Statistical Methods & Applications (2004) 13: 241–258
DOI: 10.1007/s10260-004-0093-3

SMA

Structural changes during a century of the world's most popular sport

Ignacio Palacios-Huerta*

Department of Economics, Brown University, Box B; 64 Waterman Street; Providence, RI 02912, USA
(e-mail: ipalacios@brown.edu.)

Abstract. The analysis in this paper employs a methodology for dating structural breaks in tests with non-standard asymptotic distributions. The application examines whether changes in the rules of a game and major social and political events during the past century had significant effects upon various outcomes of this game. The statistical methodology first applied here proves successful as most breaks can be traced to specific events and rule changes. Dating these breaks allows us to obtain useful insights into production and competition processes in this industry. As such, using empirical tests we illustrate the utility of a valuable statistical technique not applied previously.

Key words: Structural changes, changepoint estimators, time series analysis, football, sports industry

1. Introduction

In this paper, we apply a recently developed statistical methodology for dating structural breaks in time series. The application is concerned with what is considered to be a preeminent social phenomenon and an economically important industry.

Football (soccer) is the world's most popular sport. The game has unrivalled worldwide appeal. Hundreds of millions of children and teenagers throughout the globe practice it. Many more adults follow it. Countless football teams, especially in Europe and Latin America but also in Africa and Asia, are adored in their home cities.

* I am grateful to Gary S. Becker, Tony Lancaster, Robin Lumsdaine, Kevin M. Murphy, Gabriel Perez-Quiros, Ana I. Saracho, Amy Serrano, an associate editor and a referee for useful suggestions. I am also indebted to Barry Blake, Vicki Bogan, Salwa Hammami and Karen Wong for able research assistance, Tony Brown at the Association of Football Statisticians for the data the Hoover Institution for its hospitality, and the Spanish Ministerio de Ciencia y Tecnologia for financial support (grant BEC 2003-08182). The Gauss programs used in this paper were kindly provided by Bruce E. Hansen. Any errors are mine alone.

Even in North America, where other sports rule, many more teenagers kick footballs than hit baseballs and there are twice as many football teams on college campuses as American football teams.[1] Around the globe, the game attracts its fans at an early age and their passion does not typically diminish with age. Not surprisingly, star players are often more famous than religious and political leaders. Media proprietors consider fan interest almost insatiable and supply it by TV, satellite, radio and newsprint. Sports newspapers, whose pages are mostly occupied by football news, are typically the best-selling newspapers in most European and Latin American countries. World Cup football matches are the biggest TV shows in the world, drawing a global television audience surpassed by no other event. World Cup 2002, for instance, drew a cumulative audience of 42.5 billion people, or the world's population seven times over. Football, not surprisingly, has become an important medium to secure brand visibility and many corporations consider it key to gaining a global market share for their products. The game is indeed often considered one of the most important social phenomena of the 20th century.[2]

The global following of this sport, as a sociological and economic phenomenon, deserves a careful study of what the game produces that makes it so attractive. Social sciences in general, and economics in particular, have not ignored the study of various aspects of football as a game, an industry, and a product subject to one of the greatest demands on the planet. The demand for football has been studied, for instance, by Kuypers (1995), Baimbridge et al. (1996), Peel and Thomas (1988, 1992), and Jenett (1984).[3] The economic and revenue sharing structure, as well as the incentive and agency problems of sport leagues have been examined, for instance, by Atkison, Stanley and Tschirgart (1988), Fort and Quirk (1995), El-Hodiri and Quirk (1971), and Sloane (1971).[4] More recently, different policy, financial and legal aspects of the football industry and European football leagues have attracted substantial research attention (see Hoehn and Szymanski 1999, Szymanski and Smith 1997, Quirk and Fort 1992, Bourg and Gouget 1988).[5]

Despite these efforts, however, there is an important aspect of this industry that has not been studied in the economics and statistical literature, one which has important implications for all of the aforementioned aspects. In particular, little is known about the temporal properties of the outcomes of the game. This paper is an attempt to provide the first statistical analysis of the behavior over time of the outcomes of the game and of the structural changes that these outcomes may have experienced during the 20th century.

---

[1] Data reported by the National Collegiate Athletic Association at http://www.ncaa.org.

[2] The term football, instead of soccer, will be used throughout the paper since this is the way the game is referred to in most parts of the world. Modern football was essentially invented and "coded" at the end of the 19th century in the United Kingdom. References to predecessors of the game have been found in ancient China and Japan, as well as in ancient Greece and Rome, in the British Isles during the Middle Ages, in William Shakespeare's *King Lear* (Act I, Scene IV), and in Italy and France through the 16th, 17th and 18th centuries (see http://www.fifa.com, Walvin 1994, Vallet 1998). The first official league competition took place in 1888 in England. From there, professional and amateur competitions spread quickly and widely all over the world early in the past century.

[3] See also Dixon and Coles (1997), Ridder et al. (1994), Maher (1982), and Reep et al. (1971).

[4] See also Bulow et al. (1985) and Dobson and Goddard (1998).

[5] See also the analyses by Deloitte and Touche (1998a, 1998b).

We consider that the substantive question is important and that the empirical tests can illustrate the application of a valuable statistical technique. The results may be relevant to statistics, economics, and other social sciences for various reasons.

First, the analysis offers what to the best of our knowledge is the first application of a recently developed methodology to compute the *p*-values of a wide class of tests of structural change in econometric models. An important shortcoming in many of these tests is that the distributions are non-standard, and thus *p*-values cannot be computed from previously published information. This critical inconvenience is overcome in Hansen (1997) which presents a method to compute convenient parametric approximation functions to the *p*-value functions for various asymptotic distributions.

Second, the past decade has seen considerable empirical and theoretical research on the detection of breaks in economic time series. Most of the series studied are concerned with macroeconomic variables. This paper shows how this statistical methodology may be useful to economists and social scientists attempting to date structural breaks in other types of time series as well, in particular in time series of outcomes. This application is also relevant since the last few years have seen increasing interest in the practical applications of game theoretical analysis in economics and other social sciences. Empirical applications of game theory models fundamentally rely on data of the outcomes of the games under study.

Lastly, the application may be relevant in that social scientists are interested in the production function of industries, especially those that hold large economic and social importance. Often, however, many final "products" are not quantifiable in the sense that the quantity and quality of goods such as cars, clothes, and others can be measured. Final outcomes of games, in particular, reflect many aspects of the production function of sport industries. Also, outcomes have major effects on revenues (e.g., ticket sales, TV rights, merchandise), that is, on both the demand and supply sides. The structure of outcomes, in turn, has an effect upon the economic structure and the revenue sharing, incentive, and agency problems of competition. Also, it can be a determinant of different policy, legal, and structural aspects of the industry. Thus, the results of the application of this statistical technique may be valuable in the context of various underlying questions of interest.

