Rev Ind Organ (2017) 50:27–41
DOI 10.1007/s11151-016-9526-z

Structural Change in Competitive Balance in Big-Time College Football

Steven Salaga[1] · **Rodney Fort**[2]

Published online: 18 May 2016

Abstract Competitive balance is an important element of fan preferences in sports industries. We analyze the time series behavior of competitive balance measures over the entire histories of each of the current US. "Power 5" football conferences. Competitive balance has been remarkably stable. All series are stationary by unit root tests. None of the very few structural break points that we do find coincide with economy-wide shocks (wars and the Great Depression) or with any particular college-football-wide policy alteration. This has important implications for sports researchers and for policy in the "big-time" college football industry.

Keywords Sports industry · Competitive balance · Big-time college football

1 Introduction

Competitive balance is an important characteristic of fan preferences in sports industries.[1] While there has been some attention to competitive balance relative to historic events in US college football, to our knowledge there is no formal, comprehensive time series assessment of competitive balance in Football Bowl Subdivision (FBS) conferences. In this paper we analyze the time series behavior of competitive balance measures over the histories of each of the FBS "Power 5"

[1] The idea that balance matters to fans can be traced in economics back to Rottenberg (1956).

✉ Rodney Fort
rodfort@umich.edu

[1] Texas A&M University, 342BB Blocker, 4243 TAMU, College Station, TX 77843, USA

[2] University of Michigan, 1402 Washington Heights, Rm. 3150, Ann Arbor, MI 48109-2013, USA

conferences through the 2010 conference season: the Atlantic Coast (ACC); Big 12; Big Ten; Pacific-12; and Southeastern (SEC).[2]

College sports are important to university administrators because they generate a variety of values across the rest of a university campus.[3] These values—as well as athletic department revenues that determine the portion of the university budget that will be spent on athletics—may depend on competitive balance. Thus, competitive balance becomes a management issue for university administrators through their college conferences and on up through the National Collegiate Athletic Association (NCAA). Further, competitive balance becomes a public policy issue since most athletic programs receive general university budget support, which, in turn in most cases, is partly publicly funded.

It is common in time series analysis of balance to measure uncertainty about game outcomes, playoff outcomes, and outcomes over consecutive seasons. However, playoffs have only limited and recent applicability in college football (from a time series perspective). Consequently, we employ what is now referred to as the "BP Approach" (which is named for its originators: Bai and Perron, 1998, 2003) to analyze the time series behavior of uncertainty over game and consecutive season outcomes.[4]

Our most general result is that the time series of two game-uncertainty measures and one consecutive-season-uncertainty measure are quite stable for all Power 5 conferences. The series are stationary for all Power 5 conferences by unit root tests. Further, there are no structural "break points" that are detected by the BP Approach at all for the SEC (1933–2010). For the other four conferences, there are only six total break points, and there are never more than two conferences that share a proximate break point.

None of the structural breaks are associated with either World War I or II, or the Great Depression, and none occur after the early 1990s. This remarkable stability occurred despite major changes in college football policy, most notably: the imposition of the original amateur requirement in 1906; the introduction of athletic grant-in-aid in 1956; and TV "deregulation" in 1984.[5] Rather than a form of negative "statistical finding", in this case the comparison is relevant vis-a-vis Rottenberg's (1956) invariance proposition: Is competitive balance changed (or not) by particular policy implementations? From the policy perspective, it is essential to

[2] Our data end point (2010) is arbitrary. Things are always changing in college sports conferences, and no doubt there will be room for future work as more data are generated. Our choice to exclude, say, the Big East when it had FBS football, or the Southwest as a powerhouse of its time, was due to our focus on FBS college football *as it moves forward*. There is little in the way of policy observation to offer for defunct conferences.

[3] Goff (2004) reviews a substantial literature on these other values. The literature that followed included Tucker (2004, 2005), Humphreys (2006), Tucker and Amato (2006), Humphreys and Mondello (2007), Stinson and Howard (2007), Pope and Pope (2008), Smith (2008), and Fort and Winfree (2013).

[4] The BP Approach has been used on all four North American pro sports leagues and the English Premier League. On competitive balance, see Lee and Fort (2005, 2012) and Fort and Lee (2007). On attendance, see Lee and Fort (2008) and Mills and Fort (2014).

[5] Scholarship limits and the determination of the national champion, as well as passing notice about the introduction of the GI Bill, are covered later in the paper.

answer this question in order to assess whether or not competitive balance is indeed influenced by any particular policy.

Our paper proceeds as follows: In Sect. 2 we describe our data and the measures of competitive balance that are employed in the analysis. The results of unit root and break point tests are in Sect. 3.[6] Section 4 contains our assessment of the stability of competitive balance in college football. Conclusions and suggestions are in Sect. 5.

2 Data and Measurements of Competitive Balance

The raw data for our measures that follow were collected from two primary sources for crosschecking accuracy: Quirk and Quirk (2012); and the college football pages at Sports-Reference.com (2014). In order to control for differences in out-of-conference scheduling between college football programs, only conference games are included in the analysis. The data collected for each conference cover the period of time from the creation of each conference through the completion of the 2010 season.

