many inter-season player transfers changing the relative strengths of teams), not to mention the problem which rank numbers to assign to newly promoted teams.[14]

[14] It would make sense to use the league table of the preceding season if one would want to report the development of an *ex ante* CB-index during the season, e.g. after each round the newspaper can report that the surprise index for the games of the last weekend was x% and for the season so far y%.

Appendix C. Fact sheet of indices 1945-2003 (in %).

Year	N	S	S 3-3	S20	S19	S18	S18	S17	S16	S15	S1’	S2’	S3’	S4’
1945/46	18	71	37				46	60	65	56	44	51	79	71
1946/47	20	68	25	58	51	72					51	61	54	71
1947/48	18	62	32				28	50	73	79	49	39	35	65
1948/49	18	69	38				65	61	58	69	37	59	63	70
1949/50	18	62	26				41	56	65	70	38	60	60	62
1950/51	18	75	44				52	63	80	91	73	77	75	76
1951/52	18	66	26				39	46	66	60	54	57	47	79
1952/53	18	70	37				63	67	78	64	36	61	72	81
1953/54	18	66	30				49	65	69	60	57	53	45	62
1954/55	18	82	40				83	86	84	83	71	72	80	89
1955/56	18	69	22				53	49	56	67	67	67	71	72
1956/57	18	68	25				54	70	68	58	52	60	51	68
1957/58	18	71	46				62	65	64	73	54	63	58	73
1958/59	20	64	34	51	64	55					52	46	49	50
1959/60	20	60	32	45	58	41					34	52	55	68
1960/61	20	62	15	33	57	60					48	46	48	73
1961/62	20	73	34	61	57	75					78	75	68	57
1962/63	20	70	33	57	68	68					64	71	79	73
1963/64	18	78	34				63	75	63	107	74	75	81	68
1964/65	18	77	41				54	81	68	82	64	76	67	80
1965/66	20	69	21	84	51	61					33	57	56	86
1966/67	20	69	24	38	54	80					52	96	51	75
1967/68	20	72	32	54	54	85					53	81	67	81
1968/69	18	65	24				49	75	64	78	42	32	63	65
1969/70	18	65	37				53	72	51	56	32	59	59	63
1970/71	20	72	29	73	56	77					38	58	77	72
1971/72	20	63	37	51	55	47					38	50	62	75
1972/73	20	67	29	58	49	68					38	60	75	56
1973/74	20	73	20	65	64	78					46	65	79	64
1974/75	20	75	40	63	83	60					60	64	78	77
1975/76	20	74	27	46	54	86					54	69	75	84
1976/77	20	62	25	53	54	64					38	53	63	58
1977/78	20	65	25	42	83	64					47	59	55	65
1978/79	20	62	34	38	49	65					54	47	46	57
1979/80	20	59	37	22	40	63					34	44	32	47
1980/81	20	64	26	54	54	73					45	44	51	60
1981/82	20	63	25	47	47	64					41	36	73	57
1982/83	20	70	29	69	58	72					43	59	57	72
1983/84	20	64	25	38	62	47					48	42	64	59
1984/85	20	64	23	52	72	80					42	41	68	53
1985/86	20	68	32	37	75	70					53	43	65	57
1986/87	20	67	36	31	64	76					48	59	68	51
1987/88	20	76	28	58	80	96					50	79	80	78
1988/89	20	62	39	38	42	62					42	41	60	51
1989/90	20	73	25	63	70	74					55	51	75	75
1990/91	20	75	18	69	66	78					53	54	61	97
1991/92	20	67	26	66	63	64					41	46	58	74
1992/93	20	63	32	52	54	57					51	56	43	65
1993/94	20	61	19	51	46	59					34	56	69	63
1994/95	20	64	31	37	46	69					43	51	55	56
1995/96	20	67	29	47	60	71					45	61	66	78
1996/97	20	62	45	32	49	66					37	62	57	60
1997/98	18	64	37				59	59	45	55	61	49	62	62
1998/99	18	63	18				62	71	75	53	53	41	67	51
1999/00	18	77	39				56	74	91	89	59	65	78	67
2000/01	18	69	33				43	74	66	77	48	58	49	61
2001/02	18	69	24				52	72	65	72	66	54	66	52
2002/03	20	62	42	37	44	56					58	60	53	49
1945-2003		**68**	**31**	**50**	**58**	**68**	**54**	**66**	**67**	**71**	**50**	**57**	**62**	**67**

Figure 1. The overall index S in the period 1945-2002.

