Economics Bulletin

Volume 31, Issue 3

Measurement of competitive balance in professional team sports using the Normalized Concentration Ratio

Vasileios Manasis
University of Peloponnese

Vassiliki Avgerinou
University of Peloponnese

Ioannis Ntzoufras
Athens University of Economics and Business

J. James Reade
University of Birmingham

Abstract

Competitive balance is an important concept in professional team sports; its measurement is, therefore, a critical issue. One of the most widely used indices, which was introduced for the estimation of seasonal competitive balance is the Concentration Ratio, which is a relatively simple index and measures the extent to which a league is dominated by a particular number of teams. However, it is shown that both the total number of league teams and the number of dominant teams under examination affects the index's boundaries, which results in a misleading interpretation concerning the level of competitive balance. Thus, we introduce the Normalized Concentration Ratio for the study of competitive balance across leagues or seasons.

Citation: Vasileios Manasis and Vassiliki Avgerinou and Ioannis Ntzoufras and J. James Reade, (2011) ''Measurement of competitive balance in professional team sports using the Normalized Concentration Ratio'', *Economics Bulletin*, Vol. 31 no.3 pp. 2529-2540.
Submitted: Aug 31 2011. **Published:** September 02, 2011.

1. Introduction

Competitive balance is a key issue in professional team sports. Its importance derives from the fact that it creates an uncertainty of outcome, which instigates sport fans' interest and thus leads to an increased demand for attending and viewing sport events (El-Hodiri & Quirk, 1971; Rottenberg, 1956). Fans actually purchase the excitement generated by the uncertainty or unpredictability of the event outcome (Dobson & Goddard, 2001). Therefore, competitive balance is the "key characteristic" of the professional team sports product.

Various indices for the measurement of competitive balance have been proposed in the literature. As competitive balance is essentially concerned with inequalities amongst teams, the borrowing of indices from industrial organization theory is not surprising. One of the most popular is the *Concentration Ratio* (*CR*), which as an index is rather simple and easy to understand. Most importantly, it is a very helpful index for the study of seasonal competitive balance since it measures the extent to which a league or a championship is dominated by a small number of teams.

In the present paper we examine the *CR* index and the implications derived from its application to team sport setting. In Section 2 we discuss the fundamentals of the *CR* index and its importance for the study of seasonal competitive balance. Following that, in Section 3 we review the various ways in which it has been applied to team sports. In particular, we examine the effects of the variation on both the number of teams which make up the league and on the number of dominant teams under examination. It is shown that both the upper and the lower bound of the index are affected, which needs to be taken into consideration to avoid misleading results. In Section 4 we propose the derived expression of the *Normalized Concentration Ratio* (NCR_K), which rules out upper- and lower-bound violations. Finally, in Section 5 we conclude with a summary of the key points addressed in this paper.

2. The Concentration Ratio

Simplicity and limited data requirements make the *CR* index one of the most frequently used indices in industrial organization for the measurement of a market's share (usually expressed in turnover terms), which is accounted for by the market's *K* largest firms. The selection of the number of firms to be included in the *CR* index is a rather arbitrary decision; however, a preference for a small number is evident, since it enables a clear delineation of a market into dominant and fringe firms (Djolov, 2006). The mathematical expression is defined by the summation of the market shares of the *K* largest firms in the market and it takes the form:

$$CR = \sum_{i=1}^{K} S_i \qquad \textbf{(1)}$$

where S_i refers to the market share (expressed as a proportion) of the *i*th firm. The *CR* index ranges from zero to unity. The index approaches zero for an infinite number of equally sized firms (given that the number of *K* firms under examination is relatively small as compared to the total number of firms in the industry). The larger the *CR* index, the more monopolistic the industry is. The *CR* index reaches its upper value when the *K* largest firms completely cover the market.