Each of the Power 5 conferences began play at a different point in time; those dates and the subsequent conference realignments are in Table 1. Only the ACC and the SEC have kept their original names through the history in Table 1, and some teams changed conferences after 2010.[7] Our convention is to refer to conferences by their name as of the end of our analysis in 2010: the ACC; Big 12 (rather than the Big 6 or Big 7); Big Ten (rather than Western); Pac-10 (rather than Pac-12); and SEC.

Sloane's (1976) original ideas on measures of competitive balance were categorized in Cairns (1987) by game uncertainty (GU), playoff uncertainty (PU), and consecutive season uncertainty (CSU). PU has only limited and recent applicability in college sports, so we do not examine it with the BP Approach.[8]

We note in passing that first-place finishes are concentrated in every conference among a very few teams. Every Power 5 conference had at least one and more typically two highly successful teams that were below 5.00 years/championship, which is double the equal-share rate (i.e., one championship every 10 years) for a ten-team conference.[9]

[6] Presentation of the entire results would result in a paper that is unsuited to the usual journal length. The full test and regression results are in the *Data and Statistical Tests Appendix* available online at https://www.dropbox.com/s/reo9ym5uoohbmym/SalagaFortAppendix.docx?dl=0.

[7] Others might take the history of the Big 12 back to the Missouri Valley Intercollegiate Athletic Association (1907–1927) and the Pac-12 only back to the Athletic Association of Western Universities (1959–1967), rather than including the Pacific Coast Conference (1916–1958) that was broken up due to scandal.

[8] College conference championship games have only been around since Arkansas and South Carolina joined the SEC in 1992–1993. One could envision "post season" access to bowl games, but the purpose of nearly none of them is to crown a champion. Indeed there was no national championship game, per se, until the Bowl Championship Series put one in place for the 2007 bowl season.

[9] The full data and calculation results for historical conference championships are in Tables A1–A5 in the online appendix (see footnote 6).

Table 1 Power 5 conference realignment through 2010–2011. *Source:* Quirk and Quirk (2012). Sports-Reference.com (2014)

ACC (1953, 7)	Pac-10 (Pacific Coast, 1916, 4)
+Virginia (1954, 8)	+Washington St. (1917, 5)
−S. Carolina (1972, 7)	−Washington St. +Stanford (1918, 5)
+GA Tech (1983, 8)	+Washington St. (1919, 6)
+Florida St. (1992, 9)	+S. Cal +Idaho (1922, 8)
+VA Tech +Miami (2004, 11)	+Montana (1924, 9)
+Boston College (2005, 12)	+UCLA (1928, 10)
	−Idaho −Montana −Oregon −Oregon St. −Stanford −Washington −Washington St. (1943, 3)
Big 12 (Big 6, 1928, 6)	+Washington (1944, 4)
+Colorado (Big 7, 1948, 7)	+Washington St. +Oregon +Oregon St. +Montana (1945, 8)
+Oklahoma St. (Big 8, 1960, 8)	+Stanford +Idaho (1946, 10)
+Texas +Texas Tech +Texas A&M +Baylor (Big12, 1996, 12)	−Montana (1950, 9)
	−Washington St. −Oregon St. −Oregon −Idaho (AAUW, 1958, 5)
	+Washington St. (1962, 6)
Big Ten (Western, 1896, 7)	+Oregon St. +Oregon (1964, 8)
+Indiana +Iowa (1900, 9)	Renamed Pac-8 (1968, 8)
−Northwestern (1906, 8)	+Arizona +Arizona St. (Pac-10, 1978, 10)
−Michigan (1907, 7)	
+Northwestern (1908, 8)	SEC (1933, 13)
+Ohio St. (1913, 9)	−Sewanee (1940, 12)
+Michigan (1917, 10)	−Alabama −Auburn −Florida −Kentucky −Mississippi −Mississippi St. −Tennessee −Vanderbilt (1943, 4)
−Chicago (1940, 9)	+Alabama +Auburn +Florida +Kentucky +Mississippi +Mississippi St. +Tennessee (1944, 11)
+Michigan St. (Big Ten, 1953, 10)	+Vanderbilt (1945, 12)
+Penn St. (1993, 11)	−GA Tech (1964, 11)
	−Tulane (1966, 10)
	+S. Carolina +Arkansas (1992, 12)

Parenthetical includes new conference name, if any, the year the listed team joined or left the conference, and the final number of teams in the conference after the change. For example, under the Big 12, +Colorado (Big 7, 1948, 7) means that when Colorado joined the Big 6 in 1948, it changed name to the Big 7 with 7 teams

Mills and Fort (2014) provide a survey of the various measures that have been utilized in past work. GU measures cover the dispersion of team winning percentages and the degree of game closeness. Measures of CSU have attempted to capture the degree of variance in team quality over successive seasons. The use of each of these three measures of dispersion, separately as a measure of balance, allows for a comprehensive view of college football balance over the history of the sport.

For dispersion of winning percentages (GU), we use the "ratio of standard deviations" (RSD) that was introduced by Noll (1988) and Scully (1989).[10] $RSD_t = ASD_t/ISD_t$, where ASD_t is the "actual standard deviation" of winning percent at the end of conference play and ISD_t is an "idealized standard deviation" for a conference where the probability that one team beats the other is 0.5.