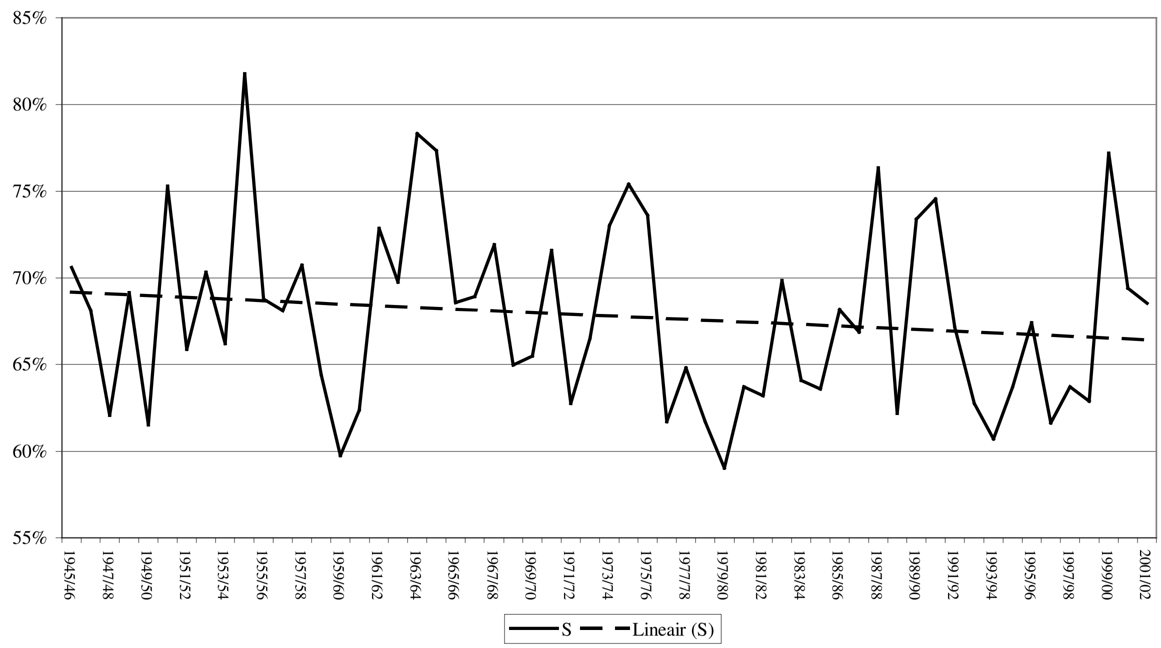

Figure 2. The surprise-index (S) pitched against the standard deviation (STD*) and the concentration ratio (C3*), 1947-1958.

