Scottish Journal of Political Economy

Scottish Journal of Political Economy, Vol. 59, No. 3, July 2012

COMPETITIVE BALANCE: TIME SERIES LESSONS FROM THE ENGLISH PREMIER LEAGUE

Young Hoon Lee and Rodney Fort***

Abstract

Structural break points in the First Division/English Premier League time series of competitive balance identify an Early Period, a Pre-World War II Period, a Post-War Period, and a Modern Period. The Early Period corresponds to technology diffusion (defense and tactics) along with important economic structural imposition by leagues. The war periods are common to many time series. The Modern Period's sharp decline in balance corresponds to the newest version of the Champions League in 1994/1995 and the Bosman Ruling of 1995. Rottenberg's invariance principle suggests that it would be the former, rather than the latter, responsible for the historical rate of decline that follows this structural break.

I Introduction

Assessment of time series of competitive balance measures in sports leagues has been done extensively for North American leagues – Lee and Fort (2005) and Lee (2009) for Major League Baseball and Fort and Lee (2007) for the National Basketball Association, National Football League, and National Hockey League. However, much less work has been done on the time series behavior of competitive balance in world football (soccer). Szymanski and Kuypers (1999), Szymanski and Smith (2002), and Groot (2008) all incorporate *simple tracking* of competitive balance over time in European football, and Palacios-Huerta (2004) analyzes break points in the time series of football *production components on the field*. In this paper, we apply what has come to be called the 'BP method' (Perron, 1989; Bai and Perron, 1998, 2003) to measures of competitive balance in the Football League First Division/English Football Association Premier League, 1888–2007 (henceforth, the EPL).

Both competitive balance itself and the time series behavior of measures of competitive balance are important. On the first count, under Rottenberg's (1956) uncertainty of outcome hypothesis, competitive *imbalance* can be sufficient to

*Sogang University
**University of Michigan

drive down team attendance and, eventually, league revenues. Since leagues pursue a variety of policy interventions ostensibly to improve balance, understanding the behavior of measures of competitive balance can help inform all interested in these choices by leagues.

On the second count, the time series behavior of measures of competitive balance is important as follows. First, a common approach when data are non-stationary after ordinary unit-root testing is the first-difference approach. If y_t is the time series, where $t = 1, \ldots, T$ indexes the time series length, then the first-difference series $y_{t+1} - y_t$, $t = 1, \ldots, T$ is invariably stationary. But this first-difference approach creates some problems for the usual economic interpretations (e.g., there is no longer any elasticity interpretation). The BP method uses regression analysis with a constant and a time trend to identify break points in a time series. If breaks are detected, the technique adds a dummy variable for the break year to estimate the significance and direction of the break. Thus, the BP method is useful in identifying stationary subsets of the data; in subsequent regression analysis within the subsets, there is no need to use first-differences and give up elasticity interpretations.

Second, in their studies assessing demand for European football – rather than just tracking competitive balance – Davies *et al.* (1995) and Jones *et al.* (2000) show that ignoring time series behavior in sports data could lead to spurious correlations. Identification of stationary periods in the time series is clearly an aid to subsequent level-data analysis using cross-section and pooled regression techniques.

Last, but not least, matching the break points to specific corresponding historical occurrences can suggest fruitful lines of future inquiry. Sometimes (as in this paper), clear tests of existing theory are suggested. On other occasions, new theories are suggested. As the reader will see, it is a long-enough endeavor to generate the insights themselves. Actually carrying out the applications of theory is beyond the scope of this (already lengthy) paper.

We find that there are statistically significant break points in the data that help define an 'Early Period' (1888 to the early 1900s), a 'Pre-War Period' (mid-1900s to later 1930s), a 'Post-War Period' (mid-1940s to the mid-1990s), and the 'Modern Period' (after the mid-1990s). The Early Period corresponds to 'technology diffusion' involving innovations in defense and tactics (also observed by Palacios-Huerta, 2004; Groot, 2008) as well as important economic structural imposition by leagues (equalized admissions prices and fixed player payments). The World War II impact is common in nearly all time series generated during that cataclysmic event.

We associate the Modern Period with three occurrences. First, there were alterations in the format of the Champions League in the mid-1990s that may have led to a change in talent choice by top national teams. Second, revenue equality grew dramatically at this time. Third, the Bosman Ruling of 1995 changed the definition of the transfer system. Rottenberg's (1956) invariance principle would argue that the Bosman Ruling should simply change the distribution of value created by players, but not the distribution of their playing talent across the EPL. We reiterate that analysis of this nature awaits further study beyond our scope here.

The paper proceeds as follows. In Section II, we lay out the time series approach. Estimated break points, their impacts on balance, and our characterization of the episodes they define are in Section III. In Section IV, we tie break points to historical occurrences, within the limits of the break point technique. Conclusions round out our analysis in Section V.

II Time Series Approach

Lee and Fort (2005) and Fort and Lee (2007) apply this BP method to North American pro sports leagues, and we follow their approach in our application to the EPL. In the BP method, the dependent variable is a measure of competitive balance and the independent variables are a constant term and time trend variable. We allow both the constant term and the trend term the chance to exhibit breaks in our model that follows, compared with more restricted versions. It also ends up that allowing both the constant and trend the chance to exhibit break points results in higher adjusted-R^2 compared with more restricted versions.