In the context of team sports, a team's "market share" is interpreted as the number of points won by the team as a proportion of the total points won by all teams in the course of the season (Depken II, 1999). Essentially, the *CR* index, measures the degree of domination by the top *K* teams. One important criticism of the *CR* index in the context of sports leagues is that it examines the behavior of a slice of the league, that is, the top *K* teams. More specifically, it depends on only one point in the concentration curve[1]. Consequently, as it is depicted in Figure 1, for many fluctuations in the concentration curve the index could remain unchanged. Despite this significant weakness, the *CR* index is so widely employed for three important reasons:

a) It is easily understood.
b) It is highly correlated with more sophisticated measures (Groot, 2008; Kamerschen & Lam, 1975).
c) It clearly captures the degree of domination of the top *K* teams, which is the major cause for the decline of competitive balance in European soccer (Michie & Oughton, 2004).

However, the application of the *CR* index in team sports is not straightforward. In contrast to the standard industry, there are two main issues in which the fundamentals of the *CR* index are markedly different when applied to team sports.

Firstly, the total number of teams which make up the league, denoted by *N,* is rather limited, whereas the relevant number of firms in the standard industry could be infinitely large. This feature has implications for the value of the lower bound of the index which concerns cases of perfect balance, that is, the top *K* teams win on average the same number of points as the rest of the teams. Consequently, when the index is applied to team sports, its lower bound, which equals *K/N*, substantially deviates from zero which is the theoretical lower bound in the standard industry.

Secondly, it is impossible for the top *K* teams to gather all the points in a championship, since the remaining teams also have to play each other and will therefore gain at least some points. This is a well-known characteristic of the distribution of points in sports leagues and has repercussions on the upper bound of the index (Owen, 2009; Utt & Fort, 2002). The upper bound concerns cases of complete domination by the top *K* teams, that is, a league in which the best *K* teams always win any team with lower ranking. Therefore, the upper bound, which is defined as the ratio of the maximum number of points that the top *K* teams can gain over the total number of points in the league, is lower than unity (which is the case in a monopolistic standard industry). Consequently, for the application of the *CR* index in team sports, an appropriate adaptation is required since the index's boundaries differ substantially from the conventional ones.

3. Existing applications of the *CR* index in team sports

The first application of the conventional *CR* index was made by Koning (2000). He introduced his own version of the concentration ratio in relation to soccer, for which he employed the notation CR_K. Koning defined CR_K as the ratio of the total number of points

[1] The concentration curve (Figure 1) is created if we plot the cumulative point share against the ranking of the teams. The height of the curve above any point on the horizontal axis measures the percentage of the league's total points accounted for by the largest *K* teams. The curve is rising from left to the right and reaches its maximum height of 100 % at a point which corresponds to the total number of teams in the league (Bikker & Haaf, 2002).

obtained by the top K teams to the maximum number of points these K could possibly obtain; it is expressed as follows:[2]

$$CR_K = \frac{\sum_{i=1}^{K} P_i}{2K(2N-K-1)} \qquad (2)$$

where P_i is the number of points achieved by the ith team. The CR_K index accounts for the upper bound, since the expression in the denominator is the maximum number of points the top K teams could possibly collect. The upper bound of the CR_K index is unity, and is obtained for a completely dominated league by the top K teams. The more CR_K deviates from unity, the more balanced, or less dominated, the league becomes. The upper bound is well defined since it is constant and therefore insensitive both to N and K. However no provision has been taken for its lower bound. The lower bound is obtained for a perfectly balanced league as defined in Section 2. The number of points the top K teams win in a perfectly balanced league equals $2K(N-1)$. As a result, based on the equation (2) above, the mathematical expression of the lower bound of the CR_K (CR_{K_LB}) is as follows:

$$CR_{K\_LB} = \frac{2K(N-1)}{2K(2N-K-1)} = \frac{N-1}{2N-K-1} = \frac{N-1}{2(N-1)-(K-1)} = \frac{1}{2-\frac{K-1}{N-1}} \qquad (3)$$

From equation (3), it is obvious that the CR_{K_LB} is an increasing function of the number of K teams considered in the index. For K=1, the CR_{K_LB} is constant and equal to 0.5 which is its minimum value while for any K>1 it increases. Similarly, we can infer that the CR_{K_LB} is a decreasing function of the size of the league N. The variation of the CR_{K_LB} for selected N and K is presented in Table I and is graphically illustrated in Figure 2. The possible values of the CR_{K_LB} is rather large ranging from 0.5 to 0.64. Therefore a normalized version of the CR_K index which will account for both the lower and upper bounds is required for the analysis of competitive balance across leagues or seasons with different number of competing teams N and/or different number of top teams K examined.

