Hidden power law patterns in the top European football leagues

Sergio, Da Silva and Raul, Matsushita and Eliza, Silveira

Federal University of Santa Catarina

2013

Hidden power law patterns in the top European football leagues

Sergio Da Silva[1], Raul Matsushita[2], Eliza Silveira[1]

[1] *Department of Economics, Federal University of Santa Catarina, Florianopolis SC, 88049-970, Brazil*
[2] *Department of Statistics, University of Brasilia, Brasilia DF, 70910-900, Brazil*

Abstract. Because sports are stylized combat, sports may follow power laws similar to those found for wars, individual clashes, and acts of terrorism. We show this fact for football (soccer) by adjusting power laws that show a close relationship between rank and points won by the clubs participating in the latest seasons of the top fifteen European football leagues. In addition, we use Shannon entropy for gauging league competitive balance. As a result, we are able to rank the leagues according to competitiveness.

PACS: 89.65.Gh; 05.10.Gg

Keywords: Power laws; Wars; Sports; Terrorism; Football; Soccer

1. Introduction

Sports are, arguably, stylized combat because they are contests of aiming, chasing, or fighting, complete with victors and the vanquished [1]. But they transcend that. Because sports are competition without killing, they fulfill a positive role of assuaging bitterness, seeking reconciliation, attempting conciliation, and pursuing courtesy. In this sense, sports are a useful substitute for war. They can divert nations from military aggression because of the heightened national self-esteem due to sporting success [2]. Whatever the deepest meaning of sports, however, their mechanics undeniably mimics that of war. What is more, while wars do not conserve the number of participants, sports do. This ultimately makes the analogy between wars and sports non trivial in that sports provide an ideal laboratory for studying competitions [3]. Unlike wars, the records of sports events are accurate, complete, and widely available [4].

Here, we are particularly interested in the mechanics of football (soccer), the number one sport. Apart from the United States and a handful of countries, soccer is the most popular game on Earth. We take data from the latest seasons of the top fifteen European football leagues, show similarities with the formal results already found for war in literature, and present novel results.

Earlier studies on war [5] found that wars with low death tolls far outnumber high fatality conflicts. Though this seems to be highly expected, what is not obvious is the finding that the number of wars with a given number of fatalities follows an approximate power law statistical distribution as a function of the number of fatalities. The link between the severity and frequency of conflicts follows a power law. This implies that extreme events such as the world wars cannot be considered as anomalies; they are expected to occur occasionally, given the frequency with which conflicts take place. Such power law also applies to individual clashes and acts of terrorism [6–8], and even to scientists career progress [9]. The possible explanation for the phenomenon is the fact that, for a wide range of human activities, the time taken to complete a given challenging task decreases with successive repetitions, following an approximate power law progress curve [8]. We start our analysis by suggesting that the same fact extends to football, given that it is stylized combat.

The rest of this article is organized as follows. Section 2 describes the data used, Section 3 carries out the analysis, Section 4 discusses the results, and Section 4 concludes the study.

2. Description of data

We collect data for the top fifteen European football leagues for the clubs participating in the latest seasons. The leagues we consider are the English, Italian, Spanish, German, Dutch, Portuguese, French, Scottish, Greek, Turkish, Belgian, Austrian, Danish, Polish, and Irish. The data source was soccerstats.com. Table 1 shows every league considered, the season for which data are available, total number of clubs participating in one league, one league's rounds per season, the number of datapoints per season, and total number of datapoints.

At the end of each season *j* of league *k*, we recover from the data the total points won by each team *i* and denote these by p_{ikj}. Then, we rank them in decreasing order, $p_{ik(1)} \geq p_{ik(2)} \geq \cdots \geq p_{ik(n)}$, where *n* is the number of clubs. For further convenience we denote $p_r \equiv p_{ik(r)}$, where *r* is rank. Actually, we first take the fraction of points won. Then, using the percentages of wins, draws and upsets we get the total points won through linear transformation. The win percentage is arguably a better measure of team strength than the points won are, because there are small variations between the numbers of matches played by various clubs across the leagues [4].

Figure 1 shows the outcomes of the top fifteen European football leagues for the 2011/12 season. There is a clear pattern of green squares (meaning matches won) for the clubs that ended up on top. And there is still a hidden power law pattern, as we will show next.

Table 1. Description of data sets for the top fifteen European football leagues

League	*Season*	*Total clubs*	*Rounds per season*	*Datapoints per season*	*Total datapoints*
Austrian	2009/10, 2010/11, 2011/12	10	36	180	540
Belgian	2011/12	16	20	160	160
Danish	2009/10, 2010/11, 2011/12	12	33	198	594
Dutch	2009/10, 2010/11, 2011/12	18	34	306	918
English	2008/09, 2009/10, 2010/11, 2011/12	20	38	380	1520
French	2008/09, 2009/10, 2010/11, 2011/12	20	38	380	1520
German	2008/09, 2009/10, 2010/11, 2011/12	18	34	306	1224
Greek	2009/10, 2010/11, 2011/12	16	30	240	720
Irish	2011/12	11	36	198	198
Italian	2008/09, 2009/10, 2010/11, 2011/12	20	38	380	1520
Polish	2010/11, 2011/12	16	30	240	480
Portuguese	2008/09, 2009/10, 2010/11, 2011/12	16	30	240	960
Scottish	2011/12	12	33	198	198
Spanish	2008/09, 2009/10, 2010/11, 2011/12	20	38	380	1520
Turkish	2009/10, 2010/11, 2011/12	18	34	306	918

Source: soccerstats.com

3. Analysis

In a league format of competition, every club plays every other club. In particular, each club hosts every other club exactly once during a season. Of course, the outcome of a match is subject to location, weather, injuries, red cards in the previous match, and a multitude of other factors. These make the outcome of a single game unpredictable to some degree; and, thus, this is so for the championship outcome. Thus, randomness is inherent to such a form of competition [4].

