Economic Inquiry

WINNING IN PROFESSIONAL TEAM SPORTS: HISTORICAL MOMENTS

HAYLEY JANG, YOUNG HOON LEE and RODNEY FORT*

Our aims in this paper are to (1) examine the higher moments of the distribution of winning percentages and (2) discover economic implications of such an examination. The results prove useful to both current sports league policy questions and future research. We speculate that the institutional differences between North American pro leagues and European soccer leagues will prove fruitful areas for future research on the determination of competitive balance. (*JEL* C1, L83, Z20)

I. INTRODUCTION

Despite the fact that it is often the focus of analysis in sports economics, there is very little work on the characteristics of the statistical distribution of winning percentage. Basic tests of the form of the distribution are few, and the higher moments of skew and kurtosis have not been examined at all.[1] Our aims in this paper are to (1) examine the higher moments of the distribution of winning percentages and (2) discover economic implications of such an examination. The results are useful to both current sports league policy questions and future research.

For example, the analysis and testing of Rottenberg's "uncertainty of outcome hypothesis" seeks to determine the impact of competitive balance on fan demand (see the surveys in Szymanski 2003; Fort 2006; and Martins and Cro 2018). Similarly, analysis and testing of his "invariance principle" seeks to determine the impacts of either endogenous league action (e.g., imposing a player draft, revenue sharing, payroll cap, or payroll tax) or exogenous factors (e.g., world wars or racial integration) on competitive balance (see the survey in Fort, Maxcy, and Diehl 2016).[2] Quite often, competitive balance is measured by some version of the standard deviation of winning percentage or by the behavior of winning percentage in the tails of its distribution. But, again, nearly nothing is known about the distribution of winning percentage in the first place.

*This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2016S1A2A2912186).

Jang: Research Associate, Center for Research on Innovation and Competition, Sogang University, Seoul 121-742, South Korea. Phone +8210-7456-3389, Fax +822-705-8750, E-mail hayley.85j@gmail.com

Lee: Professor, Department of Economics, Sogang University, Seoul 121-742, South Korea. Phone +822-705-8772, Fax +822-705-8750, E-mail yhnlee@sogang.ac.kr

Fort: Professor, Department of Sport Management, University of Michigan, Ann Arbor, MI 48109-2013. Phone 734-647-8989, Fax 734-936-1925, E-mail rodfort@umich.edu

1. Fort and Quirk (1995), citing Mandelbrot (1963, 1967) and then Fama (1963, 1965) and Fama and Roll (1968a, 1968b, 1971) in the finance literature, tested normality against non-normal stable distributions for winning percentage, but only for MLB (1952–1985), NBA (1975/76–1992/93), and NFL (1930–1941). Lee, Jang, and Hwang (2015) compared the win production efficiency distributions of four European football leagues and found that symmetry of the distribution is different across leagues. Groot (2008) argues for, and applies the Poisson distribution to winning. DiNardo and Winfree (2010) assessed the distribution of MLB home runs.

2. Fort and Maxcy (2003) detail the importance of distinguishing "tracking" analysis of balance, itself, and testing the impact of balance on demand. Ours is definitely a "tracking" exercise.

ABBREVIATIONS

AL: American League
ASD: Actual Standard Deviation
BP: Bai and Perron
EPL: English Premier League
ESLs: European Soccer Leagues
GB: German Bundesliga
ISA: Italian Serie A
ISD: Idealized Standard Deviation
MLB: Major League Baseball
NALs: North American Leagues
NBA: National Basketball Association
NFL: National Football League
NHL: National Hockey League
NL: National League
RSD: Ratio of Standard Deviations
SLL: Spanish La Liga
TV: Television

Economic Inquiry
(ISSN 0095-2583)
Vol. 57, No. 1, January 2019, 103–120

doi:10.1111/ecin.12702
Online Early publication August 6, 2018

We analyze major leagues worldwide. In addition to the well-covered English Premier League (EPL), we extend analysis to three other premier European soccer leagues (ESLs)—the German Bundesliga (GB), Italian Serie A (ISA), and Spanish La Liga (SLL). We also analyze the five major North American leagues (NALs)—the American and National Leagues (AL, NL) in Major League Baseball (MLB), the National Basketball Association (NBA), National Football League (NFL), and the National Hockey League (NHL). This choice facilitates comparison, since past work primarily covers these nine leagues (Fort 2006; Fort, Maxcy, and Diehl 2016; Martins and Cro 2018; Szymanski 2003).

Previewing our results, by-and-large we fail to reject normality against stable non-normal distributions of winning percentages. However, there are enough rejections that we urge testing the distribution of winning percentages as the research area moves forward. All of the rest of our results reveal stark contrasts between NALs and ESLs. One of the results reinforces the finding in Lee and Fort (2012) and Fort and Lee (2013) that imbalance over time has increased in the EPL while it has decreased in MLB. However, we are able to extend that same conclusion to the rest of the EPLs as well.

One of our novel and, perhaps, most important results is that all NALs exhibit negative skew (longer left tail), indicating that competitive imbalance is attributable to the weaker teams in the league. The opposite is true in the ESLs. Positive skew (longer right tail) is evident, so that competitive imbalance is attributable to the strongest teams. In addition, the rejections of normality we do find are decidedly platykurtic (deficient tail observations) in NALs. Fort and Quirk (1995) discovered this as well on their more limited samples of NALs. The opposite is true in ESLs where rejections of normality occur against the leptokurtic alternative (excess tail observations).

The novelty of our skew/kurtosis results led us to speculate on its cause and to note its league policy implications. We suggest the source is the existence of highly unequal television (TV) revenue outcomes and unequal super-competition revenue access for ESLs (Champions League, Europa League, Super Cup). For policy, contrary to current practice in NALs, if a goal is to reduce imbalance, then mechanisms chosen should be designed to enhance the competitiveness of the worst teams in NALs, but to reduce the superiority of the best teams in ESLs.

We also examine the time series behavior of the higher moments in all leagues using what is now referred to in the sports economics literature as the "Bai and Perron (BP) method" (Bai and Perron 1998, 2003, 2006). For a popular measure of the standard deviation of winning percentage, we find results in keeping with previous findings on MLB (Lee and Fort 2005), the NFL (Fort and Lee 2007), and the EPL (Lee and Fort 2012). The new results for soccer leagues mimic the EPL, except for SLL.