This econometric study is concerned with the behavior over time of the main outcomes of the game of football. It studies the final scores of more than 120,000 games over more than 100 years across both professional and amateur leagues. The analysis focuses on the English league competitions, where an unusually rich data set is available from the *Association of Football Statisticians* for the period 1888–1996.

We will begin by describing the behavior over time of outcomes such as the average and the variability of goals scored, the margins of victory, the frequency of scores, and the percentages of home team wins, losses and ties. The behavior of these and other outcomes over time may result in part from "learning" to play the game and also from changes in inputs, production functions, and other aspects of the industry. Although an explicit analysis of strategic behavior and technological progress is beyond the scope of analysis of the paper, the evidence to be presented will be suggestive of both changes in strategic behavior and technological progress over time.[6]

We then will estimate, in each of the about two hundred time series of outcomes that we study, the dates at which a structural change is statistically likely to have occurred. These structural changes may have been caused, for instance, by rule changes, socioeconomic or political events (e.g., world wars), or other factors. Obviously, different changes may have had an effect in some time series but not in others (e.g., they may have affected the variability of goals scored but not their average, or outcomes in professional leagues but not in amateur leagues). Rather than testing for structural changes at specific candidate dates, we let the structure of each time series reveal the likelihood, at each date, that a structural change happened. In this sense we follow the typical *post hoc ergo propter hoc* approach followed in the literature. We find that once a specific date has been identified as one where, in a statistical sense, a structural change occurred, it is often possible to identify the particular event or rule change that likely caused the structural break. The fact that most of the dates of structural breaks can be associated with specific events shows that the methodology is highly accurate, reliable, and hence valuable as a statistical technique.

The rest of the paper is organized as follows. Section 2 describes the data and the basic features of the time series of some of the main outcomes of the game. Section 3 describes the methodology of the tests of structural change. Section 4 applies these tests to the time series we study and discusses the results. Section 5 concludes with a summary and some final remarks.

2. Data description

The data come from the Association of Football Statisticians, which has collected data on British league competitions since their creation. They consist of the final scores of *all* the games in the professional English leagues (Division I, or Premier League, and Division II) during the periods 1888–1996 and 1893–1996 respectively, and in the amateur English leagues (Divisions III and IV) during the period 1947–1996. There were 39,550 games in Division I, 39,796 in Division II, 20,424 in Division III and 20,017 in Division IV, during these periods.[7] With these data, time

[6] For instance, players may learn to adopt more efficient offensive and defensive strategies over time, their skills may improve, and inputs may also increase as players and coaches become more "professional." The study of professional and amateur divisions, therefore, may be valuable to determining whether differences in certain outcomes of the game are related to differences in the skills of players, since professional players have greater skills than amateurs. Divisions that have different promotion and relegation rules (e.g., because in top (bottom) divisions no team can be relegated (promoted) from a better (worse) division), may induce differences in the distribution of skills within divisions that could help to explain differences in outcomes across divisions such as the number of goals scored, the proportion of wins, margins of victory, and others.

[7] League competitions typically start in the Fall of one year and conclude in the Spring of the following year. Teams play each other twice during the season, once in each home field. Because of World Wars I and II, leagues did not compete in England during the seasons 1915/16–1918/19 and 1939/40–1945/46. Professionalization of most players in Divisions I and II did not occur until the 1960s and 1970s (see

Table 1. Percentages of Scores for 1888–1996

	1888–1915		1920–1939		1947–1982		1983–1996		Entire period	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Percentage of wins										
Premier League	0.797	0.040	0.762	0.029	0.746	0.036	0.727	0.027	0.761	0.042
Division II	0.813	0.040	0.766	0.031	0.740	0.036	0.730	0.024	0.763	0.046
Division III					0.731	0.029	0.737	0.013	0.733	0.024
Division IV					0.741	0.032	0.737	0.017	0.740	0.028
Percentage of home wins										
Premier League	0.566	0.042	0.547	0.029	0.503	0.027	0.474	0.042	0.526	0.047
Division II	0.596	0.033	0.550	0.028	0.515	0.028	0.486	0.020	0.539	0.047
Division III					0.516	0.030	0.485	0.023	0.505	0.031
Division IV					0.522	0.030	0.480	0.029	0.507	0.036
Percentage of away wins										
Premier League	0.231	0.029	0.215	0.017	0.244	0.028	0.253	0.023	0.236	0.028
Division II	0.217	0.025	0.215	0.018	0.255	0.021	0.244	0.017	0.244	0.022
Division III					0.215	0.024	0.252	0.018	0.228	0.028
Division IV					0.219	0.020	0.257	0.035	0.232	0.032
Percentage of ties										
Premier League	0.203	0.403	0.238	0.029	0.254	0.036	0.273	0.027	0.239	0.042
Division II	0.187	0.040	0.234	0.031	0.260	0.036	0.270	0.024	0.237	0.046
Division III					0.269	0.029	0.263	0.013	0.267	0.024
Division IV					0.259	0.032	0.263	0.017	0.260	0.028

Notes: League competitions were stopped during 1915/16–1918/19 and 1939/40–1945/46 because of World Wars I and II. Division II data start in 1893.

series are constructed for several different outcomes. This section offers a brief description of the data.

Table 1 collects the mean and standard deviation of the percentage of wins, home and away, and ties in the four divisions. On average, roughly 52 percent of all the games played during the century concluded in wins by the home team, 24 percent in wins by the visiting team, and 24 percent in ties. Undoubtedly, playing at home was very valuable. Yet, judging from the persistent decline over time in the proportion of home team wins, the advantage of playing at home seems to decrease steadily during the century. With respect to the variability of the different series, the most variable one is the percentage of home wins in the Premier League during the last decade of data. Interestingly enough, the greater the quality of the division, the greater its variability tends to be. Thus, a lower variability appears to be associated with lower skill levels in these series. One last aspect worth highlighting in the table is the very large extent to which the different time series are similar across the four divisions. This characteristic of the data is also present in most of the remaining series that will be studied, and will also be confirmed by the tests of structural change that will be implemented in Sect. 4.

Brown 1997). In what follows, seasons will often be referred to simply by the year in which the season concluded.