For this reason, we first take the normalized distributions of the points won p_r^* by the clubs participating in the leagues:

$$p_r^* = \frac{p_r}{\sum_{r=1}^{N} p_r}, \tag{1}$$

where N stands for the total number of clubs in one league, and $\sum_{r=1}^{N} p_r$ is the actual points won. Thus, we have

$$\sum_{r=1}^{N} p_r^* = 1, \tag{2}$$

so that equation (1) represents a probability distribution.

Second, we conjecture that there is a law describing a close relationship between rank r and the normalized points won p_r^* as a power law of the form

$$p_r^* = a \cdot r^b, \tag{3}$$

where p_r^* changes as if it were a power of r. The problem is then to verify the conjecture and determine a and b. Coefficient a is the normalizing constant, and considering (3) and (2), we have

$$a^{-1} = \sum_{r=1}^{N} r^b. \tag{4}$$

Rather than taking logs on both sides of (3), in order to satisfy (2) and (4) we directly carry out a fit of (3) using nonlinear regression.

Figure 2 displays the empirical values of p_r^* along with the fitted power laws for the leagues. Coefficients a and b were found by running nonlinear regressions on (3) for each league (Table 2). As p_r^* represents a probability distribution we can calculate its Shannon entropy H, defined as

$$H = -\sum_{r=1}^{N} p_r^* \log_2 p_r^*. \tag{5}$$

The usefulness of H is related to the fact that it gives the degree of uncertainty and then can be interpreted as referring to the degree of competitiveness of one league given in bits.

The insight for the use of entropy comes from classical mechanics, in particular from a would-be analogy between particle allocations, energy levels, and rankings. In statistical mechanics, the analogy refers to a system of K particles which are distributed in R energy levels. In such a situation, the entropy is the log of the number of possible configurations. Here, K means the total points and R is ranking. The total points are distributed between the R teams according to a distribution rule given by a power law. We find a similar analogy in literature with the rank-citation profile of scientists [9].

The hypothetical maximum entropy occurs when the distribution of normalized points won is uniform, in which case $H_{max} = \log_2 N$. In the model given by equation (3), H_{max} occurs if b = 0. By contrast, if $b \to -\infty$, then $H \to 0$. Thus, one league's competitiveness is greater as H approaches H_{max} or, equivalently, as b approaches zero. Table 2 sorts the leagues for every season according to the value of their Shannon entropy. The entropies were found empirically from the normalized data H_{emp} and also from the power law equation (3), H_I. From the gauge provided by H, we then found the top five as follows: French Ligue 1 for the 2010/11 season, English Premier League for the 2010/11 season, Spanish La Liga for the 2008/09 season, French Ligue 1 for the 2011/12 season, and Italian Serie A for the 2009/10 season. The full rankings of competitive balance can be seen on the first column in Table 2.

Next, we found an interesting parallel, linear relationship between H_{emp}, H_I, and $H_{max} = \log_2 N$ (Figure 3), following the configuration

$$H_{\bullet} = \log_2 N - \delta_{\bullet}, \tag{6}$$

where $\bullet$ stands for either the empirical data or the power law (3). The dashed line in Figure 3 represents $H_{emp} = \log_2 N - \delta_{emp}$, the dotted line is $H_I = \log_2 N - \delta_I$, and the solid line is the reference $H_{max} = \log_2 N$. Strikingly, such linearity suggests that the distributions of normalized points won for the distinct leagues across different seasons can be represented by a sole model. However, there is minor discrepancy between the lines H_{emp} and H_I, which amounts $|\delta_{emp} - \delta_I| = 0.07$ bits.

This fact prompted us to try out a second power law model of the form

$$p_r^* = a \cdot b^{(r-1)^c}, \tag{7}$$

where there is dampened exponential decay of the normalized points won p_r^*. Parameter a is still the normalizing constant, $b \in (0,1]$ is the decay rate, and c is the dampening rate. We assume $b^{0^c} \equiv 1$. As a result, the linearization takes place with double log transform. Now, b = 1 refers to maximum uncertainty. There is near uniformity if c = 0, where $p_1^* > p_2^* = ... = p_N^*$. If c = 1, the power law decays geometrically. And if $c \to +\infty$, then $p_1^* > p_2^* > p_3^* = ... = p_N^* = 0$.

Figure 4 shows estimates of the power law in equation (7) for every league across different seasons, whereas Table 3 displays the coefficients adjusted by nonlinear regression. Table 3 also shows the entropy measures H_{II} calculated on the basis of equation (7). Again, we found a parallel, linear relationship between H_{emp}, H_{II}, and $H_{max} = \log_2 N$. But there is further good news. The lines H_{emp} and H_{II} almost overlap one another, meaning that the discrepancy is reduced seven fold, that is, it drops to $|\delta_{emp} - \delta_I| = 0.01$ bits. We thus conclude that the power laws described by equation (7) (Figure 4) describe the data better than the ones in equation (3) (Figure 2).