In addition to the research methods implications of the existence of break points (Davies, Downward, and Jackson 1995; Dawson and Downward 2005), all of the break points we discover do coincide directly with the episodes of non-normality that we do detect, and there are as many break points for skew and kurtosis together as there are for just the variance. This suggests something deeper going on in these episodes that deserves further attention. Episodes of upheaval, as evidenced by breaks, should prove especially interesting in understanding the determinants of competitive balance. There also is an assortment of differences between NALs and ESLs in the break points for skew and kurtosis.

No paper can do everything so we do not delve into any causal analysis of our findings. But we do offer general speculation about the general differences in ESLs and NALs that might aid future research. NALs operate in closed talent markets, compared with open talent markets for ESLs. Promotion and relegation in ESLs adds a sense of heightened competition to the bottom teams in the first division that may influence the distribution of winning percentages. Three of the four NALs have payroll caps, and MLB has a payroll tax, all likely to impact the stronger teams that typically spend the most on talent. There also are variations in the forms of ownership across the continents. ESLs produce contenders for super competitions (e.g., Champions League) while NALs have only their own within-league championship. Finally, TV revenue sharing arrangements are different across these leagues.

The paper proceeds as follows. In Section II, we describe the sample data, our measures of the moments of the winning percentage distribution, and the results of our interleague comparison of winning percentage distributions. Section III reviews the BP approach and presents our break point analysis of the moments of the winning percentage distributions. Summary discussion of the results of all of our tests is in Section IV. Concluding remarks are in Section V.

II. SAMPLE DATA AND ANALYSIS OF STANDARD DEVIATION, SKEW, AND KURTOSIS

We constructed a panel dataset of winning percentages for our nine different professional sports leagues from popular sources. A measure of each of the higher moments of the winning percentage distribution was calculated for analysis.[3] Tests of normality against non-normal stable alternatives were performed and the behavior of the standard deviation, skew, and kurtosis measures were documented. Formal time series analysis of these three measures of the moments of the winning percentage distribution is in the next section.

The ratio of standard deviations (RSD), founded by Noll (1988) and Scully (1989), and more rigorously established in Quirk and Fort (1992) and Fort and Quirk (1995) is the most popular measure of dispersion in the literature. RSD can be visualized as follows. Let ASD be the actual sample standard deviation. Let ISD be the "idealized" standard deviation of a league where the probability that any team in the league beats any other is 0.5. For the binomial distribution, without ties, Fort and Quirk (1995) show that $\mathrm{ISD} = \frac{0.5}{\sqrt{G}}$, where G is the number of games in the season.[4] Then RSD = ASD/ISD. As $\mathrm{RSD} \to 1$, the league is as balanced as the "idealized" league with equal win probability in all games. The larger is $\mathrm{RSD} > 1$, the more imbalanced the league becomes relative to the "idealized" league.[5]

There is growing dissatisfaction with using RSD to compare across leagues, or even for a given league when season length changes (see Lee, Kim, and Kim 2016; Owen and King 2015). So, we present our RSD findings below solely for what they can tell us about a given league, and for comparison purposes with past work that relied on this measure.[6]

The third moment, skew, measures asymmetry of the distribution. Negative skew (longer left tail) would suggest imbalance is primarily attributable to weaker teams. Positive skew (longer right tail) would suggest imbalance is primarily attributable to stronger teams. We used STATA to calculate skew. Let $m_r = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^r$, where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, n = number of observations, and r = chosen moment. Skew is calculated as $\frac{m_3}{m_2^{3/2}}$.

Kurtosis is the fourth moment of the distribution, equal to 3 for the normal distribution. It measures whether the distribution has either an excess of observations in the tail (values greater than 3 are "leptokurtic") or a deficiency of observations in the tail (values less than 3 are "platykurtic") relative to the normal distribution. Coupled with skewness, kurtosis adds to the detection of imbalance. For example, negative skew attributes the imbalance to weaker teams and if there were also excess tail observations (leptokurtic), it would be an excess of weak teams. The behavior in the tails has been used in past analysis of competitive balance, primarily in terms of "tail likelihood" (Fort and Quirk 1995) or excess (positive or negative) tail frequency (Lee 2004).[7] While it has yet to see any use to this end, kurtosis is a likely addition to this list of competitive balance variables since it measures the tails of the distribution directly. We also used STATA to calculate kurtosis as $\frac{m_4}{m_2^2}$.

We check the normality of the nine league winning percentage distributions. There are various methods to test for normality (Anscombe and Glynn 1983; D'Agostino 1970; Doornik and Hansen 2008; Jarque and Bera 1987; Kolmogorov 1956; Shapiro and Wilk 1965). We attempt to figure out the cause of non-normality if a normal distribution is rejected. Therefore, we conduct a test of skewness and a test of kurtosis separately instead of using omnibus tests that detect deviations from normality due to either skewness or kurtosis. In tests of skewness, the null hypothesis is normality and the alternative is non-normality due to skewness.

We start with the skewness test in D'Agostino (1970) which presents a test statistic that is approximately normally distributed under the

3. Admittedly, sample sizes in any given year for any given league might be small (around eight teams in some years). This does reduce the power of our chosen tests but there is at least the virtue that all leagues run at least 52 years (GB), two as high as 114 years (AL and NL), and one 116 years (EPL).

4. For the case of ties, see Cain and Haddock (2006), Fort (2007), and Owen (2012).

5. There actually is quite a lively debate over measures of variation and concentration for competitive balance. On the first, see Eckard (2001), Humphreys (2002), Lee and Fort (2005), Krautmann and Hadley (2006), Cain and Haddock (2006), Fort (2007), Owen (2010, 2012), and Owen and King (2015). On the latter, see Depken (1999), Schmidt and Berri (2001), Utt and Fort (2002); and Owen, Ryan, and Weatherston (2007).

6. Lee, Kim, and Kim (2016) demonstrate that ASD is also biased for competitive balance so we do not provide any further investigation of the actual standard deviation.