Technically, prior to 1992, the EPL was the Football League First Division and the Football Association English Premier League after that. Our melding of the two does no harm since the original 20 EPL clubs formed their league out of that First Division. The general BP model is as follows, estimated on competitive balance measures for the EPL, 1888–2007:

$$y_t = \alpha x'_t + \beta_j z'_t + u_t,\ t = T_{j-1} + 1, \ldots, T_j,\ j = 1, \ldots, m + 1, \tag{1}$$

where y_t is the dependent variable at time t; x_t ($p \times 1$) and z_t ($q \times 1$) are vectors of covariates and α and β ($j = 1, \ldots, m + 1$) are the corresponding vectors of coefficients; u_t is the disturbance at time t. The break points indexed ($T_1, \ldots, T_m$) are treated as unknown. For example, if one finds breaks in years $t = 3$ and $t = 12$, then $i = 1, \ldots, 3$ identifies three regimes: up to $t = 3$, from $t = 4$ to $t = 12$, and after $t = 12$.

This is a partial structural change model since the parameter vector α is not subject to change. In this paper, we allow both coefficients to change ($p = 0$ and $q = 2$). Therefore, this model is a pure structural change model where all the coefficients are subject to change as follows;

$$y_t = \beta_j z'_t + u_t,\ t = T_{j-1} + 1, \ldots, T_j,\ j = 1, \ldots, m + 1. \tag{2}$$

Ours is clearly not a structural model since it uses time trends and break points without independent variables to capture aggregate-level balance determinants. We assume that our break points, and interactions of break points and time trend, explain the average movement of balance. Break points should be picking up whether structural changes really did matter in fundamental ways to competitive balance and we'll offer historical occurrences associated with the break points to fill in the details. We hasten to note that this analysis determines only the behavior of competitive balance time series. To best fit the *impact of balance on attendance*, or to assess the *causes of competitive balance*, other measures may prove much more enlightening (Fort and Maxcy, 2003).

With respect to the specification of balance variables, a tried and true approach to capturing competitive balance in the sports economics literature is to examine the dispersion of winning percentages at the end of the regular season. The more evenly matched the teams were during the season, the more 'tightly bunched' would have been the team winning percentages at the end of the season. Lee and Fort (2005) and Fort and Lee (2007) used the well-known Scully (1989)–Noll (1988) 'ratio of standard deviations' and the 'tail likelihood' measure from Fort and Quirk (1995) as measures for tracking competitive balance in North American leagues. For the EPL, a number of factors depicted in Table 1 impact the use of these dispersion measures that have proven successful in tracking balance in North American leagues.

First, calculation of the Scully–Noll ratio based on winning percents is appropriate in nearly all North American Leagues. However, like the National Hockey League, calculation of the ratio for the EPL will involve a points system for wins. Cain and Haddock (2006) derive the different formulas for the Scully–Noll measure for winning percent, 2-points-for-a-win, and 3-points-for-a-win shown in Table 2. Fort (2007) shows that there should be no substantial difference in regression analysis using any of the three measures since their correlation coefficients are close to 1. For example, the 3-points-for-a-win version has correlation coefficients about 0.994 with the other two Scully–Noll versions in Table 2. So we lose nothing restricting ourselves to 2-points-for-a-win measure (operative for the vast majority of the years in our time series) in subsequent regression analysis.

Similarly, the tail likelihood measure falls short beyond the binary distribution for winning percent as shown by Koning (2000). Fortunately, he provides us with a 'concentration ratio' alternative also shown in Table 2. Koning's ratio is defined as the number of points obtained by a pre-specified number of the top teams divided by the maximum number of points that they could have

Table 1
Complications: divisions, teams, point systems

Period	Divisions	Period	EPL teams	Period	Points/win
1888–1891	1	1888–1890	12	1888–1980	2
1892–1919	2	1891	14	1981–2007	3
1920	3	1892–1897	16		
1921–2007	4	1898–1904	18		
		1905–1918	20		
		1919–1986	22		
		1987	21		
		1991–1994	20		
		1991–1994	22		
		1995–2007	20		

Notes: Play was precluded during the wars, 1915–1918 and 1939–1945.

Table 2
Correlation coefficients between various competitive measures

	RSD1	RSD2	RSD3	CR3U2	CR3U3	CR4U2	CR4U3	CR3L2	CR3L3	CR4L2
RSD2	1.000									
RSD3	0.994	0.994								
CR3U2	0.915	0.915	0.915							
CR3U3	0.895	0.895	0.905	0.983						
CR4U2	0.933	0.932	0.932	0.985	0.970					
CR4U3	0.906	0.905	0.915	0.964	0.984	0.980				
CR3L2	−0.898	−0.898	−0.884	−0.716	−0.714	−0.750	−0.745			
CR3L3	−0.913	−0.912	−0.908	−0.735	−0.715	−0.764	−0.738	0.977		
CR4L2	−0.919	−0.919	−0.908	−0.731	−0.718	−0.760	−0.743	0.985	0.976	
CR4L3	−0.917	−0.917	−0.915	−0.735	−0.701	−0.757	−0.718	0.943	0.982	0.972