For the simple binomial win-or-lose outcome (no ties), Fort and Quirk (1995) show that $ISD_t = 0.5/\sqrt{G_t}$, where G_t is the number of games in season t. As $RSD_t \rightarrow 1$, the conference emulates a league where the chance that any team beats another is literally 0.5. One nice feature of RSD_t is that it accounts for both changes in season length and the number of teams in a conference in any year.

Table 2 contains the decade RSD_t averages for each Power 5 conference. After some differences early on, from the 1960s onward the college conference RSD_t values are comparable to those (in the last column) of the National Football League (NFL). Indeed, the values are more similar to the NFL than to any other pro sport (Fort 2011, Table 6.3, p. 168). More recently, there has been a clear-cut leader in terms of balance as the Pac-10 has the lowest cumulative RSD_t value since the 1960s. The Big 12 has been the least balanced conference over the same time period. The averages across all decades reveal the occasionally volatile behavior of this measure of balance, decade-to-decade.

For closeness of games (also relevant to GU), we need a measure that is applicable over the entire history of the sport, accounts for the changing nature of the contest outcomes in terms of total scoring, and aggregates to one observation per season. We use the following "margin of victory ratio" (MVR_t) *for each conference* in each season t from Salaga (2015) and Mills and Salaga (2015):

$$MVR_t = \frac{1}{G_t}\sum_{g=1}^{G}\frac{MV_{g,t}}{TP_{g,t}},$$

where t is conference year t = 1, …, T, $MV_{g,t}$ is the margin of victory in conference game g = 1, …, G_t in year t, and $TP_{g,t}$ is total points scored in a conference game g in year t. A tie is the most balanced game with $MV_{g,t} = 0$. A season of ties would have $MV_{g,t} = 0 \forall g$, and so produce $MVR_t = 0$ for that conference in that year (we arbitrarily set it to zero for 0–0 ties). A shutout generates $MV_{g,t} = TP_{g,t}$ so that $\frac{MV_{g,t}}{TP_{g,t}} = 1$ for that game. A season of shutouts would have $\sum_{g=1}^{G} 1 = G_t$, so that $MVR_t = 1$ for that conference for that year.

We normalize on total points as well as the number of games, because it is clear that total points have been rising over time. For example, from 1900 to 1910 the total points that were scored in a Big Ten conference game averaged 25.07 (Salaga 2012). Apparently over time offensive prowess either developed or became more important in winning (or the type of winning that fans wanted to see). From 2000 to

[10] The "tail likelihood" measure from Fort and Quirk (1995) puts the attention on the outliers (big losers and big winners) rather than on outcomes around the mean of the winning percentage. In the data to follow, RSD takes on values that are much lower than in sports such as baseball or basketball (Fort 2011, Chapter 6), which suggests that there really isn't as much going on in the tails for college football competitive balance.

Table 2 Power 5 conference and NFL RSD_t decade averages

	ACC	Big 12	Big Ten	Pac-10	SEC	NFL
1890–1899	–	–	1.59	–	–	–
1900–1909	–	–	1.92	–	–	–
1910–1919	–	–	1.64	1.35	–	–
1920–1929	–	1.38	1.52	1.57	–	1.96
1930–1939	–	1.40	1.61	1.60	1.65	1.70
1940–1949	–	1.47	1.46	1.66	1.56	1.85
1950–1959	1.60	1.51	1.43	1.58	1.53	1.52
1960–1969	1.31	1.67	1.47	1.43	1.64	1.69
1970–1979	1.56	1.52	1.62	1.51	1.63	1.62
1980–1989	1.51	1.68	1.63	1.37	1.50	1.47
1990–1999	1.68	1.64	1.61	1.45	1.62	1.51
2000–2010	1.41	1.53	1.48	1.52	1.51	1.56
1960–2010	1.49	1.61	1.56	1.46	1.58	1.57
All years	1.50	1.55	1.58	1.51	1.58	1.65

Initial values for each conference represent averages for truncated decades based on when each conference began play

2010 the average total points scored in a Big Ten game had increased to 51.49 (Salaga 2012).

Normalizing on total points also casts higher scoring games as closer than lower scoring games with the same margin of victory. Consider a 28-21 game and a 14-7 game. Normalization casts the former as closer than the latter because there were so many more scoring plays in the former, yet the margins of victory for the two games were the same.

Like RSD_t, MVR_t accounts for changes in season length and the number of conference teams in any season. $0 \leq MVR_t \leq 1$ and game uncertainty is greater, and contests more balanced, as $MVR_t \rightarrow 0$. Alternatively, contests are less balanced, and game uncertainty is less, as $MVR_t \rightarrow 1$.

Decade averages for MVR_t are in Table 3, and it is clear that contests have become closer (so that GU by this measure is greater) over time in every Power 5 conference. The underlying data reveal that this is because total points scored (the denominator in our measure) has gradually increased over time while the numerator has remained relatively flat over time.