7. A review of these measures and comparison in some uses is in Mills and Fort (2014).

TABLE 1
Skewness Descriptive Statistics and Normality

League	Total Seasons	Ave	*SD*	Total Rejections	Percentage	Positive	Negative
AL	114	−0.14	0.48	7	6.1	3	4
NL	114	−0.11	0.47	6	5.3	0	6
NBA	69	−0.15	0.35	2	2.9	1	1
NFL	94	−0.08	0.26	0	0.0	0	0
NHL	97	−0.2	0.48	9	9.3[a]	1	8
EPL	116	0.31	0.46	20	17.2	17	3
GB	52	0.28	0.46	7	13.5	6	1
ISA	83	0.33	0.47	16	19.3	15	1
SLL	84	0.33	0.62	19	22.6	16	3

Note: The skewness test is in D’Agostino (1970).
[a]Test results are based on 5% significance level. Skewness is not calculated for seasons where there are fewer than eight teams. Therefore, for NHL 1917–1925 and 1938–1966 periods were excluded in this table.

null. We test the normality of each season’s win distribution. For example, the AL sample is 114 seasons so we repeat the normality test 114 times. The results are summarized in Table 1.[8] By-and-large, the test fails to reject normality but caution on this dimension is suggested. First, there are rejections in every league except the NFL and from 2.9% of the seasons in the NBA up to 22.6% in the SLL. Second, there is a much higher number of rejections, as well as a higher percentage of rejections, across the ESLs than across the NALs. Finally, and as an object of discussion in the next section, all of the average skew values in NALs are negative, while the opposite is true for the ESLs.

We also check normality with the kurtosis test in Anscombe and Glynn (1983) and the well-known Studentized range kurtosis test.[9] The results are in Table 2. By-and-large, the tests fail to reject normality but, again, we urge caution. First, there are rejections in every league, from less than 6% in the GB to 22.7% in the NHL. Second, when rejections occur, they appear to be decidedly platykurtic (deficiency of tail observations) in NALs but leptokurtic (excess tail observations) in ESLs.

Examining our three measures for their own characteristics, rather than as elements of normality tests, as we note above, whether RSD is suitable for such comparisons is in question. For that reason, we display the RSD results for insights about each league singly and for comparison with past works. For example, Table 3 shows the descriptive statistics and Figure 1 charts RSD behavior over time for each league. While both MLB leagues appear to show improved balance (declining RSD), it is difficult to see any trend in imbalance in the NFL or NHL. There also appears to be an upward trend in each of the ESLs.

Moving on to skew, as we already noted, Table 1 reveals that all NALs have negative skew while all ESLs have positive skew. No particular trend jumps out for NALs from Figure 2 while there may be a bit of an upward trend in ESLs. Figure 2 also shows the negative skew in the NALs and the positive skew in the ESLs. Kernel density estimates in Figure 3 make the differences in skew, NALs versus ESLs, visually quite clear.

That leaves a few remaining observations on kurtosis. Table 2 shows all kurtosis values at 3.02 (SLL) or less. Table 2 also shows that kurtosis

8. All of the tests in Tables 1 and 2 are one-sided tests. We also conduct several two-sided tests (Doornik and Hansen 2008; Jarque and Bera 1987; and Shapiro and Wilk 1965). Both Shapiro and Wilk (1965) and Doornik and Hansen (2008) produced similar results, while Jarque and Bera (1987) produced a smaller number of rejections. Nonrejection of normality may be reflection of the low power of the Jarque and Bera test in small samples (Thadewald and Büning 2007). Note that our samples are generally small ranging from 3 to 30. On the other hand, the Doornik and Hansen test is an omnibus test that controls well for size even with small samples according to their simulation results. Therefore, we consider the results of two-sided tests to support our results in Tables 1 and 2. We appreciate a referee who suggested attempting various normality tests for robustness.

9. If the random sample $x_i \sim N(0, 1)$, i = 1,..,*n*, and another random variable *s* with $\upsilon s^2 \sim \chi^2$, independent of the x_i, then $\max\left[\frac{x_i - x_j}{s}\right]$ has the Studentized range distribution with υ degrees of freedom. Tables of the critical values for the Studentized range statistic originally are in May (1952), but easily found using any browser. Fama and Roll (1971) show that the Studentized range test had higher power than a broad array of distribution-free goodness-of-fit statistics for testing normality against non-normal stable alternatives (e.g., Kolmogorov–Smirnov, chi-squared) and generally outperformed the Shapiro–Wilk test which was designed specifically for a normal null hypothesis.

TABLE 2
Kurtosis Descriptive Statistics and Normality

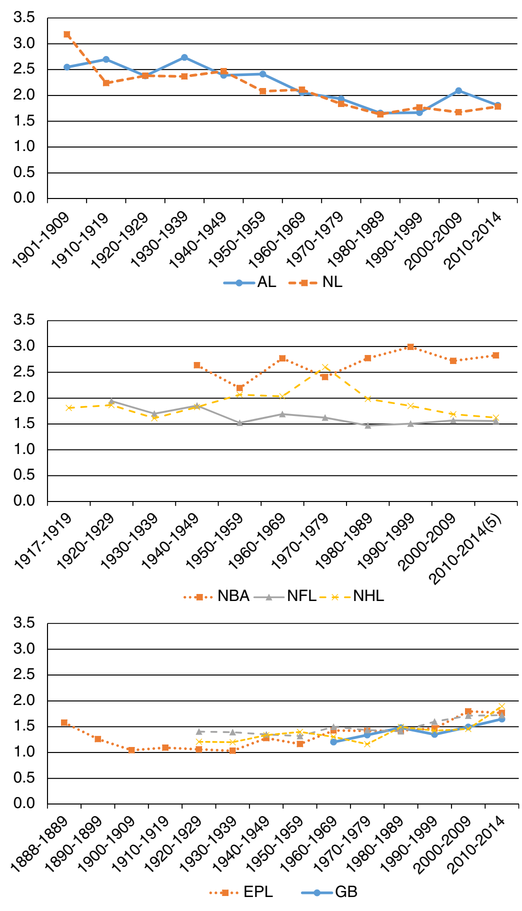
				Anscombe and Glynn (1983)				Studentized Range			
League	Total Seasons	Ave	*SD*	Total Rejections	Percentage	Leptokurtic	Platykurtic	Total Rejections	Percentage	Leptokurtic	Platykurtic
AL	114	2.28	0.66	9	7.9	2	7	13	11.4	4	9
NL	114	2.20	0.52	6	5.3	1	5	11	9.6	2	9
NBA	69	2.28	0.4	10	14.5	1	9	3	4.3	1	2
NFL	94	2.21	0.4	11	11.7	0	11	11	11.7	1	10
NHL	97	2.55	0.75	4	4.1[a]	2	2	22	22.7	6	16
EPL	116	2.92	0.81	18	15.5	12	6	17	14.7	12	5
GB	52	2.73	0.62	3	5.8	2	1	2	3.8	0	2
ISA	83	2.86	0.75	6	7.2	5	1	8	9.6	6	2
SLL	84	3.02	1.06	16	19.0	14	2	13	15.5	10	3