Finally, the discrepancy between one empirical distribution q_r^* and that of one model $p_{r\bullet}^*$ can be formally assessed by reckoning the Kullback-Leibler divergence [10], defined as

$$KL = \sum_{r=1}^{N} q_r^* \log_2 \frac{q_r^*}{p_{r\bullet}^*}. \tag{8}$$

The results in Table 4 do confirm that the fittings using the power laws in the model described by equation (7) beat the ones from the model in equation (3).

4. Discussion

In the literature of war and terrorist attacks which have inspired this work, the existence of power laws means that conflicts can be understood without considering local factors such as geography and politics. The data in Johnson et al. [8], for example, refer to the timing of attacks and number of casualties from more than 54,000 events across nine insurgent wars, including those of Iraq between 2003 and 2008 and of Sierra Leone between 1994 and 2003. By plotting the distribution of the frequency and size of events, the power law for insurgent wars was that the frequency of attacks decreases with increasing attack size to the power of 2.5. For any insurgent war, an attack with 10 casualties is 316 times more likely to occur than one with 100 casualties, because 316 is 10 to the power of 2.6. This law is surprising in light of the fact that wars are all fought in different terrains and under different circumstances. The explanation provided was that insurgent groups form and fragment when they sense danger, and strike in well-timed burst to maximize their media exposure. The authors also claim that the relative persistence of these power law distributions over time allows one to estimate the severity of future wars or clashes within an ongoing war [6, 7]. Similar arithmetic can be amassed for the 30 power laws in Figures 2 and 4. We are also optimistic about the relative persistence of such power laws over time after having considered the football leagues for distinct seasons.

The power laws describing a close relationship between rank and points won for the leagues studied can be accommodated by the general insight that the time taken to complete a given challenging task decreases with successive repetitions, following a progress curve, as observed. People adapt to circumstances, learn how to do things better, and productivity improves as a result. The twist in football matches is that two antagonistic sides are doing the adapting. Like predators and prey in constant competition, there emerges stasis, as each adaptation by one in countered by an adaptation by the other. The co-evolution between the antagonistic sides eventually reaches equilibrium and a fairly regular power law takes place.

Take the English Premier League for the 2011/12 season, for instance. As the league rounds progressed, the clubs involved learned better and better how to outsmart rivals, and this at an increasing rate. Champions Manchester City gained momentum as time passed and the same could be said of the other participants. Like an arms race evolving over time the clubs on bottom struggling to escape the relegation zone also learned better and better how to prevent that to occur. But because perfect counter-adaptation is unfeasible, and also because there are favorites and underdogs, some clubs inevitably ended up within the relegation zone.

Our finding of a parallel, linear relationship between the entropies H_{emp}, H_I, and H_{max} suggests a unique model for the distributions of normalized points won for the distinct leagues across different seasons, as observed. This fact can perhaps justify a unifying model readily applicable to a variety of competition formats, such as the one put forward by Ben-Naim et al [4].

Their model is based on a single (but fixed) parameter, the upset probability that a weaker team upsets a stronger team. The lower bound of the upset probability (zero) corresponds to predictable games where the stronger team always wins. The upper bound ($\frac{1}{2}$) corresponds to random games. Interestingly, this exactly matches our entropy measure of competitive balance, where maximum entropy (say, their upset probability = $\frac{1}{2}$) corresponds to maximum competitiveness, where the underdog is more likely to upset the favorite. After all, the more random the outcome of an individual match is, the higher the degree of parity between the clubs in a league [4].

One interesting result is that their model predicts an optimal upset frequency ≈ 0.4. And this is not at odds with data from a variety of sport competitions. They also find that single-elimination tournaments can crown a champion faster than leagues do, although the champion in tournaments is less guaranteed to be the best team. They then search for an optimal competition format and find it as a hybrid schedule consisting of a preliminary round (tournament) and a championship round [4]. Intriguingly enough, we observe that this stands as the opposite design of both the current UEFA Champions League format and that of the Soccer World Cup, where (roughly) the championship round actually precedes the tournament round.

Another interesting development [9] that is worth mentioning is the use of a discrete generalized beta distribution to model rank profiles that exhibit the progress curve scaling behavior, such as the one shown in this study.

5. Conclusion

Like wars, individual clashes, acts of terrorism and scientists' productivity, sports may follow power law progress curves. This occurs because the time taken to complete a given challenging task decreases with successive repetitions. Progress curves are a consequence of people adapting to circumstances and learning how to do things better. Warfare and, as we claim, sports, are just as capable of productivity improvements as any other activity.

We then adjusted power laws suggesting a close relationship between rank and points won by the clubs participating in the latest seasons of the top fifteen European football leagues. One insight that proved to be useful was to calculate the Shannon entropies, which track the degree of uncertainty and thus can be interpreted as the competitive balance of one league in a given season. We then were able to rank the leagues according to competitive balance.

The full significance of the existence of hidden power law patterns in the football leagues can be appreciated as one takes one practical implication into account. As long as the power laws are stable over time, one club manager can get a rough estimate of his club's rank at the end of the season based only on targeted points to be won, and the information conveyed in previous seasons can possibly be useful.