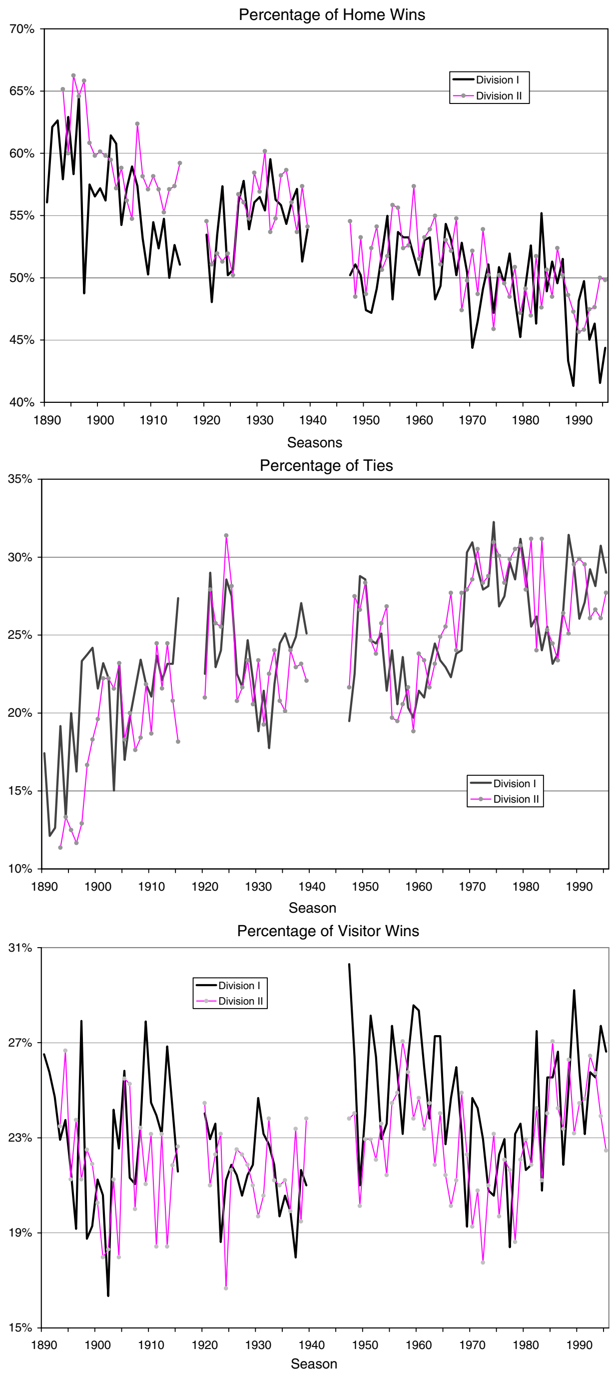

Fig. 1. Percentage of home wins, ties, and visitor wins (1889–1995)

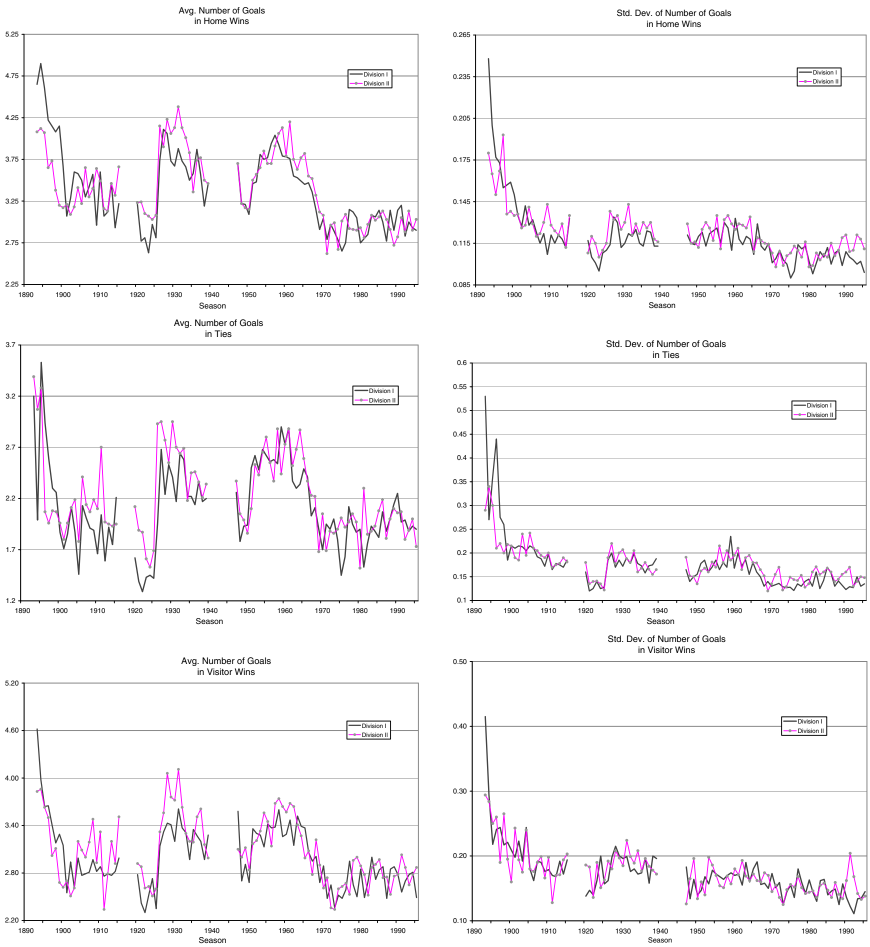

Fig. 2. Number of goals average and std. deviation 1893–1995

Figure 1 shows the percentage of home team wins, ties, and visitor wins for Divisions I and II. The percentage of games in which one team wins in these divisions decreases during the century from around 80 percent to around 72 percent. Conversely, the proportion of ties increases from around 20 percent to around 28 percent. The percentage of home team wins decreases from around 57 percent to around 47 percent, while that of away teams increases from 23 percent to 26 percent. With respect to Divisions III and IV (not shown), their trends and patterns are very similar to those of the top divisions. If anything, the percentage of home (visiting) team wins during the last half of the century is slightly lower (greater) in the top divisions.

Table 2 and Fig. 2 show how the average number of goals per game decreases over time. The mean number of goals in every division is around 2.6 over the last decade, which is substantially lower than the average of more than 4 goals at the beginning of the century. This may explain why during the last few decades the