Notes: RSDi denotes the Scully–Noll ratio for $i = 1$ (winning percent), 2 (2-points-for-a-win), and 3 (3-points-for-a-win). Cain and Haddock (2006) show that $\text{RSD1} = \frac{\sigma_p}{\sqrt{0.25N}}$, $\text{RSD2} = \frac{\sigma_2}{\sqrt{0.75N}}$, and $\text{RSD3} = \frac{\sigma_3}{\sqrt{1.734375N}}$. N is the number of matches and sigma is the actual standard deviation of either winning percent or points, as indicated. Koning's 'concentration ratio' is defined by $\text{CR}_K = (\sum_{k=1}^{K} P_{(k)})/(\text{KW}(2J - K - 1))$, where $P_{(k)}$ is the number of points obtained by the kth best team, W is the number of points awarded for a game win, and J is the total number of teams in a league. In the table, CRjkp denotes Koning's concentration ratio for j = 3 or 4 teams, $k = U$ (upper) or L (lower), and p = 2- or 3-points-for-a-win. Since all subsequent analysis uses the 2-point-for-a-win versions of all competitive balance measures, identifying indexes are suppressed in all remaining tables and charts.

obtained. If the concentration ratio is high, the top teams did not lose many points to weaker teams, so less competitive balance. Although the Scully–Noll ratio represents overall dispersion of wins or points, Koning's ratio focuses on dispersion of wins or points for top teams. While he did not mention using his ratio for teams at the bottom of the standings, we apply it to them as well. A higher value of Koning's ratio calculated for *teams at the bottom* implies *more* balance.

As any concentration ratio is sensitive to the number of firms, Koning's ratio is sensitive to league size. For example, even if the competitive balance level remains constant, the top three teams in, say, a 20-team league should be better than the top three teams in a six-team league. However, as shown in Table 1, the number of EPL teams only varied between 20 and 22 teams from 1905 on, so the sensitivity of Koning's ratio to league size should not be a problem in practice for our analysis.

Table 2 also shows the correlations between versions of Koning's ratio calculated for the top and bottom three and four teams and for both the 2-points-for-a-win and 3-points-for-a-win historical variants. The correlation coefficients between the 2-points-for-a-win and 3-points-for-a-win versions are at least above 0.980 for the top teams and at least above 0.972 for the bottom teams. As with the Scully–Noll ratio, we lose nothing in our subsequent regression analysis by using only the 2-points-for-a-win version of Koning's ratio (in operation for most of the sample period). We use the ratio calculated for three and four (top and bottom) teams for comparison purposes.

Tables 3 and 4 give descriptive statistics of our chosen competitive balance measures for the overall sample and by decade averages, respectively. Table 4 shows that temporal variations are similar between our measures. For the earliest decades, from Table 1, the number of teams increased rapidly in the 1800s (from 12 to 18) and Koning's ratio for the upper teams is supposed to increase by definition even if the level of competitive balance remains constant. On the other hand, the ratio for the lower teams is supposed to decrease by the same definition. However, in Table 4, the ratio for the upper teams *decreased* from 0.724 to 0.670 and the ratio for the lower teams *increased* from 0.370 to 0.402. This implies a dramatic improvement in competitive balance for the 1890s and early 1900s.

Competitive balance improved before WWII, but the imbalance level jumped up right after WWII, again for the 1960s, and again for the 2000s. In

Table 3
Competitive balance measures, descriptive statistics (2-point-for-a-win versions only)

	RSD	CR3U	CR3L	CR4U	CR4L
Mean	1.494	0.704	0.375	0.705	0.401
Median	1.495	0.698	0.379	0.702	0.404
Std. Dev.	0.346	0.047	0.041	0.043	0.036
Maximum	2.481	0.829	0.475	0.836	0.490
Minimum	0.742	0.617	0.279	0.628	0.314

Table 4
Competitive balance measures, averages by decades (2-point-for-a-win versions only)

	RSD	CR3U	CR3L	CR4U	CR4L
1888–1900	1.498	0.724	0.370	0.727	0.405
1901–1910	1.225	0.670	0.402	0.675	0.428
1911–1920	1.293	0.662	0.381	0.668	0.408
1921–1930	1.239	0.672	0.400	0.670	0.423
1931–1940	1.130	0.666	0.417	0.665	0.437
1941–1950	1.450	0.683	0.376	0.690	0.401
1951–1960	1.378	0.687	0.389	0.686	0.411
1961–1970	1.667	0.720	0.356	0.721	0.383
1971–1980	1.622	0.711	0.356	0.712	0.383
1981–1990	1.656	0.715	0.357	0.715	0.387
1991–2000	1.685	0.730	0.364	0.731	0.389
2001–2007	2.086	0.797	0.328	0.786	0.358

comparison, the percentage changes associated with the decreases in balance are largest for the Scully–Noll ratio and smallest for Koning's ratio calculated for the top teams. All measures show the worst balance in the current decade, 2001–2007. Pretty much across all measures, the EPL was most balanced in the 1930s (although the teens are close for Koning's ratio for the top teams).