Finally, we use the year-to-year team conference winning percentage correlation for CSU (Butler 1995; subsequent uses are reviewed in Mills and Fort 2014): $WPC_t = CORREL(W_t, W_{t-1})$, where W_t is winning percentage in year t, and W_{t-1} is winning percentage in year $t - 1$. This measure is used to determine the degree of "churning" in the season-to-season conference standings, with $-1 \leq \mathrm{WPC_t} \leq 1$.

A correlation of 1 suggests no movement in conference standings on a year-to-year basis, which represents the least possible balance. A value of -1 indicates perfect negative correlation, which represents the greatest possible balance since conference standings would turn over completely on a year-to-year basis for this value. A value of 0 suggests no correlation in conference standings, so it shows improved balance relative to a value of 1, but less balance in comparison to a value

Table 3 Power 5 conference MVR_t decade averages

	ACC	Big 12	Big Ten	Pac-10	SEC
1890–1899	–	–	0.6955	–	–
1900–1909	–	–	0.6914	–	–
1910–1919	–	–	0.6835	0.7009	–
1920–1929	–	0.6382	0.6472	0.6773	–
1930–1939	–	0.6241	0.6334	0.6396	0.6281
1940–1949	–	0.5917	0.5158	0.6018	0.5847
1950–1959	0.5436	0.5069	0.4686	0.4853	0.5235
1960–1969	0.4200	0.5027	0.4609	0.4681	0.5005
1970–1979	0.4303	0.4262	0.4711	0.3696	0.4397
1980–1989	0.3367	0.4757	0.3927	0.3564	0.3657
1990–1999	0.3894	0.4007	0.3757	0.3277	0.3770
2000–2010	0.3260	0.3710	0.3248	0.3344	0.3253
1960–2010	0.3805	0.4353	0.4050	0.3712	0.4016
All years	0.3992	0.4896	0.5196	0.4815	0.4600

Initial values for each conference represent averages for truncated decades based on when each conference began play

of −1. Decade averages are in Table 4. The correlation in winning percentages, season-to-season, is comparable to the NFL early on; but since the 1960s the professional version of football shows much less correlation than the college version. The Big 12 clearly displays the highest correlation in winning percentages.

3 Stationarity and Break Point Results

The schematic for untangling sports time series is in Fort and Lee (2006). Our approach is identical to the Bai and Perron "BP Approach" in Lee and Fort (2005) and Fort and Lee (2007).[11] The BP approach uses regression with a constant and a time trend. If break points are detected, dummy variables for the year of the break points are added to the regression to estimate the significance and direction of structural changes. The basic assessment in the BP Approach follows from a multiple regression with m breaks (m + 1 regimes):

$$y_t = x_t'\alpha + z_t'\beta_j + u_t, \quad t = T_{j-1} + 1, \ldots, T_j, \quad j = 1, \ldots, m+1.$$

The dependent variable is y_t, which in our case is the competitive balance time series of interest. The vectors of covariates are x_t and z_t (p × 1 and q × 1, respectively). The corresponding coefficient vectors are α and β_j. The disturbance term is u_t. $T_1,\ldots,T_m$ are treated as the unknown break points.

There are four possible tests for multiple structural changes, depending on the a priori choice of alternative number of breaks. *Sup* $F_T(k)$, k = 1,…,5, *UDMax*, and

[11] Andrews (1993) focused on a single structural change, while Bai (1997, 1999) and Bai and Perron (1998, 2003) moved on to consider multiple structural changes.

Table 4 Power 5 conference and NFL WPC_t decade averages

	ACC	Big 12	Big Ten	Pac-10	SEC	NFL
1890–1899	–	–	0.5244	–	–	–
1900–1909	–	–	0.4145	–	–	–
1910–1919	–	–	0.3625	0.1244	–	–
1920–1929	–	0.2958	0.3895	0.5007	–	0.4958
1930–1939	–	0.3469	0.5187	0.5331	0.5999	0.5903
1940–1949	–	0.7440	0.2182	0.4194	0.4885	0.5529
1950–1959	0.5899	0.6970	0.3572	0.3801	0.4251	0.4433
1960–1969	0.4435	0.5369	0.4422	0.6780	0.6591	0.5107
1970–1979	0.4312	0.5584	0.6981	0.5408	0.5607	0.6155
1980–1989	0.5845	0.7747	0.6847	0.7017	0.5006	0.4214
1990–1999	0.5324	0.7909	0.6167	0.3382	0.5749	0.3712
2000–2010	0.5635	0.5469	0.4262	0.5405	0.5048	0.2885
1960–2010	0.5110	0.6416	0.5736	0.5598	0.5600	0.4385
All years	0.5202	0.6195	0.4674	0.5025	0.5356	0.4738

Initial values for each conference represent averages for truncated decades based on when each conference began play

WDMax each tests for at least one break point, and then the *SupF* tests for additional break points.[12] The BP Approach favors sequential application of these four tests. We use the freely available GAUSS program that was originally provided by professors Bai and Perron for non-profit academic work.

We first test for unit root in the three competitive balance time series: RSD_t; MVR_t; and WPC_t.[13] If non-stationarity were detected, the results of the BP Approach with the assumption of stationarity would be misleading. Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests reject the hypothesis of unit root nearly uniformly for both the ADF and PP tests at the 99 % level for all three measures. There are only two exceptions and even those still reject the hypothesis at the 90 % level.[14] Given this, we take the series to be stationary for all of the Power 5 conferences.