[a]Test results are based on 5% significance level. Kurtosis test of Anscombe and Glynn (1983) is not calculated for seasons where there are fewer than five teams. The period of 1917–1923 for NHL was not included in the Anscombe and Glynn (1983) test.

TABLE 3
RSD Descriptive Statistics

League	Years	Mean	*SD*
AL	1901–2014	2.22	0.56
NL	1901–2014	2.14	0.6
NBA	1946–2014	2.72	0.44
NFL	1922–2015	1.6	0.21
NHL	1917–2014	1.91	0.44
EPL	1888–2014	1.33	0.32
GB	1963–2014	1.41	0.25
ISA	1929–2014	1.48	0.23
SLL	1928–2014	1.38	0.25

FIGURE 1
Decade Average RSD



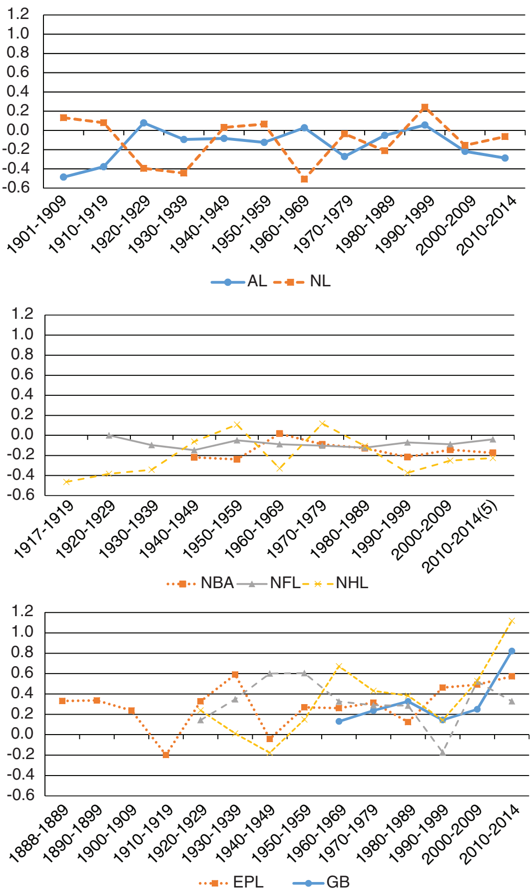
values are consistent across the NALs, and consistently higher across the ESLs than across the NALs. Finally, Figure 4 shows no particular temporal trend for either the AL or NL, and an ever so slight possible upward trend in the rest of the NALs. Kurtosis appears to have declined over most of the history of European soccer but with increases in the last 15 years in all but the EPL.[10] We list our insights and draw our conclusions after the following analysis of break points.

III. BREAK POINT ANALYSIS

Bai and Perron (1998) developed a comprehensive analysis and tests for break points in time

10. We also performed *t*-tests, pairwise, on each of the ESLs one against the other and were unable to detect any ranking in terms of differences at the mean. All we could detect is that both skew and kurtosis were different at the mean for SLL versus GB.

FIGURE 2
Decade Average Skew



series using a multiple structural change model.[11] They considered the following multiple regression with m breaks ($m+1$ regimes):

$$(1)\quad y_t = x_t'\beta + z_t'\delta_j + u_t,\quad t = T_{j-1},\ \dots\ ,T_j - 1;$$
$$j = 1,\ \dots\ ,m+1,$$

where y_t is the observed dependent variable at time t—RSD, skew, and kurtosis of winning percentages.[12] x_t ($p \times 1$) and z_t ($q \times 1$) are vectors of covariates, and β and δ_j are corresponding vectors of coefficients. u_t is the random component at time t. The break points ($T_1,\ \dots\ ,T_m$) are treated as unknown, and this model estimates coefficients of covariates with break points together. When β is assumed not to change as expressed as Equation (1), the model is a "partial structural change model."

We assume β and δ_j are allowed to change and employ the "pure structural change model":

$$(2)\quad y_t = z_t'\delta_j + u_t,\quad t = T_{j-1},\ \dots\ ,T_j - 1;$$
$$j = 1,\ \dots\ ,m+1.$$

If the model finds one break, say $t = T_1$, then the sample is partitioned into two segments, one starting time to $T_1 - 1$ and the second segment is from T_1 to the end of the sample. When unknown breaks are examined, an ordinary least squares regression with dummy variables for significant break point year(s) is estimated. The results yield the direction and statistical significance of structural change. We use a constant term only or constant and trend terms together as z_t and set a trimming parameter to be 0.10 or 0.15. Then we select estimation results based on the adjusted R^2.