Michie & Oughton (2004) followed a different approach for the application of the *CR* index also in soccer. They introduced the *C5 Index of Competitive Balance* (C_5ICB), which basically examines the degree of inequality between the top five teams and the rest of the teams. The C_5ICB index is defined as the ratio of the actual cumulative share of points of the top five teams to the cumulative share of points of the top five teams in a perfectly balanced league. The C_5ICB index is mathematically defined as:

$$C_5ICB = \frac{\sum_{i=1}^{5} S_i}{5/N} \qquad (4)$$

[2] We consider a round robin tournament championship in which every team plays twice against all others; in case of a win two points are awarded, whereas in case of a draw one point is awarded while no points are awarded in case of a loss.

where S_i stands for the share of points of the *i*th team. Essentially, the C_5ICB index is the *CR* index controlled for the case of a perfectly balanced league[3]. Consequently, the value of the lower bound of the C_5ICB index is unity and is reached when the top five teams win on average the same number of points as the rest of the teams. Any increase in the C_5ICB index implies a reduction in competitive balance and an increase in the dominance of the top five teams. The lower bound is well defined, since it is constant and hence insensitive both *N* and *K*. However, the upper bound of the index, which is the case of perfect domination by the top five teams, it is not specified in the index.

We can generalize the C_5ICB index and, following the same procedure as for the CR_K index, we can investigate the estimation of the C_KICB index's upper bound. As it is noted in equation (2), the total number of points the top *K* clubs could possibly obtain in a completely dominated league equals 2*K*(2*N*-*K*-1), whereas the total number of points allocated to all teams in the league can be estimated as 2*N*(*N*-1). Consequently, following equation (4), the upper bound of the C_KICB index (C_KICB_{UB}) is calculated as follows:

$$C_K ICB_{UB} = \frac{\frac{2K(2N-K-1)}{2N(N-1)}}{K/N} = \frac{2N-K-1}{N-1} = \frac{1}{CR_{K\_LB}} \quad \textbf{(5)}$$

Interestingly enough, the C_KICB_{UB} equals the inverse of the CR_{K_LB}. It can easily be derived from equation (5) that, for *K*=1, the C_KICB_{UB} is constant and equals 2 which is the maximum value over different *K*; for any *K* greater than one, the C_KICB_{UB} decreases. In particular, the magnitude of the decrease is affected both by *N* and *K*. This effect can be verified by the inverse inferences deducted from the differentiation of CR_{K_LB} with respect to *N* and *K* respectively. Therefore, for *K*>1 the C_KICB_{UB} is an increasing function of the size of the league *N*. Consequently, the larger the *N*, the closer the C_KICB_{UB} gets to its maximum value. Moreover, C_KICB_{UB} is negatively related to *K*. This implies that the larger the number of *K* teams under examination, the smaller the upper bound becomes.

The variation of the C_KICB_{UB} for selected *N* and *K* is presented in Table II and is graphically presented in Figure 3. The range of the possible values of the C_KICB_{UB} is quite large taking values from 1.55 to 2. Consequently, a sufficient normalization of the C_KICB index must account for its upper bound for the reliable and comparable measurement of the competitive balance for leagues or seasons with different sizes *N* and/or number of top *K* teams examined.