We next address the determination of break points: *Sup* $F_T(k)$, k = 1,…,5, *UDMax*, and *WDMax* tests for at least one break point, and then *SupF* tests for additional break points, were performed on the three time series for each of the

[12] Sequential tests allow for the ability to determine the optimal number of breaks within a given series. The *UDMax* and *WDMax* tests identify the presence of breaks (with no set maximum number of breaks) against the null of zero breaks. The *Sup* $F_T(k)$ tests specify statistical significance for a given number of breaks *(k)* against a null of zero breaks.

[13] Bounded time series such as ours can render unit root tests misleading. However, in actuality, none of our series hit the boundaries so unit root testing remains insightful. Our reference is Cavaliere and Xu (2014).

[14] ADF and PP tests are the standard choice for unit root testing in the BP literature and test results are in Table A6 in the online appendix (see footnote 6).

Table 5 Conference break points and confidence intervals

Conference	RSD_t	MVR_t	WPC_t
ACC	1966 (1965–1981)	1970 (1969–1972)	
	1993 (1987–1994)	1988 (1987–1991)	
Big 12			1950 (1949–1967)
			1985 (1983–1986)
Big Ten	1964 (1962–1977)		1963 (1961–1975)
Pac-10		1947 (1944–1952)	
SEC			

Break point tests results are at the 95 % level of confidence. The regressions that determined the significance and sign of the breaks, along with any trend after each, produced the confidence intervals (90 % level of significance)

Power 5 conferences.[15] Subsequent regressions test the significance and direction of the break-point shifts along with trends that followed each break point.[16] At the 95 % level for the break-point determination, and with the 90 % confidence bands from the regressions, the break points are summarized in Table 5. The BP Approach is flexible enough to allow asymmetric confidence bands. For example, in the ACC case for RSD_t, while the first break in Table 5 was estimated to occur in 1966, it is more likely that it occurred after 1966 than before.

While the behavior of breaks and trends can be seen in the regression results that generated Table 5, Fort and Lee (2007) demonstrated an easier visual interpretation using actual and fitted time series. Figure 1 shows the actual and fitted time series for just the regressions that produced any break points.[17]

In what follows we adopt the following descriptive for break point discussion: YEAR (shift direction, trend direction). "YEAR" is the year of the shift; "shift direction" is either a "+" upward shift or a "−" downward shift; and "trend direction" is a "+" upward trend after the shift, a "−" downward trend after the shift, or a "0" no trend after the shift.

Figure 1 shows that there were two break points for RSD_t in the ACC. GU had steadily improved (winning percent dispersion declined) since the inception of the ACC. But then a shift for the worse occurred in 1966 (+, +), and its following trend was also for the worse. The second ACC break was also for the worse in 1993

[15] Break point test results are in Tables A7–A9 in the online appendix for the results for each of RSD_t, MVR_t, and WPC_t, respectively (see footnote 6). As a robustness check, we tried to recreate the significance and direction of the break-point results using generalized linear modeling (GLM). Conference-level measures were the dependent variable, and indicators for the years at which we find break points, plus a constant, and a trend were the explanatory variables. For all of the series where we find a single break point using the BP Approach, our GLM results give the same results for both significance and direction of the break point. For the series where we find two break points, our GLM results give the same the direction of the breaks, and significance occurs in 3 of 6 cases. The BP Approach and GLM results are in general agreement.

[16] The regression results are in Table A10 in the online appendix for the results for each of RSD_t, MVR_t, and WPC_t, respectively (see footnote 6).

[17] The actual and fitted results for all Power 5 conferences, including those not in Figure 1, are in Figures A1–A3 in the online appendix (see footnote 6).