In the BP Method, four different statistical tests help identify the existence and number of break points. The first, denoted *Sup* $F_T(k)$, is a generalization of the test in Andrews (1993) for a single break. The null of no structural break ($m=0$) is tested against the alternative of $m=k$ breaks. The second and third tests are double maximum tests that consider the null of no structural break against an unknown number of breaks given a chosen upper bound M. These tests are similar to *Sup* $F_T(k)$ but the tests are divided into *UDmax* and *WDmax* depending on some fixed weights, unity, and marginal p values, respectively. The three tests so far do not provide information on the *number of breaks* and the final sequential test, denoted *SupF((l+1)/l)*, considers l number of breaks versus the alternative hypothesis of $(l+1)$ breaks. Fortunately for us, Perron has made his Gauss program for these tests freely available.[13]

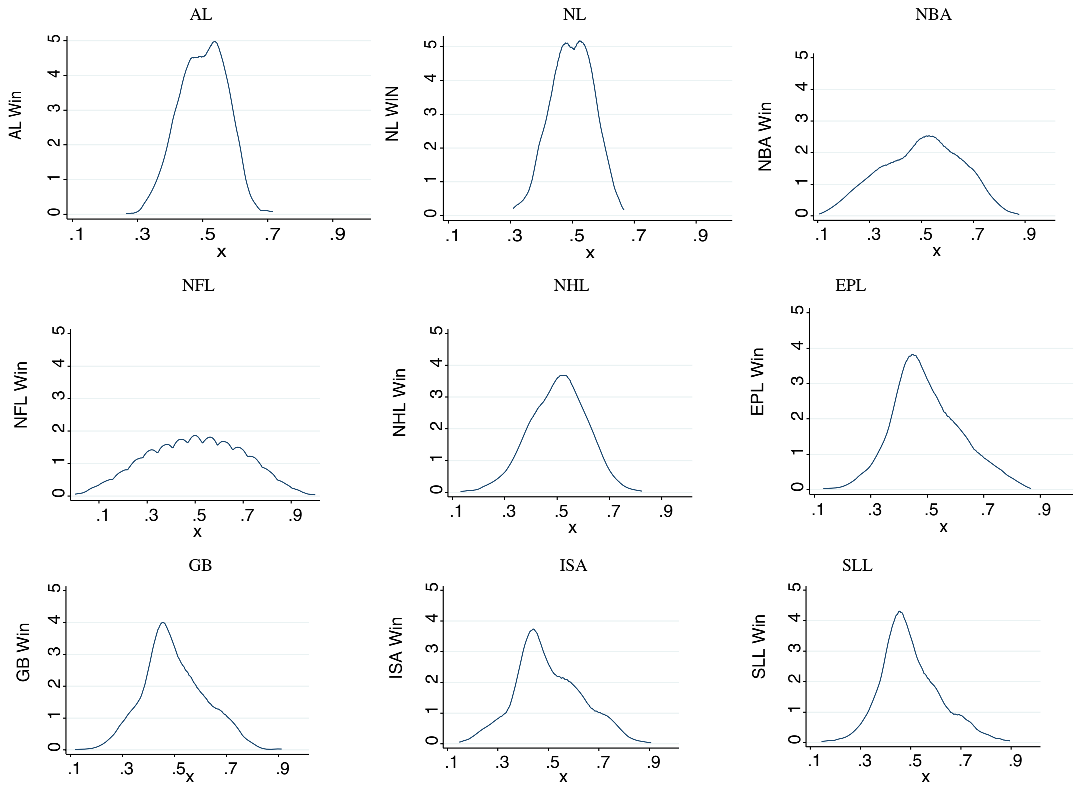
The results of our break point tests for RSD, skew, and kurtosis are in Tables 4, 5, and 6, respectively. Table 7 summarizes our final determination of the break points and their confidence

11. Stationarity is occasionally an issue in time series analysis of sports *attendance* but never has been an issue in the time series assessment of *competitive balance* in any of the major North American leagues or in the Premier League. We forego formal testing for unit roots.

12. Since we do not intend to compare between leagues, it is allowable to include RSD.

13. We also present the Bayesian information criterion (BIC) (Liu, Wu, and Zidek 1997) and the Schwarz criterion (LWZ) (Yao 1988). While these statistics provide further information on possible structural changes, we draw our inferences from the sequential test results. Bai and Perron (2006) demonstrated superiority of the sequential test over BIC and LWZ using Monte Carlo techniques.

FIGURE 3
Kernel Density Estimates, 1963–2014



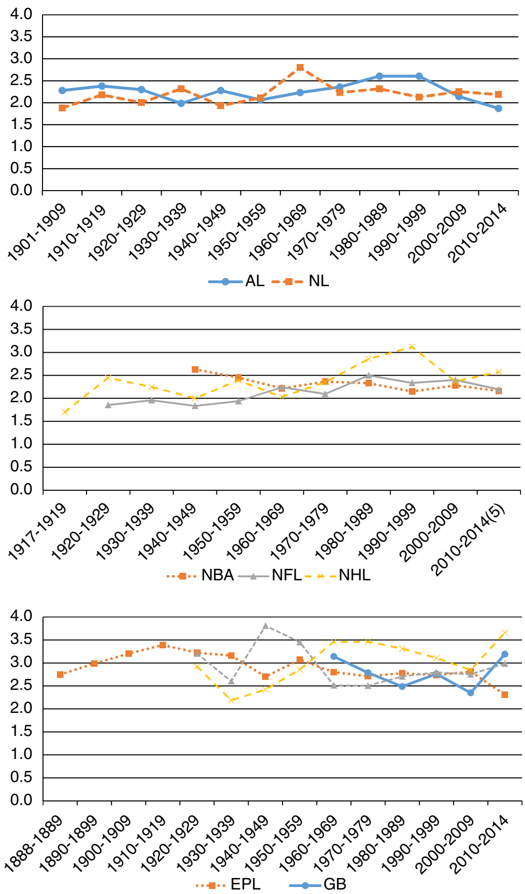
intervals. The confidence intervals are often very wide, for example, more than a decade for all the kurtosis results and several of the RSD results in Table 7. This limitation is noted again in the next section when we derive conclusions from the analysis. It is possible to discuss the results of the determination of statistical significance, direction, and possible trends directly from the regression results in Table 8. However, our experience suggests that it is easier to see these results from a plot of actual and fitted time series. Those plots are in Figures 5 (RSD), 6 (skew), and 7 (kurtosis).

Figure 5 shows breaks were detected in three of the five NAL RSD series and in all of the ESL RSD series. Structural breaks for the AL and NL, and the NFL, are generally consistent with the earlier results in Lee and Fort (2005) and generally favor improved competitive balance as measured by RSD. In addition, the breaks for the EPL are consistent with the earlier results in Lee and Fort (2012), with balance worsening each time. This is true even though the current results extend the time period on these leagues over previous studies.