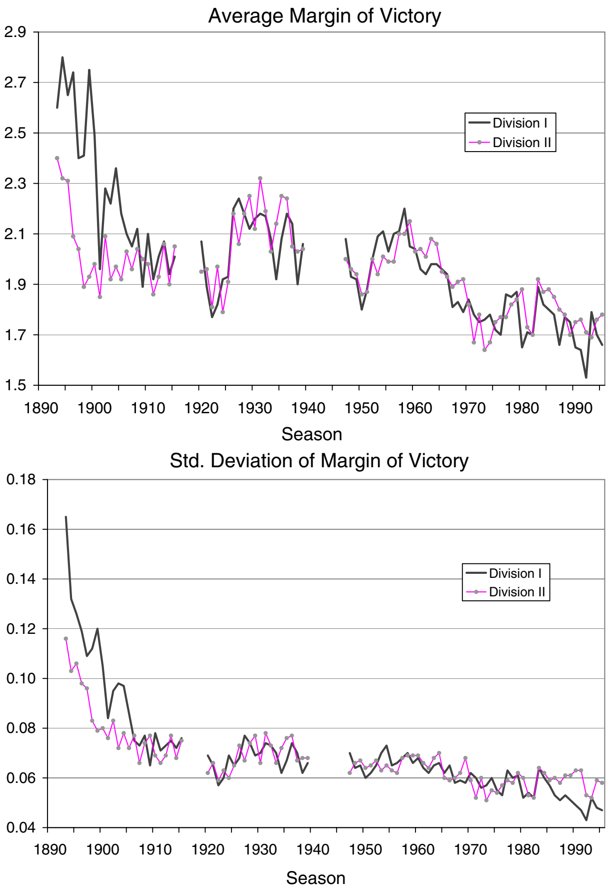

Fig. 3. Margin of victory average and standard deviation 1893–1995

governing bodies of football in England, and throughout the world as well, have attempted to promote more scoring by modifying certain rules of the game (for example, by awarding three points to wins instead of two and by establishing the backpass ruling).[8] Note that after World War I and World War II the number of goals experiences a substantial increase. This increase, in principle, does not reflect the fact that after interrupting the game for a few years players' skills diminished. Had this been the case, we should have observed that the average number of goals

[8] See The History of the Laws of the Game and Amendments to the Laws of the Game published by the *Fédération Internationale de Football Association* (FIFA). The decision to increase the points awarded to winners from two to three was established in 1981 by the English Football Association, and in 1996 in the rest of professional leagues in the world by FIFA. Of course, it is not clear that awarding three points for a win, instead of two, promotes scoring. Tournament theory predicts that such a change in the incentives to win will induce more offensive effort and more *defensive* effort as well, and hence scoring need not be promoted (see Garicano and Palacios-Huerta 2002). The backpass ruling established in 1992 worldwide penalizes players for deliberately kicking the ball to the hands of their own goalkeepers. It is estimated that this new rule has increased the amount of time the ball is in play by about 10 percent.

Table 2. Goals per match

Mean and standard deviation during different periods

Premier league								
	1888–1915		1920–1939		1947–1982		1983–1996	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Wins	3.54	0.55	3.58	0.46	3.26	0.41	2.90	0.10
Home wins	3.66	0.56	3.71	0.47	3.30	0.44	2.96	0.13
Away wins	3.27	0.57	3.27	0.48	3.07	0.41	2.80	0.14
Ties	2.32	0.57	2.32	0.44	2.23	0.36	1.95	0.14
All matches	3.30	0.57	3.28	0.47	3.00	0.42	2.65	0.10
Division II								
	1893–1915		1920–1939		1947–1982		1983–1996	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Wins	3.51	0.53	3.33	0.43	3.21	0.39	2.90	0.12
Home wins	3.65	0.56	3.44	0.45	3.30	0.41	2.98	0.13
Away wins	3.11	0.48	3.05	0.40	3.00	0.39	2.74	0.16
Ties	2.12	0.51	2.04	0.46	2.17	0.39	1.96	0.11
All matches	3.25	0.55	3.03	0.46	2.94	0.41	2.65	0.09
	Division III				**Division IV**			
	1947–1982		1983–1996		1947–1982		1983–1996	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Wins	3.06	0.31	2.94	0.14	3.07	0.29	2.94	0.08
Home wins	3.16	0.32	3.02	0.17	3.18	0.31	3.04	0.12
Away wins	2.84	0.33	2.80	0.16	2.81	0.28	2.75	0.15
Ties	2.03	0.27	2.01	0.11	2.04	0.30	1.99	0.18
All matches	2.78	0.32	2.70	0.12	2.80	0.30	2.69	0.09

Notes: League competitions were stopped during 1915/16–1918/19 and 1939/40–1945/46 because of World Wars I and II.

was greater in Divisions II, III and IV, with lower skill levels, than in Division I before these wars, which is not the case. A possible explanation may then involve the interaction between the basic nature of the game and learning to play the game *again*, at least to the extent that these aspects are independent of the skill level.[9] This idea seems plausible given that the behavior in most time series during the first few years *after* World War II very much resembles that from the decade *before* the war.

A second observation of interest is that the average number of goals per game is almost identical across the four divisions over the last few years. Also, games

[9] The nature of the game is such that, in a match, two teams compete for a single prize (winning the match) and only relative performances (in terms of goals scored) matter. Thus, this structure is identical to the typical tournament model studied in the literature on optimal labor contracts (see Lazear and Rosen (1981) and Lazear 1989).

in which the home team wins have a greater goal average than those in which the visiting team wins.[10] This is probably due to the more conservative, defensive strategies typically employed by visiting teams. The time trends of these averages by type of game (home team win, tie, visiting team win) experience a decrease over time that is similar to that of the average number of goals for all games in general. The time variability of the different time series greatly decreases over time as well, especially over the last two decades. Overall, the differences among the four divisions are slim, especially over the last decade of data where the series are almost indistinguishable from each other. Around 1925/26 and after World War II, we see the most important increase in the number of goals scored during the century. After the mid-1970s, the series stabilize at about 2.6 goals per game. With respect to the cross-sectional variability of the number of goals (Fig. 2), it also decreases slowly over time.[11] The time variability over the different periods in Table 2 is around 0.4–0.5 for the first three quarters of the century in the top divisions, declining quite sharply in the last two decades to about 0.10.

Figure 3 shows the time series of the average and the cross-section standard deviation of the margin of victory (e.g., the margin is 1 in games whose final score is 1–0, 2–1, 0–1, 3–4, etc, and 0 in ties). These series, which are also very similar across divisions, may help to explain whether some of the observed changes were induced by changes in strategies and/or in production functions. Not surprisingly, they follow patterns that are very similar to the previous series. Their detailed analysis, however, is left for future research.[12]

Table 3 is concerned with the distribution of scores. Panel A summarizes the mean and standard deviation of the relative frequency of some of the most frequent scores in the Premier League. Overall, the distribution of scores tends to narrow over time. For instance, games in which at least one team scored more than 3 goals were around 22 percent early the century; they are only around 11 percent in the last two decades. The most frequent scores over the last two decades are: 1–1 (12.7%), 1–0 (10.8%), 2–1 (9.3%), 0–0 (8.3%), 2–0 (8.2%), 0–1 (7.3%). Interestingly enough, almost exactly the same numbers are found for Division II, III and IV, where the quality of the players is substantially different from that in the Premier League. This similarity is confirmed in Panel B, which reports the values of the Kolmogorov-

[10] Home and away wins have a greater goal average than ties, which is not exceedingly surprising since goals are not a necessary condition for obtaining a tie.