Koning's ratio for the top three or four teams should vary differently over time from the ratio for the bottom three or four teams because the calculation for the top and bottom teams are negatively correlated by definition (a decline in the ratio for top teams implies an increase in the ratio of other teams). However, the correlation is not complete as seen in Table 4 (comparing the top and bottom three-team versions to each other, and similarly for the top and bottom four-team versions). Later, accounting for break points, fitted values also do not behave as exact mirror images of each other. This could be due to the fact that the top three or four teams in the EPL would be traditional powerhouses, while the bottom three or four teams would be different season-to-season because of promotion and relegation.

III Break Points and Balance Impacts

The foregoing observations, of course, assume there are no break points in the data and break point analysis gives misleading results if a competitive balance time series is not stationary. Unit-root tests lay these issues to rest. The results of augmented Dickey–Fuller and the Phillips–Perron tests for unit-root (regressions with a constant and a time trend) are shown in Table 5. The null hypothesis of unit-root is rejected at the 99% critical level. This result is consistent with earlier results on competitive balance measures for Major League Baseball (Lee and Fort, 2005) and other North American Leagues (Fort and Lee, 2007).

Perron's GAUSS code was used to estimate the break points and the results are in Table 6. Various recommended tests all reject the null hypothesis of no breaks at the 99% critical level for all of our measures. The first is a 'sup

Table 5
Augmented Dickey–Fuller and Phillips–Perron unit-root tests (2-point-for-a-win versions only)

	ADF (p)		PP (l)	
	Constant	Trend	Constant	Trend
RSD	−2.296** (1)	−7.217* (0)	−5.866* (6)	−7.318* (4)
CR3U	−1.558 (2)	−6.679* (0)	−5.988* (7)	−6.912* (6)
CR3L	−7.873* (0)	−9.036* (0)	−8.520* (6)	−9.123* (4)
CR4U	−1.488 (2)	−6.146* (0)	−5.456* (7)	−6.301* (6)
CR4L	−7.616* (0)	−9.240* (0)	−8.307* (6)	−9.281* (2)

Notes: The number of lags is determined by minimization of the Schwartz–Bayesian criterion for the ADF test. Bandwidth is determined by Newey and West (1994) for the PP test.
*Significant at the 99% critical level; **Significant at the 95% critical level.
ADF, Augmented Dickey–Fuller test; p, the number of lags; PP, Phillips–Perron test; l, bandwidth.

Table 6
BP approach results for the English Premier League: 1888–2007 (2-point-for-a-win versions only)

	Specifications						
	z_t = {1, time}	x_t = {0}	q = 2	p = 0	ε = 0.1	h = 10	M = 5
	Tests						
	Sup $F_T(1)$	Sup $F_T(2)$	Sup $F_T(3)$	Sup $F_T(4)$	Sup $F_T(5)$	UD_{max}	WD_{max}
RSD	33.80*	30.84*	25.50*	33.13*	25.11*	33.80*	54.34*
CR3U	55.55*	46.77*	36.63*	28.86*	28.65*	55.56*	58.70*
CR3L	28.55*	19.79*	15.74*	12.32*	11.38*	28.55*	28.55*
CR4U	68.94*	59.00*	46.70*	48.17*	39.85*	68.94*	79.02*
CR4L	21.49*	15.84*	13.09*	11.38*	10.74*	21.49*	21.49*
	Sup $F(2/1)$	Sup $F(3/2)$	Sup $F(4/3)$	Sup $F(5/4)$			
RSD	19.22*	12.81**	5.31	7.85			
CR3U	27.94*	12.35	8.69	8.69			
CR3L	11.84	7.26	3.23	2.01			
CR4U	24.98*	14.55***	20.30*	6.24			
CR4L	10.24	7.51	4.17	5.82			
	Number of breaks selected, sequential method						
RSD	CR3U		CR3L		CR4U		CR4L
2	2		1		2		1

Notes: $\varepsilon = h/T$, a trimming parameter; h = minimum length of each regime; M = upper bound. 1% significance level for the sequential test.
*Significant at the 99% critical level; **Significant at the 95% critical level; ***Significant at the 90% critical level.

Wald' type test of no structural break ($m = 0$) vs. a fixed number of breaks. We denote this 'Sup $F_T(k)$' for no break vs. k breaks. The second and third tests are double maximum tests, referred to as UD_{max} and WD_{max}. These tests are for the null hypothesis of no structural break against an unknown number of breaks given some upper bound M. The test of the 'Sequential Method' compares the null hypothesis of, say, l breaks, vs. the alternative hypothesis of $l + 1$ breaks. We denote this 'Sup $F_T(l + 1/l)$' for l breaks vs. $l + 1$ breaks. Based on the Sequential Method at the 99% critical level, the number of structural breaks is estimated as 2 for the standard deviation ratio, 2 for Koning's ratio on the top teams, and 1 for Koning's ratio on the bottom teams.