4. The Normalized Concentration Ratio

In this section we focus on the appropriate normalization of the *CR* index. We base our normalization on *CR* index's boundaries, which are fundamental for the proper and robust definition of our index. As it is noted in Section 2, the lower bound of the *CR* index equals *K/N* and corresponds to perfect competitive balance. However, the lower bound is not constant since it depends on *N* and *K*. More specifically, the lower bound is an increasing function of *K* while it is a decreasing function of *N*. This is presented in Table III and is graphically illustrated in Figure 4. As far as the upper bound is concerned, its value is

---

[3] The expression in the denominator K/N with *K*=5, stands for the value of the *CR* index in case of a perfectly balanced league.

obtained in case of a completely dominated league. If we consider that the total number of points allocated to all teams equals 2*N*(*N*-1) while the maximum number of points the top *K* teams could possibly collect is 2*K*(2*N*-*K*-1), the upper bound of the *CR* index (CR_{UB}) is calculated as:

$$CR_{UB} = \frac{K(2N-K-1)}{N(N-1)} \quad \textbf{(6)}$$

Based on equation (6), we can state that the CR_{UB} is not constant since it depends on *N* and *K*. More specifically, variation in the upper bound can be ascertained by differentiating (6) with respect to *N* and *K* as follows:

$$\frac{\partial \frac{K(2N-K-1)}{N(N-1)}}{\partial N} = \frac{2KN(K-N+1)-K(K+1)}{N^2(N-1)^2} < 0 \text{ for } N>1,\ K\geq 1,\ N\geq K+1 \quad \textbf{(7)}$$

$$\frac{\partial \frac{K(2N-K-1)}{N(N-1)}}{\partial K} = \frac{2N-2K-1}{N(N-1)} > 0 \text{ for } N>1,\ K\geq 1 \text{ and } N>K+1/2 \quad \textbf{(8)}$$

Equations (7) and (8) show us that the CR_{UB} is a decreasing function of *N* and an increasing function of *K*. This effect is depicted in Table III and is graphically illustrated in Figure 4 for selected *N* and *K*. In Table III it is also presented the range of the *CR* index which significantly varies depending on the various possible values of *N* and *K*. The range's variation emanates from the variation in the upper and lower bounds respectively.

This sensitivity of the range of the *CR* index that lies on different values of *N* and *K* provides convincing arguments for the development of a normalized version of *CR* index. Such a normalization should satisfy two conditions:

a) For a reliable calculation of the index, a point of reference is required. For this reason, the lower bound is chosen as a benchmark for the measurement.[4] Consequently, the subtraction of the lower bound from the observed value provides a re-located to zero measurement.
b) The value of the index has to be rescaled for the variability in both bounds. Interestingly enough, this can be achieved by dividing the re-located to zero measurement with the index's feasible range.

Consequently, following equation (1), the ratio of the above two conditions formulates the *Normalized Concentration Ratio* (NCR_K), which is mathematically defined as:

$$NCR_K = \frac{\sum_{i=1}^{K} S_i - \frac{K}{N}}{\frac{K(2N-K-1)}{N(N-1)} - \frac{K}{N}} = \frac{\frac{\sum_{i=1}^{K} P_i}{2N(N-1)} - \frac{K}{N}}{\frac{K(N-K)}{N(N-1)}} = \frac{\sum_{i=1}^{K} P_i - 2K(N-1)}{2K(N-K)} \quad \textbf{(9)}$$

[4] The upper bound could also be chosen. In that case, the observed value is subtracted from the upper bound.

The NCR_K index ranges from zero to unity regardless both of the number of the league teams N and the top K teams under investigation[5]. It approaches zero in case of a perfect balanced league, while it approaches unity in case of a totally dominated league by the top K teams.

NCR_K index provides a zero–one rescaled measurement of competitive balance. This is a major advantage since it enables us to make reliable comparisons across leagues of different size or across measurements with different number of top teams examined. This is of crucial importance if we are interested to study competitive balance across different leagues or different seasons where the size of the league is not constant. Additionally, a different number of the top K teams under examination may be required in order to study competitive balance according to the league's specific interest such as the number of teams qualifying in European competitions or experts' opinion or policy makers' aspiration. For instance, in England it is appropriate to examine the degree of domination of the top four teams since there four teams participating in Champions League whereas the relevant number in Germany is three and in Greece is two.