[11] Until the early 1950s, this decrease tends to be faster over time than that experienced by the number of goals. The decreases is slower thereafter, until the early 1980s.

[12] The average goal difference narrows over time, from around 2.1 early in the 20th century to around 1.7 towards the end of the century. The cross-sectional variability also decreases from around 0.08 in the early 1900s to about 0.06 in the 1990s. This may possibly be a reflection of (i) slow changes in strategies over time whereby teams become more defensive minded, and/or (ii) different progress of the offensive and defensive production functions. Games won by the home team have a greater difference of goals than those won by the visiting team. This suggests relatively more conservative play by visiting teams. The average number of goals per game is around 4.7 early in the century and is now around 2.6, whereas the average margin of victory was around 2.1 goals and is now about 1.7 goals. This means that the "average score" was 3.4–1.3 at the beginning of the century and is now about 2.15–0.45. As a result, on average, home teams now score 1.25 fewer goals (a 36.7% decrease with respect to 3.6 goals) and visiting teams now score 0.85 fewer goals (a 65% decrease with respect to 1.3 goals) than in the early 1900s.

Table 3. Score frequencies and differences across divisions

Panel A: Frequency of scores in the premier league

	1888–1915		1920–1939		1947–1982		1983–1996	
Score	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev	Mean	Std. Dev.
0–0	0.053	0.025	0.058	0.024	0.067	0.026	0.083	0.015
1–0	0.087	0.029	0.081	0.027	0.091	0.028	0.108	0.014
0–1	0.050	0.018	0.048	0.019	0.058	0.018	0.073	0.009
1–1	0.088	0.022	0.108	0.025	0.115	0.023	0.127	0.013
2–0	0.077	0.023	0.073	0.020	0.074	0.017	0.082	0.013
2–1	0.086	0.017	0.080	0.015	0.087	0.013	0.093	0.014
2–2	0.049	0.015	0.055	0.015	0.055	0.012	0.049	0.009
0–2	0.030	0.013	0.027	0.009	0.034	0.009	0.036	0.009
1–2	0.052	0.013	0.047	0.010	0.056	0.013	0.061	0.011
3–0	0.053	0.012	0.049	0.013	0.046	0.010	0.046	0.009
3–1	0.060	0.014	0.061	0.008	0.057	0.014	0.046	0.014
3–2	0.031	0.008	0.030	0.011	0.032	0.010	0.020	0.005
3–3	0.011	0.008	0.014	0.006	0.014	0.007	0.011	0.005
0–3	0.015	0.007	0.014	0.005	0.015	0.006	0.016	0.007
1–3	0.023	0.009	0.021	0.008	0.025	0.010	0.025	0.006
2–3	0.018	0.008	0.021	0.010	0.021	0.007	0.017	0.005
Others	0.218	0.094	0.212	0.064	0.155	0.059	0.108	0.019

Panel B: P-values of the Kolmogorov-Smirnov tests of equality of distributions across divisions

	Division I vs II		Division II vs III	Division III vs IV
	1893–1996	1947–1996	1947–1996	1947–1996
Proportion of games with:				
No Goals	0.196	0.277	0.452	0.620
Score 1–0	0.230	0.401	0.389	0.444
Score 0–1	0.178	0.387	0.603	0.256
Score 1–1	0.177	0.289	0.238	0.338
Score 2–0	0.356	0.664	0.317	0.228
Score 2–1	0.266	0.421	0.171	0.255
3 Goals	0.503	0.323	0.501	0.422
3+ Goals	0.611	0.261	0.223	0.701

Notes: League competitions were stopped during 1915/16–1918/19 and 1939/40–1945/46 because of World Wars I and II. Wilconson-Mann-Whitney rank tests for the comparison of the distributions in Panel B yield similar results in that the null hypothesis cannot be rejected at conventional significance levels.

Smirnov tests of the null hypothesis of coincidence of distributions (over time) across divisions for different scores. The results show that the null hypothesis cannot be rejected at conventional significance levels in the majority of cases.

Lastly, other time series, such as the growth rates of scores, changes in the relative proportions of the most frequent scores and others have been constructed and could be discussed. However, we believe that this description is sufficient to provide an initial flavor for the data.

3. Testing for structural changes

The basic statistical methodology that will be implemented comes from Andrews (1993) and Andrews and Ploberger (1994) who found the asymptotic distributions of a wide class of tests of structural change in econometric models. The distributions are non-standard and depend on two parameters: the number of parameters tested and the range of the sample that is examined for the break date. One disadvantage of these tests is that because the distributions are non-standard, p values cannot be calculated from previously published information. This shortcoming is overcome in Hansen (1997), who presents a method to compute convenient parametric approximation functions $p(x \mid \theta)$ to the asymptotic p-value functions $p(x)$ for the Andrews and Andrews-Ploberger asymptotic distributions.[13]

Let T_n denote a given test of structural change and T denote the associated asymptotic distribution. Let $p(x) = P(T > x)$ denote the p-value function of T and $Q(q) = p^{-1}(q)$ define its inverse function that satisfies $q = p(Q(q))$. The parametric approximation function suggested by Hansen (1997) is constructed as follows. Let $\alpha_v(x \mid \theta)$ be the vth-order polynomial in x defined as

$$\alpha_v(x \mid \theta) = \theta_0 + \theta_1 x + ... + \theta_v x^v.$$

Even though the function $p(x)$ can be well approximated by setting $p(x \mid \theta) = \alpha_v(x \mid \theta)$ for a suitable choice of the parameter θ, an improvement suggested by MacKinnon (1994) is to set $p(x \mid \theta) = 1 - \Psi(\alpha_v(x \mid \theta))$ where $\Psi(.)$ is a leading distribution function of interest. Hansen (1997) further improves the approximation by allowing the distribution to depend on an unknown parameter η and thus setting $p(x \mid \theta) = 1 - \Psi(\alpha_v(x \mid \theta) \mid \eta)$. In applications, he suggests setting $\Psi(. \mid \eta) = \chi^2(\eta)$,[14] that is, using as the approximating function

$$p(x \mid \theta) = 1 - \chi^2\left(\theta_0 + \theta_1 x + ... + \theta_v x^v \mid \eta\right),$$

where

$$\chi^2\left(z \mid \eta\right) = \int_0^z \frac{y^{\frac{\eta}{2}-1} e^{-\frac{y}{2}}}{\Gamma(\eta/2) 2^{\frac{\eta}{2}}} dy,$$

and $\theta = (\theta_0, \theta_1, ..., \theta_v, \eta)$. To fit the approximation, he uses a weighted loss function to make the difference $|p(x \mid \theta) - p(x)|$, or equivalently the difference $|p(Q(q) \mid \theta) - q|$, small. Since the uniform metric is difficult to implement numerically, he uses the L^r norm for r large (in particular, $r = 8$), and includes a weight function $w(q) \geq 0$, decreasing in q, in order to provide more precision in the approximation of small p values which are our main concern. The modified metric is thus

$$d_r(\theta) = \left[\int_0^1 |p(Q(q) \mid \theta) - q|^r \, w(q) dq\right]^{1/r}.$$

[13] See Hansen (1992) and MacKinnon (1994) for previous attempts.