More interesting are the new RSD break point results for the other European leagues. We use the following convention for discussing breaks. The pair (break, trend) denotes the direction of the break and the direction of the following trend, if any. Thus, (+,−) denotes a shift upward followed by a negative trend, (+,+) a shift upward followed by a positive trend, and (+,0) a shift upward followed by no trend. Similarly, for any negative shift.

For the GB a (+,0) break in RSD is detected around 2001 following a complete period of no trend at all prior to that. ISA has a much earlier (+,+) break in 1941 reversing a complete downward trend prior to that. SLL was the most volatile of the three and the break directions were opposite of the other two leagues. Following an extended earliest period of increasing trend, a (−,+) break occurred about 1971, followed by another (−,+) break in 1999.

FIGURE 4
Decade Average Kurtosis



As shown in Figure 6, breaks in skew were only detected for the NL in NALs, and for all but the EPL in ESLs. There was a (+,0) break in skew for the NL (but not the AL) in 1942, following a long decline. The same (+,0) break occurred for the GB in 2010, but there was no trend detected prior to that. Declining trends surrounded the (+,−) break for ISA in 2004 but the SLL was just the opposite; an increasing trend followed by a (−,+) break in 1994.

Finally, breaks in kurtosis were found three times in NALS—the NL, NFL, and NHL—and twice in the ESLs—the ISA and SLL. Following an extended period prior without any trend at all, there was a (+,0) break for the NL in 1961 and a (−,0) break in 1976. Essentially, these two breaks cancel each other out with kurtosis equal to 2.0 prior to the first break and returning to 2.0 after the second. The NFL also showed a (+,0) break in 1978 without any trend prior to that. Both of the breaks in ESLs occurred much earlier, a (−,+) break in 1958 for ISA, following an increasing trend overall prior to that. Once again, the SLL was contrary with a (+,0) break in 1953 and no trend at all prior to that.

IV. DISCUSSION

Our analysis yields five main insights—caution concerning the predominance of normality, improved balance in NALs but worsened balance in ESLs, asymmetric skew in ESLs (positive skew) versus NALs (negative skew), greater kurtosis in ESLs compared to NALs, and that break points have occurred in at least one of the higher moments of the winning percent distribution of all leagues except the NBA. We address these seriatim and offer our conclusions on the consistent asymmetry in our results between ESLs and NALs.

A. *Normality*

The upshot of our normality test results is that future research would do well to exercise caution about any assumption that winning percent is distributed normal. Typically, winning percentages have been normally distributed. However, when deviations from normality do occur, there are a variety of causes across different leagues. Normality violations have been due to both skew and kurtosis in all leagues except the NFL. In that league, normality violations have only occurred relative to kurtosis.

These results suggest that past work that relied on normality may need to be revisited. More importantly, future research would do well to test for normality of the winning percentage distribution. Both sound method and believable results depend on it.

B. *RSD: Asymmetric Behavior of Imbalance, NALs versus EPLs*

We are unable to compare imbalance across leagues due to the previously noted limits of the RSD measure. However, there is a stark contrast in terms of the behavior of this measure of imbalance, league-by-league. By Figure 1, balance has improved generally in each NAL but worsened in each ESL. This is not a new result but reinforces past findings on imbalance in past works on MLB and the EPL (Fort and Lee 2013; Lee and Fort 2012), as well as extending those findings to the comparison of NALs and ESLs, generally.

TABLE 4
BP Method Results for RSD

	Specifications										
	$z_t = \{1\}$ **or** $\{1, \text{time}\}$		$x_t = \{0\}$	$q = 1$ **or 2**	$p = 0$		$\epsilon = 0.10$ **or 0.15**	**h = 10**		**M = 5 or 6**	
	Tests										
	Sup $F_T(1)$	***Sup*** $F_T(2)$	***Sup*** $F_T(3)$	***Sup*** $F_T(4)$	***Sup*** $F_T(5)$	**UDmax**	**WDmax**	***SupF*** **(2/1)**	***SupF*** **(3/2)**	***SupF*** **(4/3)**	***SupF*** **(5/4)**
AL	76.85*	41.57*	40.99*	32.98*	27.28*	76.86*	76.86*	3.58	3.19	2.18	2.26
NL	62.92*	67.78*	40.77*	38.68*	32.16*	67.78*	86.89*	28.32*	2.77	6.63	2.50
NBA	8.09	6.95	8.02*	9.60*	5.12*	9.60**	19.06*	4.36	11.58**	6.80	0.00
NFL	27.32*	21.21*	15.81*	12.46*	8.82*	27.32*	27.85*	4.73	6.22	6.22	0.00
NHL	9.26**	39.33*	28.24*	23.36*	18.48*	39.33*	51.64*	25.32*	6.48	1.43	0.16
EPL	70.38*	65.67*	50.78*	39.11*	33.42*	70.38*	84.19*	33.78*	11.40**	3.42	1.92
GB	13.58*	8.13**	7.23**	5.49	6.31**	13.58*	16.41*	2.11	1.84	2.11	3.92
ISA	19.14*	15.39*	16.41*	17.74*	14.14*	19.14*	29.10*	11.68	13.75	3.89	3.89
SLL	105.31*	59.99*	41.51*	33.79*	29.11*	105.31*	105.31*	16.85*	14.78**	7.95	3.37

	Number of Breaks Selected								
	AL	**NL**	**NBA**	**NFL**	**NHL**	**EPL**	**GB**	**ISA**	**SLL**
Sequential[a]	1	2	0	1	0	2	1	1	2
BIC	1	2	1	1	2	3	1	1	2
LWZ	1	2	0	1	2	2	0	0	1

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

C. Skew and Kurtosis: Asymmetric Sources of Imbalance, NALs versus EPLs

Table 1 and Figure 2 make two things clear. First, ESLs exhibit exclusively positive skew (longer right tail) and this occurs for all leagues for some seasons. Second, skew in NALs, when it is detected, is more mixed. The AL and NBA exhibit both types without any predominance. The NL and NHL exhibit predominantly negative skew (longer left tail). No skew at all is ever detected for the NFL. For skew, the clear asymmetry is in the ESL entirely positive skew result compared with the more mixed NAL results.