References

[1] S. Pinker, The Blank Slate: The Modern Denial of Human Nature, Penguin Books, New York, 2002.

[2] J.A. Mangan, Making European Masculinities: Sport, Europe, Gender, Frank Cass Publishers, London, 2000.

[3] A.M. Petersen, W.S. Jung, H.E. Stanley, Europhysics Letters 83 (2008) 50010.

[4] E. Ben-Naim, N.W. Hengartner, S. Redner, F. Vazquez, Journal of Statistical Physics 151 (2013) 458.

[5] L.F. Richardson, Journal of the American Statistical Association 43 (1948) 523.

[6] J.C. Bohorquez, S. Gourley, A.R. Dixon, M. Spagat, N.F. Johnson, Nature 462 (2009) 911.

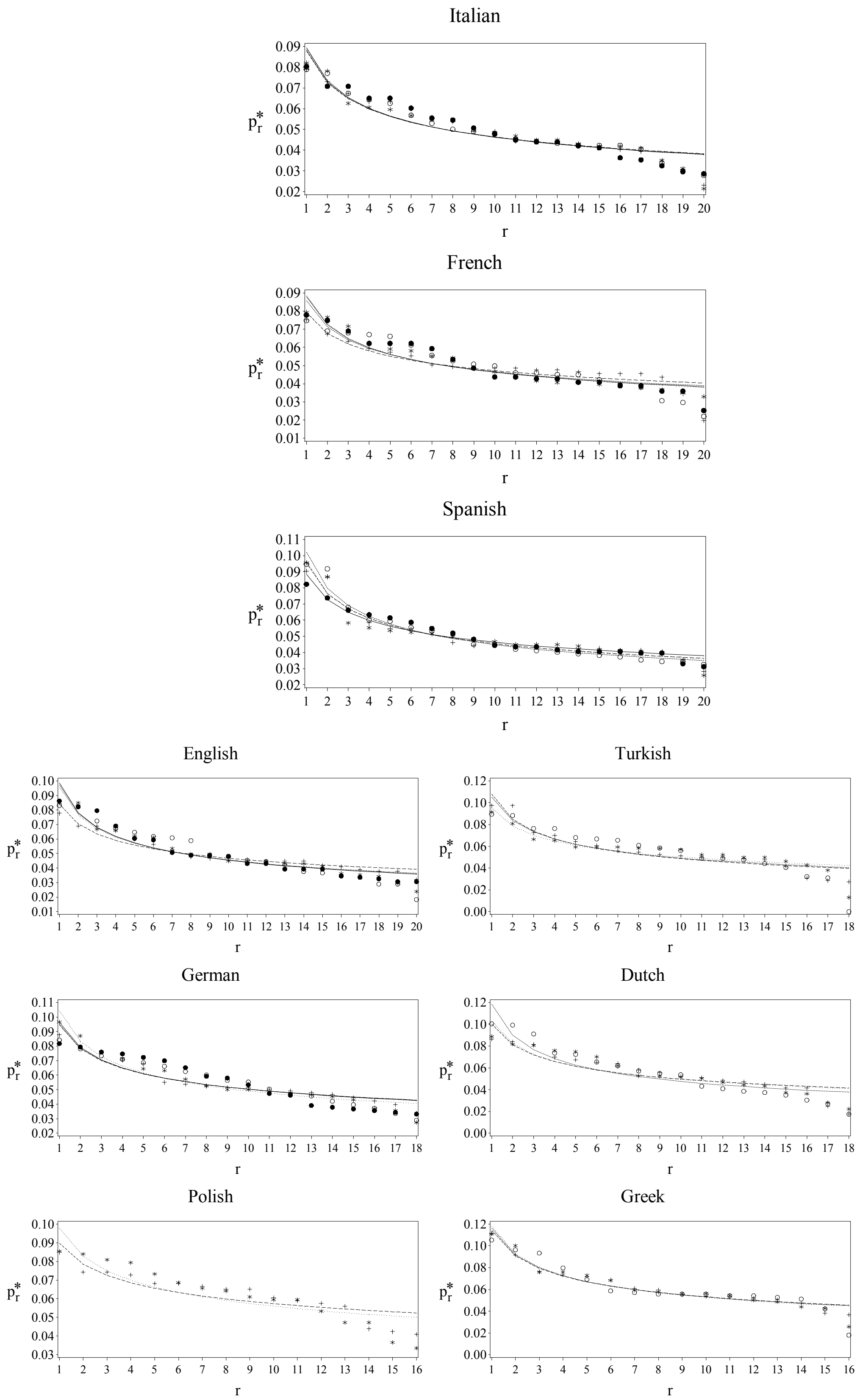
[7] A. Clauset, M. Young, K.S. Gleditsch, Journal of Conflict Resolution 51 (2007) 58.

[8] N. Johnson, S. Carran, J. Botner, K. Fontaine, N. Laxague, P. Nuetzel, J. Turnley, B. Tivnan, Science 333 (2011) 81.

[9] A.M. Petersen, H.E. Stanley, S. Succi, Scientific Reports 1 (2011) 1.

[10] R. Matsushita, A. Figueiredo, S. Da Silva, Physica A 391 (2012) 4891.