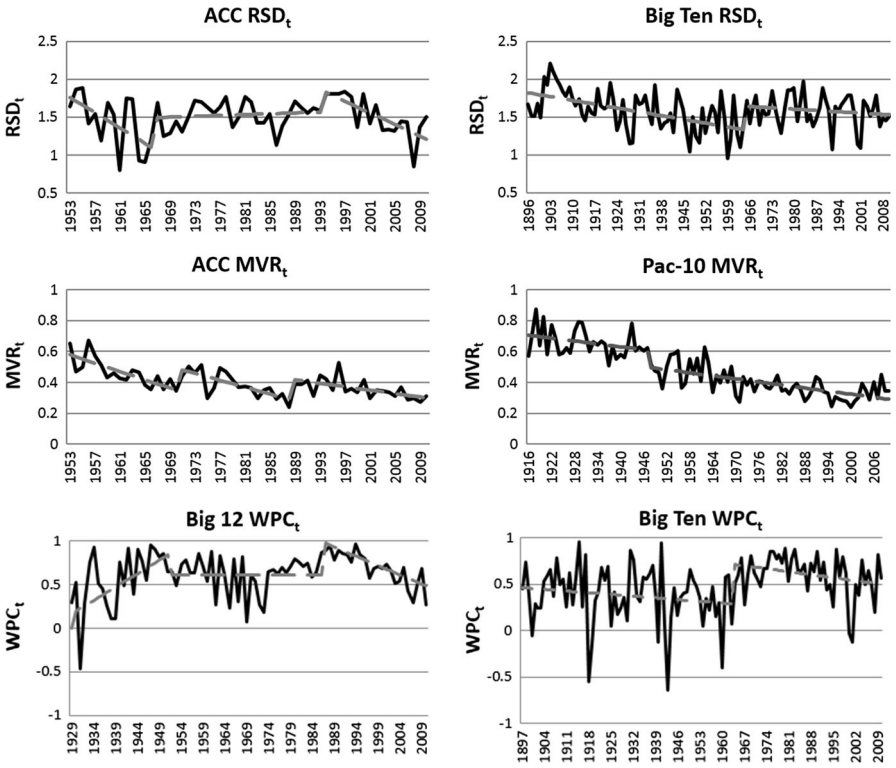

Fig. 1 Actual (*black*) and fitted (*gray dashed*) RSD_t, MVR_t, and WPC_t, for Power 5 conferences

(+, −), but the following trend was for improvement. Indeed, GU improved to the level that was observed in the early 1960s. The single break in RSD_t for the Big Ten was in 1964 (+, −). GU had been improving before and improved just slightly after that structural break.

Figure 1 also shows two break points in the ACC for our other measure of GU: MVR_t (margin of victory ratio). Break points in 1970 (+, −) and 1988 (+, −) resulted in larger margins of victory, but these appear not to have been able to stem the steady pressure toward closer games (more competitive balance) that is revealed by the trends that surround those break points. The same overall trend in MVR_t goes for the Pac-10, but its single break point in 1947 (−, −) hastened that continuing improvement in balance (decline in margins of victory).

Finally, for our measure of CSU, there were two breaks in WPC_t (win percent correlation over time) for the Big 12. The first in 1950 (−, 0) reversed the trend of increasing correlation since the inception of the conference, and the series remained completely flat for 35 years. The next break point in 1985 (+, −) saw the correlation increase again but then steadily decline (improved balance) back to around 0.50. In the Big Ten, the single shift in 1963 (+, −) reversed what had been steady improvement since the beginning of the conference in 1896. However, the following trend was back toward improved balance.

4 Assessment of the Results

The preceding results suggest three general observations: first, while there was variability—especially early on—as is indicated by our measures and estimation methodology balance got no worse and, for some conferences on some dimensions, improved over the last few decades. Second, the stability of competitive balance is remarkable compared to past work.[18] Finally, nothing about structural changes appears to have anything to do with big-time college football industry policies that have been touted as required for competitive balance.

The first observation can be seen as follows: There was no trend in RSD_t for the Big 12, Pac-10, and SEC.[19] From Fig. 1, GU has been either the same or improving by this measure for both the ACC and Big Ten since the 1960s. For MVR_t, the Big 12, Big Ten, and SEC showed a downward trend throughout their entire histories without break points. Again, from Fig. 1, despite break points that operated in the opposite direction, the downward trend in the ACC and Pac-10 are unmistakable. Finally, no trend was apparent for WPC_t in the ACC, Pac-10, and SEC. From Fig. 1, the correlation in annual winning percentages has either remained the same or declined (after contrary break points) in the Big 12 and Big Ten since the 1950s. So, for recent decades, the verdict is either constant or improved competitive balance for both GU and CSU in the Power 5.

The second observation of remarkable stability rests on a comparison with past work. Unlike Major League Baseball (Lee and Fort 2005), but more like the remaining major leagues (Lee and Fort 2007), none of the break points for the Power 5 conferences have anything to do with major non-sports events such as world wars or the Great Depression.[20] Indeed, the general stability (in terms of absence of break points) for the Power 5 conferences extends to encompass alterations in college conference membership and broader social outcomes such as racial integration.[21]

The final observation pertains to college football-wide "industry policy". No breaks occurred coincident with the imposition of the amateur requirement (1906) or subsequent relaxation of that requirement with the grant-in-aid in 1956.[22] This is consistent with Rottenberg's (1956) invariance principle that is well known in the

[18] It may be the case that our measures of competitive balance are not robust enough to show the impact of events. However, in the many other works that are referenced as using these variables, the measures were insightful.

[19] To follow part of this discussion, the reader will need to return to Figures A1–A3 in the online appendix (see footnote 6).

[20] Salaga (2015) analyzed the GI Bill (1946), a general national education policy measure thought to have influenced balance in college football. Our work reinforces his episode analysis finding that the GI Bill did not coincide with changes in competitive balance. However, Mills and Salaga (2015) do find breakpoints in NCAA college basketball that align closely with the implementation of the GI Bill.

[21] Quirk (2004) found short-term impacts on balance with conference switching. The work cited above on pro sports also found the pro version of conference switching—expansion and relocation—and integration to coincide with break points. With respect to integration, Mills and Salaga (2015) find a number of breaks proximate to this time period in college basketball.

[22] The grant-in-aid allowed coverage of tuition, room and board, and books. See Falla (1981) and Byers (1995).

sports economics literature.[23] However, the finding is counter to claims that the amateur requirement preserves competitive balance, which is a mainstay of the NCAA in its many legal defenses against antitrust suits that are aimed at removing the requirement. There may be other arguments in favor of the amateur requirement, but the findings here do not support arguments that are based on preserving competitive balance.