Further, Table 2 and Figure 3 add to the observation of asymmetry between NALs and ESLs. When rejections of normality do occur, NALs are characteristically platykurtic (deficient observations in the tails), albeit by only the Anscombe and Glynn test for the NBA and only the Studentized range test for the NHL. On the other hand, ESLs are characteristically leptokurtic (excess observations in the tails), except for the GB which has the fewest normality violations via kurtosis of all leagues studied.

Thus, in ESLs, the source of imbalance is the stronger teams (positive skew) and that there are probably more of them relative to the case of a normal distribution of winning percentage (excess observations in the tail). In NALs, the source of imbalance where skew is determinate is the weaker teams (negative skew in the AL, NL, and NHL) and that there are probably too few stronger teams, relative to the case of a normal distribution of winning percentage. Skew and kurtosis for the remaining NALs do not tell such a consistent story (indeterminate for skew) although normality violations, where they do occur for the AL, NBA, and NFL, all exhibit too few tail observations relative to the normal distribution.

To our knowledge, this is a novel result and, from both the research methods and the policy perspective, possibly the most important result in the paper. From the research methods approach, the observation is straightforward. Future research would do well to account for the type of normality rejection. Think about it like the income distribution—for NALs, it is the lower, thinner tail that is the source of imbalance; for ESLs, it is the upper thicker tail that is the source of imbalance.

From the policy perspective, if less imbalance is the goal, it is clear that the policy tools (1) will be more effective if they are chosen with implications of our skew/kurtosis results in mind and (2) must be different for ESLs than for NALs. In NALs, payroll caps (in all NALs except the AL and NL in MLB) and payroll

TABLE 5
BP Method Results for Skew

	Specifications										
	$z_t = \{1\}$ or $\{1, \text{time}\}$		$x_t = \{0\}$	$q = 1$ or 2		$p = 0$	$\varepsilon = 0.10$ or 0.15			h = 10	M = 5 or 6
	Tests										
	Sup $F_T(1)$	*Sup* $F_T(2)$	*Sup* $F_T(3)$	*Sup* $F_T(4)$	*Sup* $F_T(5)$	UDmax	WDmax	*Sup* $F(2/1)$	*SupF* (3/2)	*SupF* (4/3)	*SupF* (5/4)
AL	11.23	7.83	6.68	6.01	4.97	11.23**	11.23**	7.01	3.68	2.06	1.61
NL	14.69**	13.55*	17.65*	15.82*	13.20*	17.64*	25.96*	12.41**	9.38	9.58	3.06
NBA	1.64	3.65	4.66	3.63	2.92	4.66	7.54	6.00	1.51	1.07	0.16
NFL	1.46	1.28	2.22	2.79	1.81	2.79	5.54	1.46	3.76	2.85	0.00
NHL	8.74**	8.57**	7.07**	5.79**	4.08**	8.74	10.18**	8.50	1.81	3.72	0.00
EPL	4.30	9.12	12.65**	8.07	7.54	12.65**	16.68**	9.55	4.44	5.63	4.23
GB	40.26*	21.30*	17.08*	14.89*	10.56*	40.26*	40.26*	2.66	13.46	5.38	1.19
ISA	26.28*	20.44*	38.07*	36.32*	29.85*	38.07*	59.58*	12.37**	13.03	21.46*	4.56
SLL	14.28**	14.21*	12.08*	10.45*	9.14*	14.28**	17.83*	13.23**	5.55	3.78	4.64

Number of Breaks Selected

	AL	NL	NBA	NFL	NHL	EPL	GB	ISA	SLL
Sequential[a]	0	1	0	0	0	0	1	1	1
BIC	0	0	0	0	0	0	1	0	0
LWZ	0	0	0	0	0	0	0	0	0

Note: See Table 4.
[a] 1% significance level for the sequential test.
*Significant at the 99% critical level. **Significant at the 95% critical level.

TABLE 6
BP Method Results for Kurtosis

	Specifications										
	$z_t = \{1\}$ or $\{1, \text{time}\}$		$x_t = \{0\}$		$q = 1$ or 2	$p = 0$	$\varepsilon = 0.10$ or 0.15		h = 10	M = 5 or 6	
	Tests										
	Sup $F_T(1)$	*Sup* $F_T(2)$	*Sup* $F_T(3)$	*Sup* $F_T(4)$	*Sup* $F_T(5)$	UDmax	WDmax	*SupF* (2/1)	*SupF* (3/2)	*SupF* (4/3)	*SupF* (5/4)
AL	8.94	10.98**	8.86	7.35	6.19	10.98	12.71	7.20	4.58	2.99	3.30
NL	7.80	12.92*	9.27*	8.14*	6.65*	12.93**	16.57*	16.96*	1.40	4.21	1.00
NBA	3.80	2.78	3.45	3.35	2.02	3.80	6.64	1.43	3.17	3.04	0.00
NFL	30.43*	18.06*	12.25*	9.98*	6.92*	30.42*	30.42*	4.55	1.38	0.43	0.00
NHL	13.03*	8.04**	6.13**	5.01**	4.07**	13.03*	13.03**	5.76	5.22	1.01	0.15
EPL	7.85	4.39	3.76	3.84	3.67	7.84	7.84	2.37	1.68	3.62	1.31
GB	4.25	5.36	4.09	3.63	3.35	5.36	6.29	6.51	2.39	2.39	1.90
ISA	14.61**	24.10*	21.10*	20.88*	16.88*	24.10*	34.24*	11.74	14.15**	12.72	3.18
SLL	20.61*	10.30*	7.73**	5.94	5.71**	20.61*	20.61*	0.79	4.89	1.66	3.36

Number of Breaks Selected

	AL	NL	NBA	NFL	NHL	EPL	GB	ISA	SLL
Sequential[a]	0	2	0	1	1	0	0	1	1
BIC	0	2	1	1	2	0	0	1	1
LWZ	0	0	0	1	1	0	0	0	0

Note: See Table 4.
[a] 1% significance level for the sequential test.
*Significant at the 99% critical level. **Significant at the 95% critical level.