Table 7 shows estimates of break points and their confidence interval at the 99% critical level. There are a number of observations. First, as discovered in work on North American leagues, the results in Table 7 show the value of analyzing a variety of measures. Using just the ratio of standard deviations would have missed the 1901 and 1906 break points. Using just Koning's ratio (for top and bottom teams) would have missed the 1937 break point. Second, despite these differences, the results do exhibit some consistency; all measures detect the last break point, 1995–1997, and Koning's ratio (using both top and bottom teams, for either the top three and bottom three, or top four and bottom four) detects the 1901, 1906, and 1995/1997 breaks. Third, the confidence interval of the 1937 break is wide because of the war-precluded seasons, 1939–1945. Across the two different ratio measurement approaches, it is reasonable to say that break points occurred in 1901, 1906, 1937, and 1995/1997 (or both). To us, these suggest an 'Early Period' (1888 to the early 1900s), a 'Pre-War Period' (mid-1900s to later 1930s), a 'Post-War Period' (mid-1940s to the mid-1990s), and the 'Modern Period' (after the mid-1990s).

The next step in the BP process is to use the break points to actually estimate the competitive balance time series to determine the statistical significance of the shifts associated with these breaks as well as their direction. Table 8 shows the OLS estimation results using dummy variables based on break estimates. Using the coefficient estimates on the trend variables in Table 8 to interpret changes before and after the break points is straightforward. But it is not so easy to see what happens with the coefficient estimates of intercepts. The intercept estimates are single-season values for the first season in the sample, 1888, but not when a season occurs during a break point. So we turn to charts of the fitted values of the dependent variables for each of our competitive balance time series measures in Figures 1–3.

The charts help to visualize the Early Period (1888 to the early 1900s), Pre-War Period (mid-1900s to later 1930s), Post-War Period (mid-1940s to the mid-1990s), and Modern Period (after the mid-1990s). Interestingly, compared with World War II, we find no break point for World War I. The confidence intervals in Table 7 show the Early Period ending at its earliest around 1899

Table 7
Break point confidence interval summary, 99% critical level (2-point-for-a-win versions only)

Year	RSD	CR3U	CR3L	CR4U	CR4L
1901			1899–1904		1899–1905
1906		1904–1909		1904–1908	
1937	1935–1951				
1995/97	1991–1996	1994–1996		1994–1998	

Table 8

Estimation results (2-point-for-a-win versions only)

Year	RSD	CR3U	CR3L	CR4U	CR4L
Constant	1.511 (20.78)	0.767 (50.22)	0.318 (16.07)	0.769 (56.10)	0.365 (21.08)
Trend	−0.086 (−3.45)	−0.059 (−4.38)	0.078 (3.37)	−0.057 (−4.74)	0.060 (2.966)
D_1901			0.100 (4.62)		0.075 (3.97)
D_1901·trend			−0.085 (−3.64)		−0.066 (−3.25)
D_1906		−0.131 (−7.17)		−0.130 (−7.97)	
D_1906·trend		0.067 (4.99)		0.065 (5.38)	
D_1937	−0.562 (−2.65)				
D_1937·trend	0.161 (4.70)				
D_1995	−6.711 (−2.76)	−0.963 (−3.14)			
D_1995·trend	0.596 (2.80)	0.087 (3.26)			
D_1997				−0.662 (−1.81)	
D_1997·trend				0.061 (1.93)	
R^2	0.493	0.561	0.278	0.584	0.292

Notes: D_1901 = 0 if $t \leq 1901$ and = 1 otherwise. The values in parenthesis are *t*-statistics.

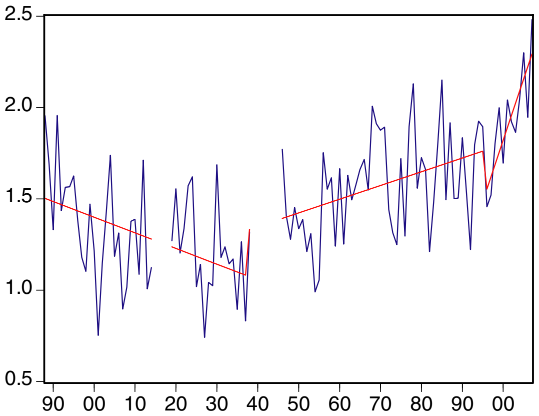

Figure 1. Fitted Scully–Noll ratio (break points: 1937, 1995).

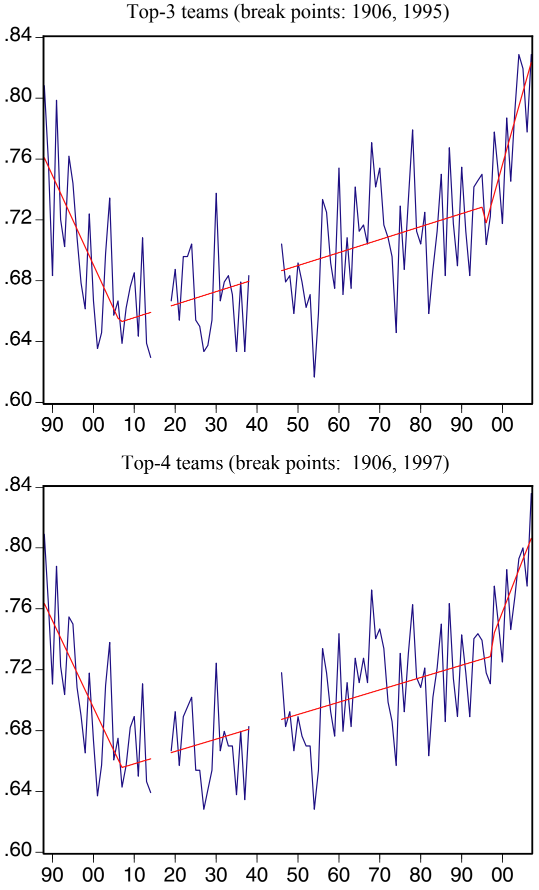

Figure 2. Fitted Koning's ratio, top teams.