5. Conclusion

The *Concentration Ratio* (*CR*) is one of most commonly used indices because it neatly captures the degree of dominance of big teams, which is one of the major problems in soccer. However, the application of the index to team sports is not straightforward. The reason why this occurs relates to the fact that, in contrast to the standard industry, the number of N teams in a league is rather limited and the top K teams cannot gather all the points allocated in the course of a season. As a result, both the upper and the lower bounds of the index are a function of the particular values of the N and K. This can lead to misleading results when dealing with different numbers N of teams; i.e. when studying competitive balance across leagues or over time. Because of the deficiencies of the existing applications, it is introduced the *Normalized Concentration Ratio* (NCR_K) of which the range remains invariant irrespective of the number of teams N that make up the league. Moreover, the NCR_K index allows for a selection in the number K of the top teams under investigation according to a league's specific interests.

References

Bikker, J. A., & Haaf, K. (2002). Measures of competition and concentration in the banking industry: a review of the literature. *Economic & Financial Modelling, 9*, 53-98.

Depken II, C. A. (1999). Free-Agency and the Competitiveness of Major League Baseball. *Review of Industrial Organization, 14*, 205-217.

Djolov, G. (2006). *The Economics of Competition: The Race to Monopoly*. Binghampton, NY: The Haworth Press Inc.

Dobson, S., & Goddard, J. (2001). *The Economics of Football*. Cambridge, UK: Cambridge University Press.

El-Hodiri, M., & Quirk, J. (1971). An Economic Model of a Professional Sports League. *The Journal of Political Economy, 79*(6), 1302-1319.

Groot, L. (2008). *Economics, Uncertainty and European Football: Trends in Competitive Balance*. Northampton: Edward Elgar Publishing Limited.

[5] Under a proper modification, the NCR_K applies to any point system. For instance, for the application of the NCR_K index in the modern system (three points for a win, one for a draw and zero for a loss) an appropriate weight must be put in the case of perfect competitive balance. More specifically, this weight must be based on the ratio of the total number of wins to the total number of draws in the league.

Kamerschen, D. R., & Lam, N. (1975). A survey of measures of Market Power. *Rivista Internazionalle di Scienze Economiche e Commercialli, 22*, 1131-1156.

Koning, R. H. (2000). Balance in competition in Dutch Soccer. *The Statistician, 49*(3), 419-431.

Michie, J., & Oughton, C. (2004). *Competitive Balance in Football: Trends and Effects*. London: The Sports Nexus.

Owen, D. (2009). Pitfalls in Using the Relative Standard Deviation of Win Percentages to Measure Competitive Balance in Sports Leagues. *Economic Discussion Papers*, No 0908, University of Otago.

Rottenberg, S. (1956). The Baseball Players' Labor Market. *The Journal of Political Economy, 64*(3), 242-258.

Utt, J., & Fort, R. (2002). Pitfalls to Measuring Competitive Balance with Gini Coefficients. *Journal of Sports Economics, 3*(4), 367-373.