Finally on industry-wide policy, no break points coincide with any of the following: The NCAA limited scholarships in 1977 and again in 1992. The conferences took over TV broadcast negotiations from the NCAA after the US Supreme Court's antitrust decision against the NCAA in *NCAA v. Board of Regents of the University of Oklahoma* in 1984.[24] The determination of the national champion was fundamentally altered under the formation of the Bowl Championship Series (BCS), begun a bit haphazardly in 1992 but in full force in 1998.[25]

There is some evidence that a few conference-specific issues may be in play: While no football-wide breaks occur for the war years, there are break points that are proximate to shrinkage and membership rebounds in both the Pac-10 and the SEC: 1942–1944 (see Table 1). Some of the break points also appear to coincide with outlier coaching performance: Bobby Bowden at Florida State University in the ACC (1976–2009, 31 bowls, 2 National Championships); Bud Wilkinson at the University of Oklahoma in the Big 12 (1947–1963, 8 bowls, 3 National Championships); and the fierce rivalry in the Big Ten between Woody Hayes at Ohio State University (1951–1978, 12 bowls, 5 National Championships) and Bo Schembechler at the University of Michigan (1969–1989, 17 bowls, 0 National Championships). But future work on any such speculation is required and beyond the scope and ambition of our analysis here.[26]

5 Conclusions and Suggestions

We apply the BP Approach to two measures of game uncertainty and one measure of consecutive season uncertainty for the current Power 5 conferences, from their inception to 2010. First, we find evidence that both game and consecutive season uncertainty have either stayed the same or improved across the Power 5 over the last few decades. Second,

[23] The invariance principle holds that the distribution of talent is invariant with respect to who keeps the value created by athletes. Talent goes to its highest valued use whether athletes are paid their marginal revenue product or less due to labor market restrictions. Our finding also reinforces the episode analysis in Salaga (2015) that finds grants-in-aid had no impact on balance in college football.

[24] 468 US 85 (1984).

[25] Scholarship limits: Sutter and Winkler (2003). TV deregulation: Bennett and Fizel (1995), Eckard (1998), Siegfried and Burba (2004), and Salaga (2015). The BCS formation is common knowledge and its website was still in operation as of this writing (www.bcsfootball.org).

[26] For example, on coaches, no break points were associated with other outlier performances such as: Fielding Yost at the University of Michigan (1901–1926, pre-bowls, 6 National Championships including 4 in a row, 1901–1904); John McKay at the University of Southern California (1960–1975, 9 bowls, 4 National Championships); Bear Bryant at the University of Alabama (1958–1982, 24 bowls, 6 National Championships); or Joe Paterno at Pennsylvania State University (1966–2011, 37 bowls, 2 National Championships).

we find remarkable stability despite world wars and the Great Depression. Finally, no structural change coincides with what has been touted as game-changing alterations in college football policy. This last point suggests two things:

On the one hand, Rottenberg's invariance principle appears operational in college football. In policy debates and antitrust lawsuits, the NCAA has adhered without fail to the idea that the amateur requirement is essential to the maintenance of competitive balance. None of our findings support that contention. On the other hand, no breaks are found that are coincident with any of the other football-wide policy changes either. Arguments that involve impacts on balance cannot justify changes in the number of scholarships, TV deregulation, or alteration in the determination of the national champion.

Our results also have implications for future work: The presence of break points dictates caution in any level-data analysis in the future. As Davies, Downward, and Jackson (1995) and Dobson and Goddard (1998) warned, level-data regressions (e.g., panel analysis) could suffer spurious correlation outcomes when the sample period spans structural breaks.

In addition, structural modeling and analysis of possible policy alterations and coaching technology relative to the break points that we find—or their absence—is in order. On policy, it is an interesting question as to whether general stability and the improvements in types of balance that are indicated in this analysis happened because of, or in spite of, centralized management of college football by university administrators. Was the general improvement in game and consecutive season uncertainty due to management through conferences, the NCAA, and the BCS (now the College Football Playoff since 2014–2015)? Or has it occurred because of the relative equalization of population across regions and the gradual nationalization of the college football talent market?

Finally, competitive balance matters for Rottenberg's other hypothesis on outcome uncertainty (i.e., fans prefer close games where their team wins, as well as equitable access to post-season play). As has been the course in time series work on pro sports, moving from the behavior of balance, per se, to the impact of balance on fan welfare (attendance analysis) is in order. From that perspective, the question is surely whether or not the type of overall stability, along with the improvement in game and consecutive season uncertainty, are in line with fan welfare.

References

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica, 61*, 821–856.

Bai, J. (1997). Estimating multiple breaks one at time. *Econometric Theory, 13*, 315–352.

Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics, 91*, 299–323.