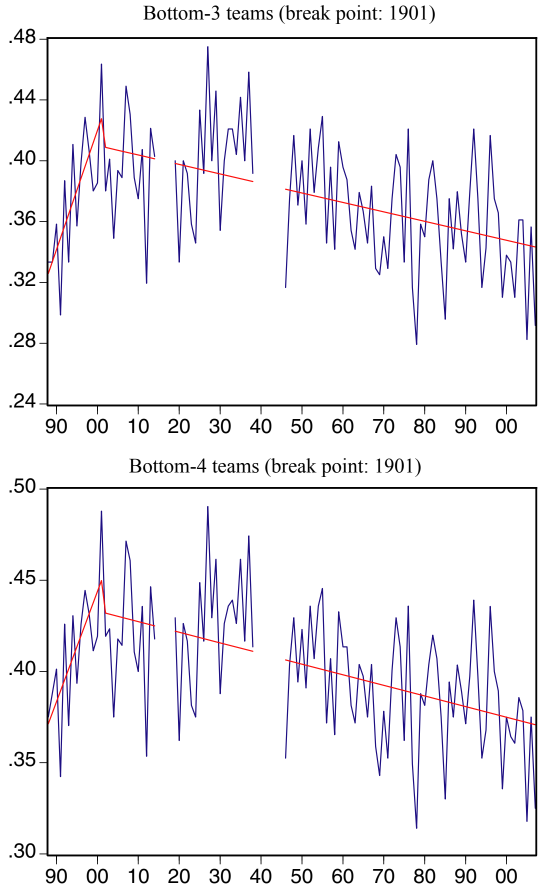

Figure 3. Fitted Koning's ratio, bottom teams.

and at its latest around 1909, well before the beginning of the 'war to end all wars' in 1914.

Figures 1–3 all show the beginning of the Early Period marked by improved balance in the EPL. However, Koning's ratio does an abrupt about-face during this period (and a shift to less balance for the bottom-team versions in Figure 3) that is missed by the Scully–Noll ratio. The Scully–Noll ratio doesn't detect the abrupt turnaround, plus a decided shift to less balance, until the Post-War Period. All fitted values in Figures 1–3 show the continued decline in balance to the Modern Period. Finally, the Scully–Noll ratio and Koning's ratio for the top teams detect a decided shift followed by a truly notable increase in the rate of decline in balance. By Koning's ratio for the top three or four teams, the annual rate of decline is just under 2% per year over the Modern Period, and almost 19% from beginning to end. Koning's ratio for the lower teams shows only a continued steady decline in balance in the EPL.

IV EPL History

As stated in the introduction, break point analysis also generates insight through association to historical events. (We are indebted to Stefan Szymanski for sharing his historical knowledge on the Early Period, income variation of clubs, the role of the maximum wage, and the modern sources of revenue inequality.) There is a steady increase in competitive balance through the Early Period, and possibly quite dramatically so by Koning's ratio. A two-fold explanation follows both changes in the game and its economic underpinnings. On the former, the number of goals scored began declining at the start of the league (Groot, 2008). This makes the increasing balance characterizing the Early Period consistent with the general diffusion of improved tactics (Palacios-Huerta, 2004). On the latter economic underpinnings, associated quite close to the 1901 break point, all teams charged the same admission price from 1890 on. In addition, a maximum wage was fixed. Limiting revenues and fixing costs could also contribute to the rise in balance early on. Further, there is a significant expansion of professional football, from two divisions to four after World War I. At the same time, variation in team revenues was small. It also is interesting that there is no break point associated with World War I, the 'war to end all wars', like there is for World War II.

It seems pretty clear that the wide confidence interval around the 1937 break point is associated with World War II, a result common to many time series. However, an explanation of why balance improved over the entire Pre-War Period but then reversed its path so clearly after the war eludes us. Although the abolition of the maximum wage in 1960 allowed the larger revenue clubs the chance to flex their muscles in the talent market, quite possibly contributing to a decline in competitive balance, this change is quite removed in time from the break point identifying the Post-War Period. And so we observe our first suggestion for future research.

We associate this mid-1990s break point with (1) alterations in the format of the Champions League, (2) growing revenue inequality, and (3) the European Court of Justice's Bosman Ruling in 1995 (Groot, 2008, also noted the first and last). Further, the Bosman Ruling may well have influenced the Champions League in important ways (noted shortly).