Figure 1. Top fifteen European football leagues outcomes for the 2011/12 season. From top to bottom (left to right) the tables cover the clubs participating in the English, Italian, Spanish, German, Dutch, Portuguese, French, Scottish, Greek, Turkish, Belgian, Austrian, Danish, Polish, and Irish leagues. Green squares mean matches won, yellow squares represent draws, and red squares stand for matches lost. The expected pattern is the presence of more green squares for the clubs that ended up on top. But there is still a hidden power law pattern, as this study demonstrates. The data are from soccerstats.com.


Italian
French
Spanish
English
Turkish
German
Dutch
Polish
Greek
p*r
r

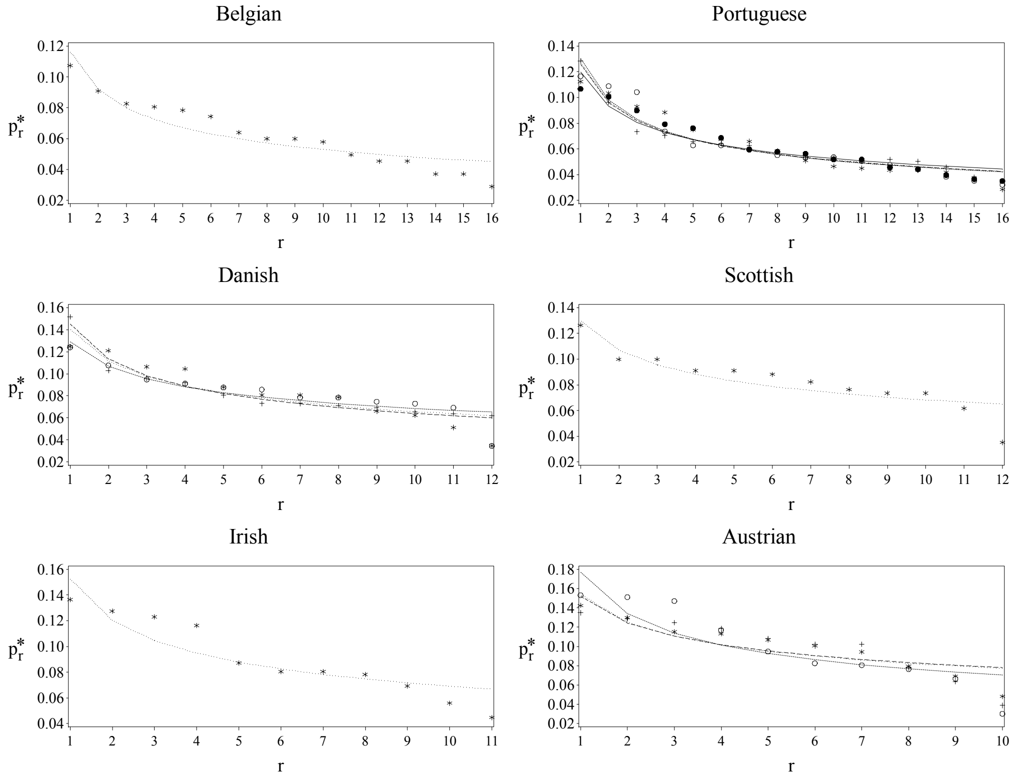


Figure 2. Empirical values of p_r^* (2008/09 season = dots, 2009/10 = circles, 2010/11 = crosses, 2011/12 = stars) and the fitted power laws $p_r^* = a \cdot r^b$ for the top European football leagues (2008/09 = solid lines, 2009/10 = long dashed lines, 2010/11 = short dashed lines, 2011/12 = dotted lines). The patterns do not change significantly across the seasons. The coefficients a and b (Table 2) were estimated by nonlinear regressions on (3).

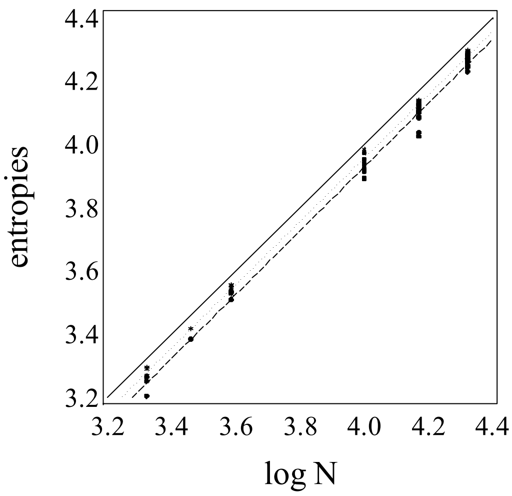


Figure 3. Interesting parallel, linear relationship between the empirical entropy H_{emp} versus $\log_2 N$ (dashed line), the power law (equation (3))-based entropy H_I versus $\log_2 N$ (dotted line), and the reference (solid) line $H_{max} = \log_2 N$. The three lines are parallel in the form $H_q = \log_2 N - \delta_q$. This finding suggests the existence of a common power law pattern underlying every league for different seasons. However, the power law fittings from model (7) is superior to this one provided by equation (3) (see Figure 5).