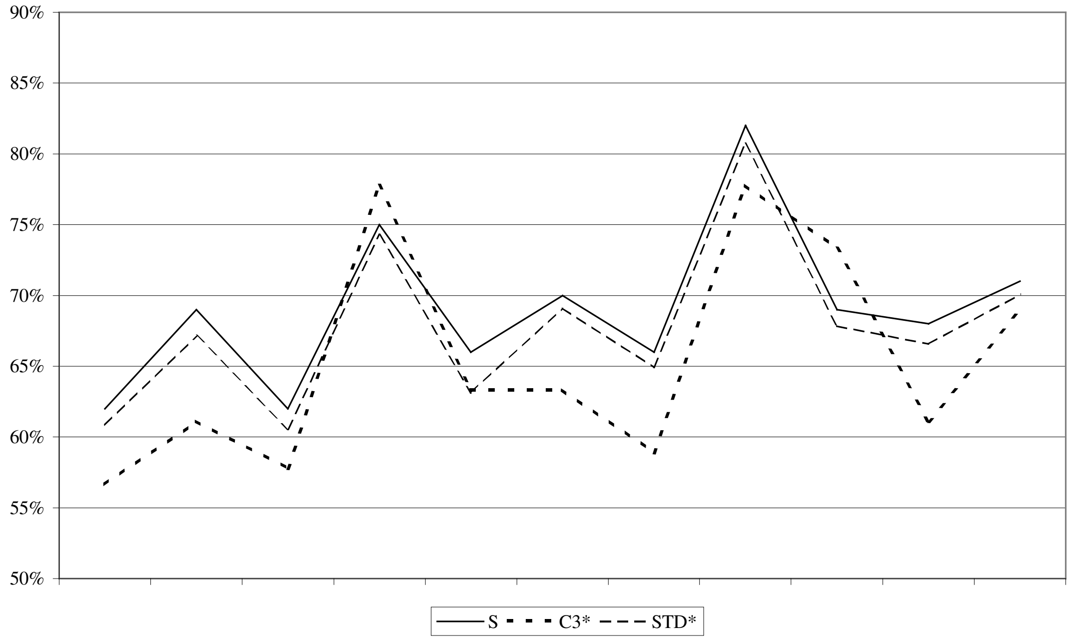

Figure 3. Top 3 against lowest 3

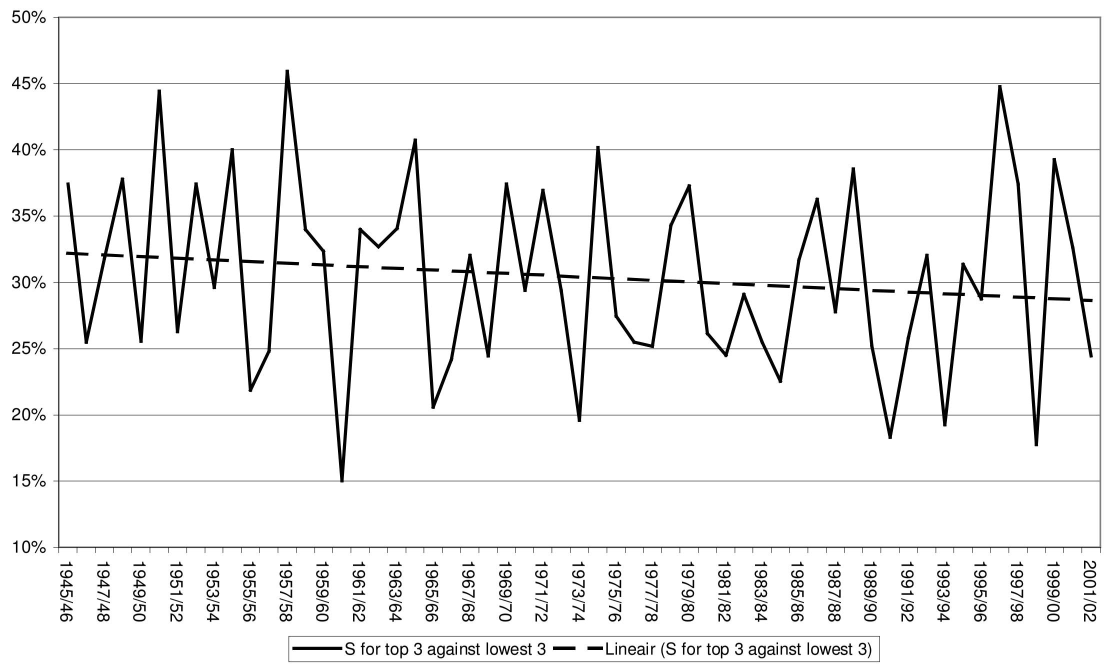

Figure 4. The index for the best team