Figure 1: Concentration Curve

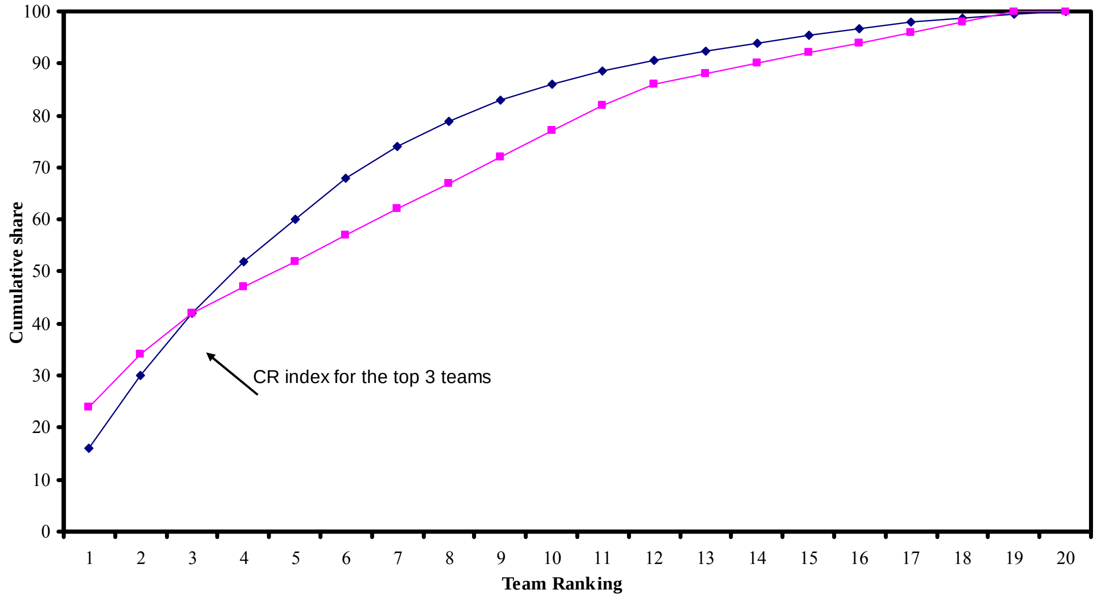

Table I: Lower Bound of the CR_K Index (CR_{K_LB})

K:	*N*:10	*N*:12	*N*:14	*N*:16	*N*:18	*N*:20	*N*:22
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	0.529	0.524	0.520	0.517	0.515	0.514	0.512
3	0.563	0.550	0.542	0.536	0.531	0.528	0.525
4	0.600	0.579	0.565	0.556	0.548	0.543	0.538
5	0.643	0.611	0.591	0.577	0.567	0.559	0.553

N: number of teams that make up the league
K: number of top teams under investigation

Figure 2: Lower Bound of the CR_K Index (CR_{K_LB})

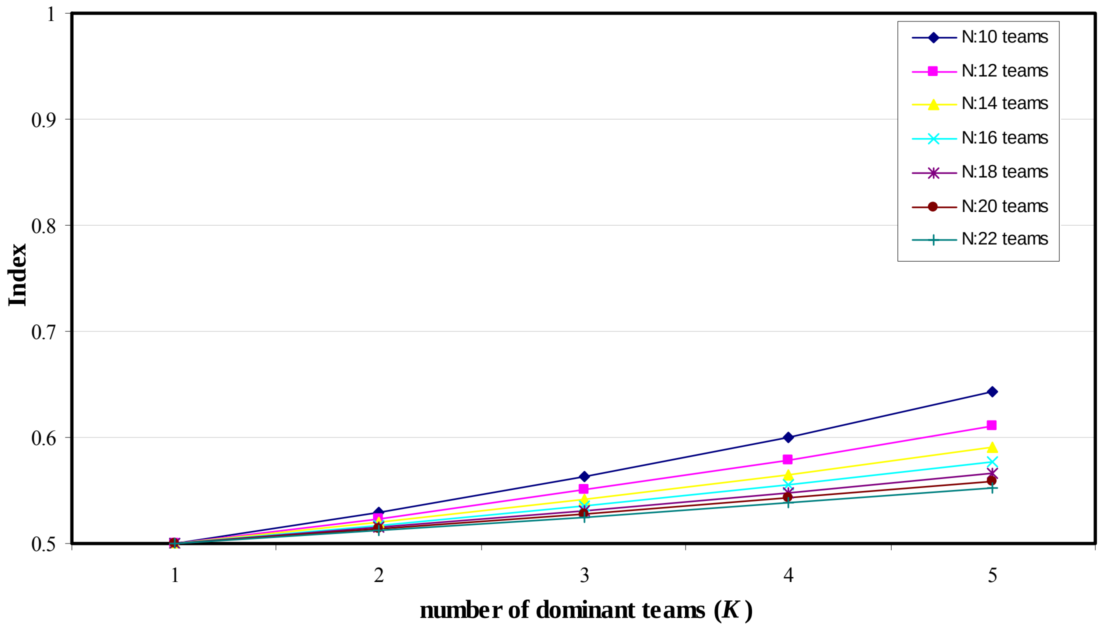

Table II: Upper Bound of the $C_K ICB$ Index ($C_K ICB_{UB}$)