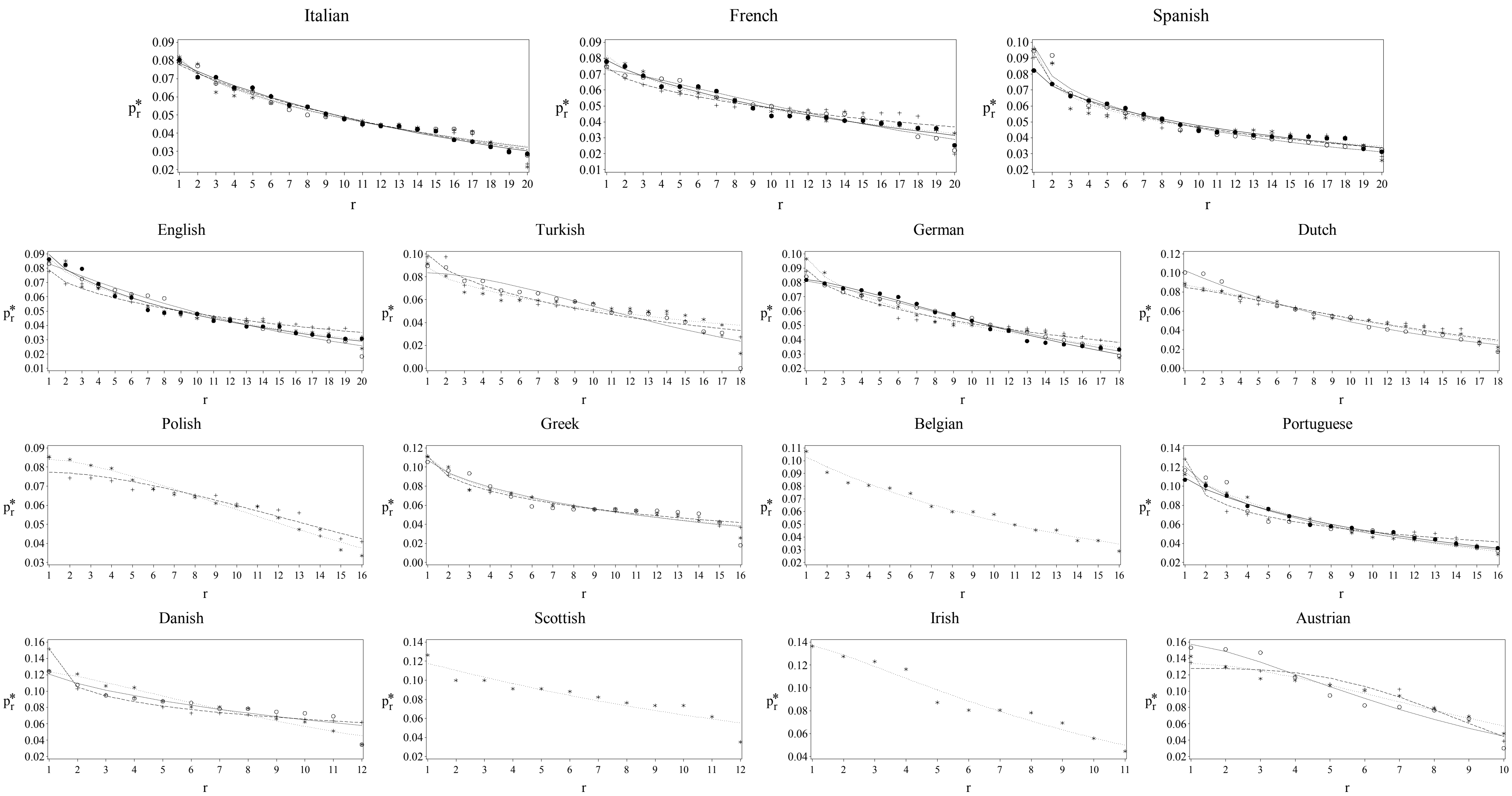


Figure 4. Empirical values of p_r^* (2008/09 season = dots, 2009/10 = circles, 2010/11 = crosses, 2011/12 = stars) and the fitted power laws $p_r^* = a \cdot b^{(r-1)^c}$ for the top European football leagues (2008/09 = solid lines, 2009/10 = long dashed lines, 2010/11 = short dashed lines, 2011/12 = dotted lines). The three coefficients *a*, *b*, and *c* (Table 3) were estimated by nonlinear regressions on (7).

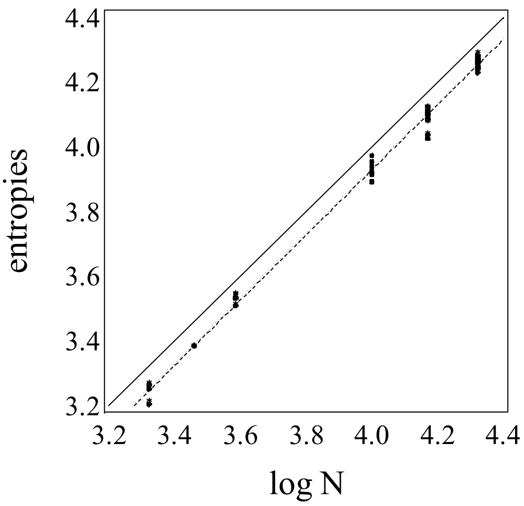


Figure 5. Interesting parallel, linear relationship between the empirical entropy H_{emp} versus $\log_2 N$ (dashed line), the power law (equation (7))-based entropy H_{II} versus $\log_2 N$ (dotted line), and the reference (solid) line $H_{max} = \log_2 N$. The three lines are parallel in the form $H_q = \log_2 N - \delta_q$. The line of the empirical entropy and that of the power law almost overlap one another and thus the fittings using equation (7) beat those using equation (3).