[14] The main reason for this choice is that numerical plots of the densities of the Andrews' (1993) and Andrews-Ploberger's (1994) tests resemble very closely those of the chi-square.

The empirical estimations in the next section use the weight function $w(q)$ that he suggests: 1 to p values in the region $0 \leq q \leq 0.1$, 0 to p values in the region $0.8 \leq q \leq 1$, and the continuous function with quadratic decay $((0.8-q)/0.7)^2$ in $0.1 \leq q \leq 0.8$. We also set $[q_1, ..., q_N] = \{0.001, ..., 0.999\}$ to approximate the integral in the modified metric. Minimization of $d_r(\theta)$ yields the parameter value

$$\theta^* = Arg \min_{\theta \in \Theta} d_r(\theta)$$

that best fits the approximation $p(x \mid \theta)$ to the true p-value function $p(x)$. Let also the loss-minimizing p-value function be denoted by $p^*(x) = p(x \mid \theta^*)$, and the approximate p-values by $p_n^* = p(T_n \mid \theta^*)$. The true quantile function is approximated using Monte Carlo simulations using a grid on $[0, 1]$ with 1,000 evenly spaced points.

With regard to the general problem of interest, it consists of testing whether a subvector $\beta \in \mathbf{R}^p$ of a parameter $\theta = (\beta', \delta') \in \Theta \subset \mathbf{R}^m$ describing some aspect of the time series x_t under consideration equals zero when the likelihood function depends on an additional parameter k under the alternative. The parameter $k \in (0, 1)$ indicates the point of structural change as a fraction of sample size, δ_1 is a pre-change parameter vector for $t < k$, $\delta_1 + \beta$ is a post-change parameter vector, and δ_2 is a parameter vector that is constant across regimes.

We will implement Lagrange Multiplier (LM) tests of the hypothesis of no structural change ($H_0 : \beta = \mathbf{0}$) for given k, which will be denoted by $LM_n(k)$.[15] The alternative hypothesis is $H_1 : \beta \neq \mathbf{0}$.[16] Let $[k_1, k_2]$ denote the range in which the date of the structural change k is known to lie.[17] Then, the Andrews' (1993) and Andrews and Ploberger's (1994) "Ave" tests implemented in this paper have the following asymptotic null distribution:

$$\begin{aligned} Ave\ LM_n &= \frac{1}{k_2 - k_1 + 1} \sum_{t=k_1}^{k_2} LM_n(k) \underset{d}{\longrightarrow} Ave\ LM(\pi_0) \\ &= \frac{1}{\pi_2 - \pi_1} \int_{\pi_1}^{\pi_2} LM(\tau) d\tau \end{aligned}$$

where $LM(\tau) = \frac{1}{\tau(1-\tau)}(W(\tau) - \tau W(1))'(W(\tau) - \tau W(1))$, $W(\tau)$ is an $m \times 1$-vector Brownian motion, $\pi_1 = k_1/n$, $\pi_2 = k_2/n$, $\pi_0 = 1/(1 + \sqrt{\lambda_0})$, and $\lambda_0 = \frac{\pi_2(1-\pi_1)}{\pi_1(1-\pi_2)}$.[18] For the $AveLM$ distributions, $v = 3$ is used for $m = 1$, $v = 2$ for

[15] In virtually all cases, the results that will be reported in the next section are very similar to those that are obtained from Wald tests, which are available upon request. Likelihood ratio tests are not implemented because they are found not to be optimal using the weight functions considered in Andrews and Ploberger (1994).

[16] To derive asymptotically optimal tests, the likelihood function under the alternative takes the form $f(\theta + B_n^{-1}h, k)$ for some $h \in \mathbf{R}^m$ and some non-random $m \times m$ diagonal matrix B_n that satisfies $[B_n^{-1}]_{jj} \to 0$ as the sample size $n \to \infty$ for all $j \leq m$.

[17] In our case the time of the change (if it occurs) is completely unknown, so no restrictions are imposed.

[18] The results from Andrews and Ploberger's (1994) "Sup" and "Exp" tests are similar to the "Ave" tests.

$m = 2$, and $v = 1$ for $m \geq 3$. Andrews and Ploberger (1994) and Hansen (1997) offer further details of the methodology and its advantages as an empirical statistical technique. As indicated earlier, despite its valuable properties, this methodology has not been applied previously. We apply it to our time series next.

4. Empirical evidence on structural breaks

Given the large amount of p-values calculated in the statistical analysis (one for each date of each time series), we will only report the years when, according to the test, there was a structural change for a p-value ≤ 0.05. We choose this p-value because, interestingly enough, besides being a conventional significance level, we found no cases in which the results of the tests would support the hypothesis of a structural change for p-values below 0.20, a value that is clearly high for conventional standards. Therefore, the results of the tests will indicate quite distinctly the presence of structural breaks in the data set. They are collected in Table 4 and divided into six different sets.

The set T1 corresponds to the proportion of wins (W), home team wins (HW), visiting team wins (VW) and ties (T). In 1903 the T and VW series experience a structural change in Division I. World Wars I and II had an effect on Divisions I and II (dates 1920 and 1947). The change in points awarded to the winner (three instead of two) in 1982 had an effect in most time series. Similarly, structural changes are found in 1988 in the HW and VW series. The change in the offside rule in the 1925/26 season has an effect only on the T series in Divisions I and II, but not on the others.[19]

The set T2 is concerned with the ratios W/T, HW/T, VW/T and HW/VW. We find that some of the breaks in the individual series are maintained, others disappear and, interestingly, new ones appear. We conjecture that a possible reason for the differences with respect to the individual series is that the numerator and denominator may be breaking in different directions and/or with different strengths. The year 1924 appears in the VW/T series for Division II, 1970 in all HW/VW series, and 1971 in Division I for the HW/VW series and in Divisions III and IV for the W/T series. The 1925/26 change in the offside rule, the two World Wars, and the years 1897, 1982 and 1988 consistently continue to appear in most time series. Interestingly, 1982 does not appear for the HW/VW series in Division I, although it does for the HW and VW series separately. The date 1926 does not appear in ratios that do not involve ties. It appears in the W/T and VW/T series but not in the HW/T series.