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

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

Bennett, R. W., & Fizel, J. L. (1995). Telecast deregulation and competitive balance: NCAA Division I football. *American Journal of Economics and Sociology, 54*, 183–199.
Butler, M. (1995). Competitive balance in Major League Baseball. *American Economist, 39*, 46–52.
Byers, W. (1995). *Unsportsmanlike conduct: Exploiting college athletes*. Ann Arbor, MI: University of Michigan Press.
Cairns, J. (1987). Evaluating changes in league structure: The reorganization of the Scottish Football league. *Applied Economics, 19*, 259–275.
Cavaliere, G., & Xu, F. (2014). Testing for unit roots in bounded time series. *Journal of Econometrics, 178*, 259–272.
Davies, B., Downward, P., & Jackson, I. (1995). The demand for Rugby League: Evidence from causality tests. *Applied Economics, 27*, 1003–1007.
Dobson, S. M., & Goddard, J. A. (1998). Performance and revenue in Professional League Football: Evidence from Granger causality tests. *Applied Economics, 30*, 16541–16551.
Eckard, E. W. (1998). The NCAA cartel and competitive balance in college football. *Review of Industrial Organization, 13*, 347–369.
Falla, J. (1981). *NCAA: The voice of college sports*. Mission, KS: National Collegiate Athletic Association.
Fort, R. (2011). *Sports economics* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
Fort, R., & Lee, Y. H. (2006). Stationarity and Major League Baseball attendance analysis. *Journal of Sports Economics, 7*, 408–415.
Fort, R., & Lee, Y. H. (2007). Structural change, competitive balance, and the rest of the Major Leagues. *Economic Inquiry, 45*, 519–532.
Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature, 23*, 1265–1299.
Fort, R., & Winfree, J. (2013). *15 Sports myths and why they're wrong* (pp. 22–40). Stanford, CA: Stanford University Press.
Goff, B. (2004). Effects of university athletics on the university: A review and extension of empirical assessment. In J. Fizel & R. Fort (Eds.), *Economics of college sports* (pp. 65–86). Westport, CT: Praeger.
Humphreys, B. R. (2006). The relationship between big-time college football and state appropriations to higher education. *International Journal of Sport Finance, 1*, 119–128.
Humphreys, B. R., & Mondello, M. J. (2007). Intercollegiate athletic success and donations at NCAA Division I institutions. *Journal of Sport Management, 21*, 265–280.
Lee, Y. H., & Fort, R. (2005). Structural change in baseball's competitive balance: The depression, team location, and integration. *Economic Inquiry, 43*, 158–169.
Lee, Y. H., & Fort, R. (2008). Attendance and the uncertainty-of-outcome hypothesis in baseball. *Review of Industrial Organization, 33*, 281–295.
Lee, Y. H., & Fort, R. (2012). Competitive balance: Time series lessons from the English Premiere League. *Scottish Journal of Political Economy, 59*, 266–282.
Mills, B., & Fort, R. (2014). League level attendance and outcome uncertainty in U.S. pro sports leagues. *Economic Inquiry, 52*, 205–218.
Mills, B. M., & Salaga, S. (2015). Historical time series perspectives on competitive balance in NCAA Division I basketball. *Journal of Sports Economics, 16*, 614–646.
Noll, R. G. (1988). Professional basketball. *Studies in industrial economics paper no. 144*. Stanford, CA: Stanford University.
Pope, D. G., & Pope, J. C. (2008). The impact of college sports success on the quantity and quality of student applications. *Southern Economic Journal, 75*, 750–780.
Quirk, J. (2004). College football conferences and competitive balance. *Managerial and Decision Economics, 25*, 63–75.
Quirk, J., & Quirk, Q. (2012). *All time major college football game scores, 1869–2011*. CDROM.
Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy, 64*, 242–258.
Salaga, S. (2012). Empirical essays in sport management. Dissertation, University of Michigan.
Salaga, S. (2015). *Competitive balance in college football: The GI Bill, grant-in-aid and the College Football Association* (pp. 116–143). XXXIII: Essays in Economic and Business History.
Scully, G. (1989). *The business of Major League Baseball*. Chicago: University of Chicago Press.
Siegfried, J. J., & Burba, M. G. (2004). The College Football Association television broadcast cartel. *The Antitrust Bulletin, 49*, 799–819.

Sloane, P. J. (1976). Sporting equality: Labour market versus product market control: A comment. *Journal of Industrial Relations, 18*, 79.

Smith, D. R. (2008). Big-time college basketball and the advertising effect: Does success really matter? *Journal of Sports Economics, 9*, 387–406.

Sports-Reference.com. (2014). Sports reference/college football, conference index. Accessed October 10, 2014 from http://www.sports-reference.com/cfb/conferences.

Stinson, J. L., & Howard, D. R. (2007). Athletic success and private giving to athletic and academic programs at NCAA institutions. *Journal of Sport Management, 21*, 235–264.

Sutter, D., & Winkler, S. (2003). NCAA scholarship limits and competitive balance in college football. *Journal of Sports Economics, 4*, 3–18.

Tucker, I. B. (2004). A re-examination of the effect of big-time football and basketball success on graduation rates and alumni giving rates. *Economics of Education Review, 23*, 655–661.

Tucker, I. B. (2005). Big-time pigskin success: Is there an advertising effect? *Journal of Sports Economics, 6*, 222–229.

Tucker, I. B., & Amato, L. T. (2006). A reinvestigation of the relationship between big-time basketball success and average SAT scores. *Journal of Sports Economics, 7*, 428–440.