Table 2. Empirical Shannon entropy H_{emp}, power law (equation (3))-based entropy H_I, and the coefficients *a* and *b* of the power laws in (3).

League	*Season*	H_{emp}	H_I	*a*	*b*
French	2010/11	4.28244	4.29844	0.07907	–0.22404
English	2010/11	4.28142	4.29160	0.08402	–0.25495
Spanish	2008/09	4.27253	4.28467	0.08837	–0.28098
French	2011/12	4.27058	4.28542	0.08864	–0.28237
Italian	2009/10	4.26780	4.28659	0.08832	–0.28037
French	2008/09	4.26634	4.28836	0.08785	–0.27743
Italian	2010/11	4.26447	4.28868	0.08778	–0.27697
Italian	2011/12	4.26441	4.28471	0.08888	–0.28378
French	2009/10	4.26232	4.29387	0.08562	–0.26382
Italian	2008/09	4.26169	4.28680	0.08939	–0.28628
Spanish	2010/11	4.25977	4.26749	0.09559	–0.32276
Spanish	2011/12	4.25913	4.26326	0.09624	–0.32687
English	2011/12	4.24839	4.27518	0.09569	–0.32195
Spanish	2009/10	4.24427	4.25650	0.10220	–0.35792
English	2008/09	4.24298	4.27104	0.09852	–0.33711
English	2009/10	4.22942	4.27849	0.09728	–0.32953
German	2010/11	4.12435	4.13535	0.09457	–0.27465
German	2009/10	4.11172	4.13948	0.09476	–0.27489
German	2008/09	4.10320	4.13975	0.09612	–0.28215
Turkish	2011/12	4.10216	4.13457	0.09533	–0.27890
German	2011/12	4.10058	4.12050	0.10449	–0.32893
Dutch	2010/11	4.08640	4.13364	0.10012	–0.30433
Turkish	2010/11	4.08638	4.11557	0.10793	–0.34666
Dutch	2011/12	4.08362	4.13073	0.10283	–0.31866
Dutch	2009/10	4.03666	4.10871	0.11860	–0.39706
Turkish	2009/10	4.02547	4.12984	0.10502	–0.32975
Polish	2010/11	3.97281	3.98449	0.08974	–0.19434
Polish	2011/12	3.95355	3.97742	0.09786	–0.24233
Greek	2010/11	3.93845	3.94868	0.11412	–0.33163
Greek	2011/12	3.92275	3.94465	0.11793	–0.35043
Portuguese	2010/11	3.91984	3.92124	0.12587	–0.39115
Belgian	2011/12	3.91719	3.95172	0.11654	–0.34240
Portuguese	2008/09	3.91634	3.94635	0.11965	–0.35791
Greek	2009/10	3.91415	3.94969	0.11611	–0.34085
Portuguese	2009/10	3.89256	3.92650	0.13009	–0.40804
Portuguese	2011/12	3.89213	3.93673	0.12711	–0.39307
Danish	2009/10	3.53579	3.55551	0.12895	–0.27398
Scottish	2011/12	3.53318	3.55484	0.12988	–0.27859
Danish	2010/11	3.53140	3.52762	0.14513	–0.35515
Danish	2011/12	3.50963	3.54577	0.14044	–0.32947
Irish	2011/12	3.38433	3.41777	0.15269	–0.34361
Austrian	2011/12	3.26623	3.29105	0.15396	–0.29943
Austrian	2010/11	3.25125	3.29494	0.15205	–0.28963
Austrian	2009/10	3.20477	3.26960	0.17685	–0.40075

Table 3. Empirical Shannon entropy H_{emp}, power law (equation (7))-based entropy H_{II}, and the coefficients *a*, *b*, and *c* in (7).