On alterations in the Champions League, there were earlier tournaments consisting of national champions, for example, the Coupe de Nations in 1930. But the competition of European champions known as the European Champion Clubs' Cup formally began in 1955 just after the first congress of the Union of European Football Associations (UEFA). It became the UEFA Champions League in 1992. The champion earns the European Cup. Table 9 shows that, with the name change, there also have been substantial alterations in format, especially in the 1990s. With these changes, and the increase in value of broadcast rights for the Champions League, it is reasonable that an additional focus on the top teams may have been behind our break in the mid-1990s that identifies the Modern Period.

Table 9
Champions league format chronology

Period	History
1955–1991	Knockout format, one club per country (the league champion) plus the defending champion
1991–1993	Three knockout qualifying rounds, group phase with two groups, two group winners meet in final, one club per country (the league champion) plus the defending champion
1993–1994	Knockout semi-finals added following group phase
1994–1997	One knockout qualifying round, group phase with four groups, group winners and all runners-up to eight club knockout phase, one club per country (the league champion) plus the defending champion
1997–1999	Two knockout qualifying rounds, group phase with six groups, group winners and two runners-up to eight club knockout phase, up to two clubs per country
1999–2003	Three knockout qualifying rounds, two group phases with eight first-phase group winners and all runners-up moving to four second-phase groups, second-phase group winners and all runners-up to eight club knockout phase, up to four clubs per country
Since 2003	Three knockout qualifying rounds, one group phase with eight groups, group winners and all runners-up to 16 club knockout phase, still up to four clubs per country

Notes: 1955: Many countries were represented by a team not the domestic champion. 1956–1959: The domestic runner-up was allowed to compete where the domestic champion was also European champion. Prior to 1970, aggregate draws were settled by a play-off and (if necessary) coin-toss. Since then, it has been via the away goals rule and (if necessary) a penalty shootout. The final retained the potential for a replay until the late 1970s.
Source: Wikipedia.com (2009).

In addition, growing revenue inequality cannot be ignored. Accelerating revenue inequality followed from three sources in the early 1990s. First, the TV money gives an edge to the more successful teams since income is not shared equally. Second, larger revenue teams enjoyed the development of merchandising revenue to a much larger extent than did smaller revenue teams. Finally, larger revenue teams generated more overseas income.

As for the Bosman Ruling, around this time, European national leagues confronted important changes in the player transfer system brought about by this ruling. The ruling was made in a consolidation of three separate legal cases, all involving Belgian player Jean-marc Bosman – *Union royale belge des societies de football association ASBL* vs. *Jean-Marc Bosman* (Case C-415/93, ECR I-4921); *Royal club liegeois SA* vs. *Jean-Marc Bosman and others*; and *Union of European Football Associations (UEFA)* vs. *Jean-Marc Bosman*. The eventual decision concerned free movement of labor covered by the treaty establishing the European Community. Essentially, out-of-contract players could move between national teams without any payment of transfer fees. This is quite similar to the original free agency outcome in North American sports leagues decades earlier. Players still under contract required payment of that fee by receiving clubs.

The case had a profound effect on the transfers of football players within the European Union (EU), banning restrictions on the movement of foreign EU members within the national leagues. This ban also allowed professional football players in the EU to move freely to another club at the end of their term of contract with their present team. As a result of this enhanced ability to move between teams, EU players move at the end of their contracts. In addition, many players born outside of the EU take advantage of EU naturalization, obtain a member-country passport, and enhance their employability with teams across Europe.

While it was best known for allowing out-of-contract players to move to other clubs without a transfer fee, there also is a direct relationship with the Champions League. Players from *EU member states* (including naturalized members) were not counted as foreigners on any given EU club. A continued push to move talent to those national leagues where eventual competition at the European Cup level was most likely may also be a contributing factor attributable to the Bosman Ruling.

The foregoing has the following research implications. First, analysts should adhere to break points in subsequent analysis of cross-section or pooled regression analysis or run the risk of spurious outcomes. Second, Early Period analysis of the diffusion of tactics and the impacts of the first economic structures is warranted. Third, we are left wanting on why there is no World War I break and any reasoning behind just why it is that balance should decline after World War II. Fourth, an episode analysis of the end of the maximum wage could prove enlightening for the Post-War Period.

Finally, the Bosman Ruling established a degree of mobility similar to that which occurred with free agency in North American leagues. Rottenberg's (1956) invariance principle would suggest that competitive balance would remain the same before and after the Bosman Ruling. The only difference would be that the value created by players would be reallocated from owners to players. Our evidence from Koning's ratio is contrary to the predictions of the invariance principle for the EPL. But remembering the setting for Rottenberg's invariance principle suggests the following line of inquiry.

First, can the rejection of the invariance principle be explained by the existence of transactions costs prior to the decision that continued on after the decision (Daly and Moore, 1981)? With significant transactions costs, players would not be so free to move as the invariance principle requires. Second, were there barriers to player movement erected *after the Bosman Ruling* that would undo the invariance principle prediction? North American leagues definitely responded to the initial institution of free agency with mechanisms designed to ameliorate its consequences (Plan A and B in Major League Baseball, the Rozelle Rule in the National Football League). National associations and international federations may have responded to the Bosman Ruling in predictable ways to try to reduce the reallocation of values away from clubs and associations to players. From this perspective, the dramatic increase in imbalance since the Bosman Ruling may have more to do with how European leagues responded, and the impacts of those responses, than with the ruling itself. Sorting this out will also be complicated by the added interaction with the evolution of the Champions League.