An interesting aspect of the evidence in these two sets is that the results seem to suggest that in some cases there was a one year "lag" between a structural change in Division I and one in Divisions II, III and IV (e.g., 1897 and 1898, 1970 and 1971, and 1947 and 1948). This conjecture, if true, would indicate that more skilled players react or are able to adapt to certain changes faster than less skilled

[19] Law XI in the 1997 edition of the Laws of the Game establishes that: "A player is in an offside position if he is *nearer* to his opponents' goal line than both the ball and the *second* last opponent." [Italics added]. Before 1925, it was the third last opponent.

Table 4. Tests of structural change

This table reports the years when the null hypothesis that a structural change took place that year cannot be rejected at the 5 percent confidence level using the approximate asymptotic p-values for the *"Ave"* LM test. Divisions are denoted by I,II,III, IV, respectively, and "All" denotes all four divisions. W denotes wins, and HW and VW denote home team and visiting team win games, respectively.

T1 – Percentage of wins (home and visitor) and ties

Wins ... 1920 (I,II), 1982 (All)
HW games ... 1897 (I), 1920 (I,II), 1950 (I,II), 1970 (I), 1982 (I,III,IV), 1988 (I)
VW games ... 1903 (I), 1947 (I), 1977 (IV), 1982 (I,III,IV), 1988 (I,III,IV)
Ties ... 1903 (I), 1920 (I,II), 1926 (I,II), 1947 (I,II), 1948 (II), 1982 (I,II)

T2 – Relative wins and ties percentages

% Wins / %Ties ... 1897 (II), 1903 (I), 1920 (I,II), 1926 (I,II), 1947 (I,II), 1971 (III,IV), 1982 (II), 1988 (I)
% HW / % VW ... 1897 (I), 1898 (I), 1903 (I), 1947 (I), 1955 (II), 1982(II,III,IV), 1993(IV)
% HW / % Ties ... 1970 (All), 1971 (IV), 1982 (IV), 1988 (I)
% VW / % Ties ... 1903 (I), 1924 (II), 1926 (II), 1947 (I), 1971 (III,IV), 1982 (I,II)

T3 – Average number of goals

In all games ... 1901 (I,II), 1920 (II), 1926 (I,II), 1947 (I,II), 1950 (I,II), 1971 (All)
In HW only ... 1901 (I,II), 1926 (I,II), 1947 (I,II), 1950 (I,II), 1971 (I,II), 1972 (III,IV)
In VW only ... 1901 (I,II), 1926 (I,II), 1948 (II), 1973 (All)
In Ties only ... 1901 (I,II), 1926 (I,II), 1927 (II), 1948 (II), 1950(I), 1969(All),1971(II,IV)

T4 – Std. deviation of number of goals

In all games ... 1926 (I,II), 1950 (I), 1972 (I,II), 1973 (I,II), 1982 (I,III,IV), 1984(II), 1991 (I), 1993 (All)
In HW only ... 1926 (I), 1965 (I,II), 1972 (I,II), 1977 (II), 1982 (All), 1993 (All), 1994 (III,IV)
In VW only ... 1921 (I,II), 1950 (I), 1973 (I,II), 1982 (I,II), 1993 (I), 1994 (All)
In Ties only ... 1920 (I,II), 1926 (I,II), 1950 (I), 1969 (I,III,IV), 1973 (I,III,IV), 1994 (I)

T5 – Average difference of goals

In all games ... 1926 (I,II), 1950 (I,II), 1973 (I,II), 1982 (I,II,III), 1993 (II)
In HW only ... 1926 (I,II), 1950 (I,II), 1967 (IV), 1973 (I), 1982 (I,II,III)
In VW only ... 1922 (II), 1950 (I,II), 1973 (I), 1982 (III,IV)

T6 – Std. deviation of difference of goals

In all games ... 1922 (I,II), 1973 (I), 1982 (I,II)
In HW only ... 1922 (I,II), 1982 (I,II)
In VW only ... 1922 (I,II), 1939 (I), 1971 (I), 1973 (I)

players. If this is correct, structural changes in divisions with less skilled players may also spread out over various years, and hence may be more difficult to observe statistically.

Sets T3–T4 correspond to the average number of goals and their variability. These are typically considered some of the most important outcomes of the game by the football fan and the governing football organizations. In general, the number of structural breaks is smaller and appear to be more clearly defined than in previous sets. The breaks in set T3 seem to determine five different subperiods: 1888–1901, 1902–1921, 1926-World War II, 1947-early 1970s, and early 1970s–1996. The years 1901, 1926, 1947 and/or 1948, and 1971 appear in most of the series. The pattern of lags mentioned earlier arises here as well, perhaps more clearly. In this case, temporal lags are not only observed across divisions (e.g., in the T series 1926 (I) and 1927 (I and II)) but across final results as well. For example, for Divisions I and II, 1947 appears only in the HW series, whereas 1948 appears in the T and VW series; then, the year 1950 appears for Divisions I (T and HW) and II (HW). In this case, it seems possible to conjecture that, first, teams react to the structural change when playing at home (to a greater extent than when playing as a visitor), then when playing as a visiting team (both to the structural break and the initial reaction of home teams) and, lastly, they adjust again when playing at home. Similar results are observed for 1971, 1972, and 1973. Therefore, the evidence suggests that structural breaks: (i) may take longer to be accommodated in lower quality divisions, and (ii) are first reflected in the HW series, and later in the T and VW series.

The set T4 shows other interesting results. First, the years 1920 and 1921 appear only in the VW and T series, the year 1926 in almost all series (except in the VW series), and the year 1950 only in the Division I series. After 1950 there are three clusters of dates: (a) the early 1970s, especially 1972 and 1973; (b) the early 1980s, especially around 1982; and (c) the early 1990s, in particular 1993 and 1994. The longitudinal transmission of the structural break over time and across divisions and results mentioned above seems to characterize these tests as well (e.g., from HW games (I,II) in 1972 to T (III, IV) games in 1973; from HW (All) games in 1993 to T (I) and VW (All) games in 1994)).

Summing up, sets T3 and T4 indicate that: (i) World War II had an effect on the average number of goals but not on its variance; (ii) the change in points awarded to the winners in 1981/82 and the establishment of the backpass rule had an effect on the variability of the number of goals but not on its average; (iii) the temporal adjustments to the structural changes in these series seem to be different across teams and divisions.

Sets T5–T6 are concerned with the margin of victory. As a general feature, fewer breaks are experienced than in the previous series. The series for the average difference of goals have their most clear structural breaks in 1926, 1950, 1973 and 1982, while the years 1922, 1973 and 1982 represent the main breaks in the standard deviation series. The change in the offside rule in 1926 induces a structural change in the average difference of goals but not in its standard deviation. Similarly, 1922 and 1950 have very different impacts and represent a break in one moment of these series but not in the other.