League	*Season*	H_{emp}	H_{II}	*a*	*b*	*c*
French	2010/11	4.28244	4.29323	0.07415	0.90816	0.67314
English	2010/11	4.28142	4.28206	0.07862	0.89389	0.67112
Spanish	2008/09	4.27253	4.27317	0.08280	0.87736	0.65418
French	2011/12	4.27058	4.26982	0.08161	0.89716	0.72314
Italian	2009/10	4.26780	4.27033	0.08029	0.91129	0.77480
French	2008/09	4.26634	4.26855	0.07874	0.92879	0.85448
Italian	2010/11	4.26447	4.26910	0.07760	0.93959	0.91206
Italian	2011/12	4.26441	4.27218	0.08159	0.89683	0.72075
French	2009/10	4.26232	4.26753	0.07271	0.97624	1.23773
Italian	2008/09	4.26169	4.26348	0.07866	0.94050	0.93305
Spanish	2010/11	4.25977	4.26160	0.09302	0.80005	0.51463
Spanish	2011/12	4.25913	4.26387	0.09695	0.75243	0.43444
English	2011/12	4.24839	4.25190	0.08673	0.89295	0.76013
Spanish	2009/10	4.24427	4.24221	0.09771	0.80542	0.56618
English	2008/09	4.24298	4.24098	0.08947	0.88609	0.75827
English	2009/10	4.22942	4.23505	0.08288	0.95076	1.06816
German	2010/11	4.12435	4.12570	0.08896	0.88089	0.66951
German	2009/10	4.11172	4.11403	0.08103	0.97221	1.23270
German	2008/09	4.10320	4.10204	0.08211	0.97659	1.32551
Turkish	2011/12	4.10216	4.12643	0.08834	0.89360	0.71632
German	2011/12	4.10058	4.10348	0.09722	0.86484	0.69300
Dutch	2010/11	4.08640	4.10031	0.08473	0.96890	1.23700
Turkish	2010/11	4.08638	4.09469	0.09933	0.86768	0.72270
Dutch	2011/12	4.08362	4.08902	0.08747	0.96366	1.20468
Dutch	2009/10	4.03666	4.04346	0.10241	0.92368	1.01740
Turkish	2009/10	4.02547	4.08038	0.08341	0.98960	1.69101
Polish	2010/11	3.97281	3.97579	0.07733	0.99343	1.66655
Polish	2011/12	3.95355	3.95691	0.08424	0.98582	1.49040
Greek	2010/11	3.93845	3.94166	0.11102	0.81089	0.56917
Greek	2011/12	3.92275	3.93040	0.11124	0.84172	0.66888
Portuguese	2010/11	3.91984	3.92322	0.12863	0.70408	0.43209
Belgian	2011/12	3.91719	3.92365	0.10294	0.92366	0.96721
Portuguese	2008/09	3.91634	3.91574	0.10867	0.89083	0.84300
Greek	2009/10	3.91415	3.92978	0.10737	0.87408	0.75672
Portuguese	2009/10	3.89256	3.89274	0.12073	0.83810	0.73087
Portuguese	2011/12	3.89213	3.89108	0.11423	0.89016	0.88766
Danish	2009/10	3.53579	3.54896	0.12072	0.90684	0.84078
Scottish	2011/12	3.53318	3.54642	0.11778	0.93800	1.03030
Danish	2010/11	3.53140	3.53186	0.15163	0.69167	0.37321
Danish	2011/12	3.50963	3.51701	0.12369	0.95559	1.29148
Irish	2011/12	3.38433	3.38805	0.13726	0.93928	1.20882
Austrian	2011/12	3.26623	3.27245	0.13439	0.97746	1.64594
Austrian	2010/11	3.25125	3.25726	0.12785	0.99832	2.93359
Austrian	2009/10	3.20477	3.21579	0.15751	0.94428	1.40458

Table 4. Kullback-Leibler divergences (in bits) between the empirical distributions and the two models of power laws given by equation (3) (model *I*) and equation (7) (model *II*). The last column shows the difference between the two models. Negative differences favor model *II*.

League	*Season*	*Model I*	*Model II*	*II – I*
Turkish	2009/10	0.061639	0.035226	–0.026414
Austrian	2010/11	0.025101	0.001128	–0.023973
Austrian	2009/10	0.023823	0.009162	–0.014661
Dutch	2010/11	0.018049	0.006761	–0.011287
Dutch	2011/12	0.012669	0.002475	–0.010194
Danish	2011/12	0.012489	0.002873	–0.009616
Austrian	2011/12	0.011333	0.001889	–0.009445
Dutch	2009/10	0.011436	0.002640	–0.008795
German	2008/09	0.010053	0.001612	–0.008441
English	2009/10	0.010499	0.002077	–0.008422
Polish	2011/12	0.009518	0.001166	–0.008352
French	2009/10	0.010529	0.002398	–0.008131
German	2009/10	0.007556	0.000141	–0.007415
Irish	2011/12	0.009882	0.003226	–0.006656
Belgian	2011/12	0.007268	0.001124	–0.006144
Scottish	2011/12	0.012980	0.008179	–0.004801
Polish	2010/11	0.006170	0.001667	–0.004503
Italian	2008/09	0.004291	0.000477	–0.003815
Italian	2010/11	0.006120	0.002372	–0.003748
Turkish	2011/12	0.021627	0.018043	–0.003584
Danish	2009/10	0.012510	0.009788	–0.002723
Danish	2010/11	0.002990	0.000441	–0.002549
Italian	2011/12	0.006652	0.004687	–0.001965
French	2008/09	0.004447	0.002638	–0.001809
Portuguese	2010/11	0.003940	0.002575	–0.001365
French	2010/11	0.010275	0.009011	–0.001263
Italian	2009/10	0.003288	0.002158	–0.001130
English	2011/12	0.003754	0.002813	–0.000941
Turkish	2010/11	0.006846	0.006062	–0.000784
Portuguese	2011/12	0.003337	0.002617	–0.000720
Portuguese	2008/09	0.001713	0.001110	–0.000602
Greek	2011/12	0.005731	0.005246	–0.000485
Greek	2009/10	0.015500	0.015204	–0.000296
Greek	2010/11	0.001635	0.001501	–0.000134
German	2011/12	0.002311	0.002253	–0.000057
Spanish	2011/12	0.006019	0.006199	0.000180
English	2010/11	0.001253	0.001567	0.000314
Spanish	2008/09	0.000914	0.001268	0.000354
German	2010/11	0.001687	0.002147	0.000460
French	2011/12	0.000706	0.001313	0.000607
English	2008/09	0.000872	0.002484	0.001612
Spanish	2010/11	0.003140	0.005519	0.002379
Portuguese	2009/10	0.002587	0.006760	0.004173