V Conclusions

We find two different ratio measures of competitive balance in the English Premier League, 1888–1907, are stationary. Using methods suggested by Bai and Perron, break points are detected in competitive balance over this history. It matters which measures are used – the two ratio measurements detect some different break points. Taking the accumulation of information across these two ratio measures, it appears there are four distinct competitive balance episodes – The Early, Pre-World War II, Post-World War II, and Modern Periods.

Knowing this is interesting enough, especially for those interested in cross-section analysis of the data. First, league actions may have improved balance enough that our technique misses their impacts. This is a place where cross-section analysis can complement our findings. Second, ours is *not* an assessment of the impacts of balance on fan demand and the interesting episodes identified here suggest just such an analysis. But we are an aid by identifying the stationary periods for balance where level-data studies will not be prone to spurious correlation problems.

For policy, our results are of some, but limited, help. We have found that, in the aggregate, arguments for policy intervention based on the idea that competitive balance has worsened in the EPL of late are well-founded. Further work is needed to determine the causes of this behavior of competitive balance during the Modern Period. One possible outcome may well be that Rottenberg's invariance principle does not hold for the English Premier League and all of the fears voiced about free agency in an international association context were on target. But we urge caution. Is it a failure of Rottenberg's logic (a logic that has fared quite well, as a matter of fact, under time series analysis of competitive balance in North American leagues) or is it that European clubs and national and international federations responded to the ruling in a way that Rottenberg's invariance principle does not address? Only future work can answer this question. In any event, competitive balance in the English Premier League has declined at a record rate for the last decade.

Acknowledgements

The National Research Foundation of Korea, Grant KRF-2009-220-B00008, funded by the Korean Government (MEST), supported this work. Lee also acknowledges financial support through a Sogang University Research Grant.

References

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**, pp. 47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, **18**, pp. 1–22.

Cain, L. P. and Haddock, D. D. (2006). Measuring parity: tying into the idealized standard deviation. *Journal of Sports Economics*, **7**, pp. 330–8.

Daly, G. and Moore, W. J. (1981). Externalities, property rights and the allocation of resources in Major League Baseball. *Economic Inquiry*, **19**, pp. 77–95.

Davies, B., Downward, P. and Jackson, I. (1995). The demand for rugby league: evidence from causality tests. *Applied Economics*, **27**, pp. 1003–7.

Fort, R. (2007). Comments on "Measuring Parity". *Journal of Sports Economics*, **8**, pp. 642–51.

Fort, R. and Lee, Y. H. (2007). Structural change, competitive balance, and the rest of the major leagues. *Economic Inquiry*, **45**, pp. 519–32.

Fort, R. and Maxcy, J. (2003). Comment: competitive balance in sports leagues: an introduction. *Journal of Sports Economics*, **4**, pp. 154–60.

Fort, R. and Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, **23**, pp. 1265–99.

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

Jones, J. C. H., Schofield, J. A. and Giles, D. E. A. (2000). Our fans in the north: the demand for British Rugby League. *Applied Economics*, **32**, pp. 1877–87.

Koning, R. H. (2000). Balance in competition in Dutch soccer. *The Statistician*, **49**(Part 3), pp. 419–31.

Lee, Y. H. (2009). The impact of postseason restructuring on the competitive balance and fan demand in Major League Baseball. *Journal of Sports Economics*, **10**, pp. 219–35.

Lee, Y. H. and Fort, R. (2005). Structural change in baseball's competitive balance: the depression, team location, and integration. *Economic Inquiry*, **43**, pp. 158–69.

Newey, W. and West, K. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economics Studies*, **61**, pp. 631–53.

Noll, R. G. (1988). *Professional Basketball*. Studies in Industrial Economics Paper No. 144. Stanford, CA: Stanford University.

Palacios-Huerta, I. (2004). Structural changes during a century of the world's most popular sport. *Statistical Models & Applications*, **13**, pp. 241–58.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, **57**, pp. 1361–401.

Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy*, **64**, pp. 242–58.

Scully, G. W. (1989). *The Business of Major League Baseball*. Chicago: University of Chicago Press.

Szymanski, S. and Kuypers, T. (1999). *Winners and Losers: The Business Strategy of Football*. London, UK: Viking (Published by the Penguin Group).

Szymanski, S. and Smith, R. (2002). Equality of opportunity and equality of outcome: static and dynamic competitive balance in European and American sports leagues. In C. P. Barros, M. Ibrahimo and S. Szymanski (eds), *Transatlantic Sport*. Cheltenham, UK: Edward Elgar, pp. 109–26.

Wikipedia.com. (2009). European Cup and Champions League History. Gleaned June 8, 2009 (available at http://en.wikipedia.org/wiki/European_Cup_and_Champions_League_history#1992_to_1996_.E2.80.94_Italian_consistency) (accessed 30 March 2012).

Date of receipt of final manuscript: 16 September 2011