K:	*N*:10	*N*:12	*N*:14	*N*:16	*N*:18	*N*:20	*N*:22
1	2.000	2.000	2.000	2.000	2.000	2.000	2.000
2	1.889	1.909	1.923	1.933	1.941	1.947	1.952
3	1.778	1.818	1.846	1.867	1.882	1.895	1.905
4	1.667	1.727	1.769	1.800	1.824	1.842	1.857
5	1.556	1.636	1.692	1.733	1.765	1.789	1.810

N: number of teams that make up the league
K: number of top teams under investigation

Figure 3: Upper Bound of the $C_K ICB$ Index ($C_K ICB_{UB}$)

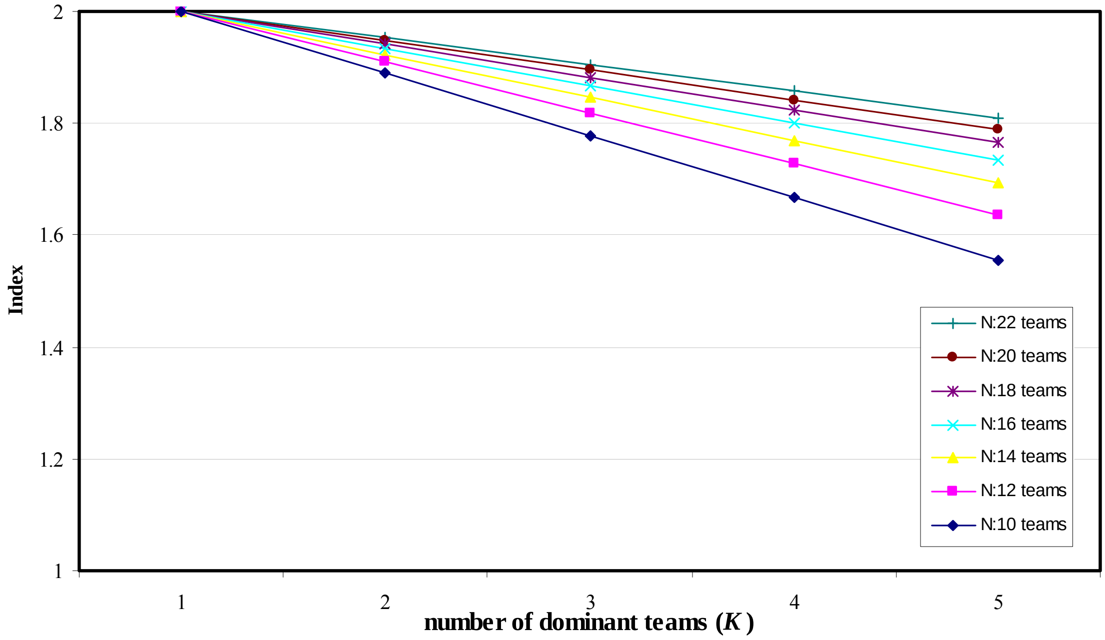

Table III: Lower Bound – Upper Bound – Range of the *CR* Index

	***N*:18**		***N*:20**		***N*:22**	
***K*:**	***Lower***	***Upper***	***Lower***	***Upper***	***Lower***	***Upper***
1	0.056	0.111	0.050	0.100	0.045	0.091
2	0.111	0.216	0.100	0.195	0.091	0.177
3	0.167	0.314	0.150	0.284	0.136	0.260
4	0.222	0.405	0.200	0.368	0.182	0.338
5	0.278	0.490	0.250	0.447	0.227	0.411
	Range		***Range***		***Range***	
1	0.056		0.050		0.045	
2	0.105		0.095		0.087	
3	0.147		0.134		0.123	
4	0.183		0.168		0.156	
5	0.212		0.197		0.184	

N: number of teams that make up the league
K: number of top teams under investigation

Figure 4: Upper & Lower Bound of the *CR* Index