Lastly, the new method developed by Bai et al. (1998) for constructing asymptotically valid confidence intervals for the date of a *single* break in multivariate time series was used to find, for each division, the most significant break in all of its time series considered simultaneously. This was World War II for Divisions I and II. When considering on a separate basis the time series pre- and post-World War II, 1925–1926 (the change in the offside rule) and 1982–1983 (the change in points awarded to winners) are the two main structural breaks in each and every division. Additional evidence such as the ARMA fits to the time series and tests for other series such as the proportions of different scores, the Gini coefficient of the distribution of points per season, and others have also been implemented. The results do not appear to uncover any interesting effects or new dates not already found in the tests presented here.

5. Concluding remarks

The analyses in this paper use an econometric methodology for dating structural breaks in tests with non-standard asymptotic distributions. The results indicate that some changes in the rules of the game of football had an effect on the mean of the distribution of certain outcomes but not on their variance and vice versa. WWII affected the average number of goals but not their variance; the amendment in the offside rule from three to two players in 1925 had an effect on the average margin of victory but not on its variability. Conversely, the change in points awarded to winners in 1981/82 and the establishment of the backpass rule in 1992 affected the variability of the number of goals but not its average. Time series that involve the number of goals per match, experience fewer and more clearly defined structural breaks than those that involve the proportion of games won and tied. For the former series, however, the change in the offside rule and the establishment of the backpass rule had a significant impact, but not for the latter series. Interestingly, the evidence shows a great degree of similarity across divisions, although it also suggests that divisions with high-skill players adapt to structural breaks faster than divisions with lower-skill players. The reactions that are observed seem to be driven mostly by the behavior of teams when playing at home.

The application of this methodology proves successful as most of the structural breaks can be traced to specific events and changes in the rules of the game.[20] Although an explicit analysis of strategic behavior and technological progress is beyond the scope of analysis of the paper, dating these breaks is also valuable in this dimension.

We have argued that as sports and their associated industries often represent socioeconomic phenomena of the first magnitude, it remains of great importance to evaluate the effects that different social, economic, and political events may have, the initiatives to change their foundations and basic structure, and the attempts to increase their attractiveness. To the best of our knowledge this is the first time that structural breaks have been dated in the sports industry. The results of the analysis may have implications for various research issues of interest in economics, given that the outcomes of the game have major effects on revenues, revenue sharing, incentive and agency problems of competition, and other aspects of the economic structure of the football industry.

References

Andrews DWK (1993) Tests for parameter instability and structural change with unknown change point. Econometrica 61: 821–856

Andrews DWK, Ploberger W (1994) Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62: 1383–1414

[20] As a referee noted, apart from the reasons that are mentioned in the paper, there are other factors related to the football industry (e.g., professionalization of players, pay tv contracts, possibility to employ foreign players in different numbers, football firms being quoted in the stock market, etc) that may have had a non-negligible impact. Although it could be argued that some of these developments in the industry are more gradual, they could certainly be taken into account as a possible explanation of certain breaks, especially the most recent ones.

Atkinson S, Stanley LR, Tschirhart J (1988) Revenue sharing as an incentive in an agency problem: an example from the national football league. Rand Journal of Economics 19: 27–43
Bai J, Lumsdaine RL, Stock JH (1998) Testing for and dating common breaks in multivariate time series. Review of Economic Studies 65: 395–432.
Baimbridge M, Cameron S, Dawson P (1996) Satellite television and the demand for football: a whole new ball game? Scottish Journal of Political Economy 43: 317–333
Bourg JF, Gouguet JJ (1988) Analyse economique du sport. Presses Universitaires de France
Brown T (1997) The ultimate football league statistics book. Association of Football Statisticians, Essex, United Kingdom, 3rd edn
Bulow JI, Geanokoplos JD, Klemperer PD (1985) Multimarket oligopoly, strategic substitutes and complements. Journal of Political Economy 93: 488–511
Deloitte & Touche (1998a) Annual review of football finance 1997
Deloitte & Touche (1998b) Financial Review of Serie A
Dixon MJ, Coles SG (1997) Modelling association football scores and inefficiencies in the football betting market. Applied Statistics 46, 265–280
Dobson S, Goddard J (1998) Performance, revenue, and cross subsidization in the football league, 1927–1994. Economic History Review 51: 763–785
El-Hodiri M, Quirk J (1971) An economic analysis of a professional sport league. Journal of Political Economy 79: 1302–1319
Fort R, Quirk J (1995) Cross subsidization, incentives and outcomes in professional team sport leagues. Journal of Economic Literature 33: 1265–1299
Garicano L, Palacios-Huerta I (2002) An empirical examination of multidimensional effort in tournaments. Brown University, Department of Economics, Working Paper 2002-17
Hansen BE (1997) Approximate asymptotic *P* values for structural-change tests. Journal of Business & Economic Statistics 15: 60–67
Hoehn T, Szymanski S (1999) The Americanization of European football. Economic Policy 28: 205–240
Jennett N (1984) Attendances, uncertainty of outcome and policy in the Scottish football league. Scottish Journal of Political Economy 31: 176–198
Kuypers T (1995) The beautiful game? An econometric study of why people watch English football. University College London Discussion Paper 96/01.
Lazear E (1989) Pay equality and industrial politics. Journal of Political Economy 97: 561–580
Lazear E, Rosen S (1981) Rank order tournaments as optimal labor contracts. Journal of Political Economy 89: 841–864
MacKinnon JG (1994) Approximate asymptotic distribution function for unit-root and cointegration tests. Journal of Business & Economics Statistics 12: 167–176
Maher MJ (1982) Modelling association football scores. Statistica Neerlandica 36: 109–118
Peel DA, Thomas DA (1988) Outcome uncertainty and the demand for football: an analysis of match attendances in the English football league. Scottish Journal of Political Economy 35: 242–249
Peel DA, Thomas DA (1992) The demand for football: some evidence on outcome uncertainty. Empirical Economics 17: 323–331
Quirk J, Fort R (1992) Pay dirt: the business of professional team sports. Princeton University Press, Princeton, NJ
Reep C, Pollard R, Benjamin B (1971) Skill and chance in ball games. Journal of the Royal Statistical Society, Series A (General) 34: 623–629
Ridder G, Cramer J, Hopstaken P (1994) Down to ten: estimating the effect of a red card in soccer. Journal of the American Statistical Association 89: 1124–1127
Sloane PJ (1971) The economics of professional football: the football club as utility maximizer. Scottish Journal of Political Economy 18: 121–146
Szymanski S, Kuypers T (1999) Winners and losers: the business strategy of football. Penguin, Harmondsworth
Szymanski S, Smith R (1997) The English football industry: profit, performance and industry structure. International Review of Applied Economics 11: 135–153
Vallet O (1998) Le football entre réligion et politique. In: Boniface P (ed) Géopolitique du Football, pp 117–131. Editions Complexe, Brussels
Walvin J (1994) The people's game. Mainstream Press, Edinburg