TABLE 7
Break Point Estimates and Their Confidence Interval Summary

	RSD	Skew	Kurtosis
AL	1958 (1955, 1962)	—	—
NL	1913 (1909, 1915) 1966 (1963, 1973)	1942 (1940, 1962)	1961 (1956, 1966) 1976 (1970, 1981)
NBA	—	—	—
NFL	1952 (1948, 1960)	—	1978 (1971, 1984)
NHL	—	—	1988 (1972, 1991)
EPL	1956 (1951, 1962) 2001 (2000, 2004)	—	—
GB	2001 (1996, 2012)	2010 (2008, 2014)	—
ISA	1941 (1939, 1946)	2004 (2002, 2005)	1958 (1956, 1968)
SLL	1971 (1961, 1972) 1999 (1997, 2000)	1994 (1993, 1995)	1953 (1938, 1956)

Note: Break points are statistically significant at the 90% critical level using the tests listed in the text.

TABLE 8
Estimation Results

	Constant	Trend	First Break	First Break·Trend	Second Break	Second Break·Trend
RSD						
AL	2.57* (35.07)		−0.71* (−7.47)			
NL	3.14* (23.67)		−0.84* (−5.49)		−0.56* (−6.25)	
NFL	1.82* (36.26)		−0.27* (−4.64)			
EPL	1.13* (27.14)		0.30* (5.77)		0.38* (7.41)	
GB	1.35* (34.07)		0.22* (3.98)			
ISA	4.03* (8.82)	−0.57* (−5.87)	−3.04* (−6.33)	0.62* (6.36)		
SLL	1.05* (7.59)	0.04** (1.99)	−1.54* (−3.52)	0.15* (3.12)	−5.12* (−8.77)	0.41* (7.64)
Skew						
NL	0.53* (4.57)	−0.20* (−6.84)	−0.64** (−1.98)	0.21* (4.87)		
GB	0.22* (3.44)		0.60* (7.48)			
ISA	0.83* (3.16)	−0.07** (−2.13)	8.73* (4.22)	−0.67* (−3.90)		
SLL	−0.54 (−1.91)	0.11* (2.99)	−7.56* (−3.20)	0.62* (3.05)		
Kurtosis						
NL	2.07* (32.09)		0.67* (6.93)		−0.56* (−5.40)	
NFL	1.97* (44.12)		0.42* (5.71)			
NHL	2.20* (32.25)		0.70* (3.03)			
ISA	0.52 (0.59)	0.48* (2.84)	1.59 (1.63)	−0.42** (−2.42)		
SLL	2.40* (22.97)		0.84* (4.83)			

Notes: t-statistics are in parenthesis. Newey-West standard errors are used.
*Significant at the 99% critical level; **Significant at the 95% critical level.

taxes (explicitly, just in the AL and NL in MLB) are in place, ostensibly to improve competitive balance.[14] Indeed, the payroll tax in MLB is called the "competitive balance" tax. Neither of these devices is in place in ESLs.

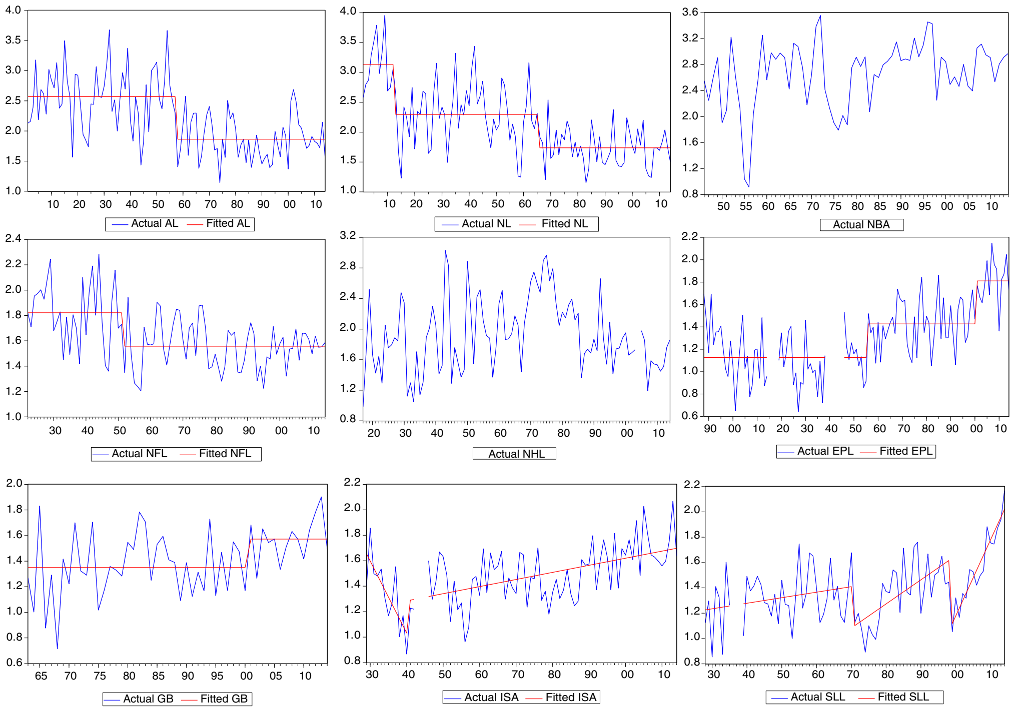
Generally, theory suggests that league policies like payroll caps and payroll taxes have the ability to improve competitive balance. They do so by altering the talent spending decisions of the *strongest* teams in a league (textbook examples are in Fort 2011, Chapter 6). But our skew and kurtosis results suggest that imbalance is attributable to the *weaker* teams in NALs. This suggests that NALs consider how to improve winning performance by the *worst teams*, rather than reduce performance by the top teams. For example, our results suggest that any proceeds from revenue sharing must be distributed in an incentive-compatible way that leads the worst teams to actually spend the money on talent.[15] Any other direct, incentive-compatible transfers to the worst teams will also be aimed correctly according to our results.

On the other hand, our skew and kurtosis results suggest that payroll caps and/or payroll

14. The remaining NALs have what are referred to as "luxury taxes" but they actually are reserve withholdings ("escrow") from player pay that facilitate unexpected transfers from players to owners in the event that payroll caps are violated.

15. We leave aside that the impact of revenue sharing on competitive imbalance is at best situation specific (Winfree and Fort 2012).

FIGURE 5
Fitted Value of RSD



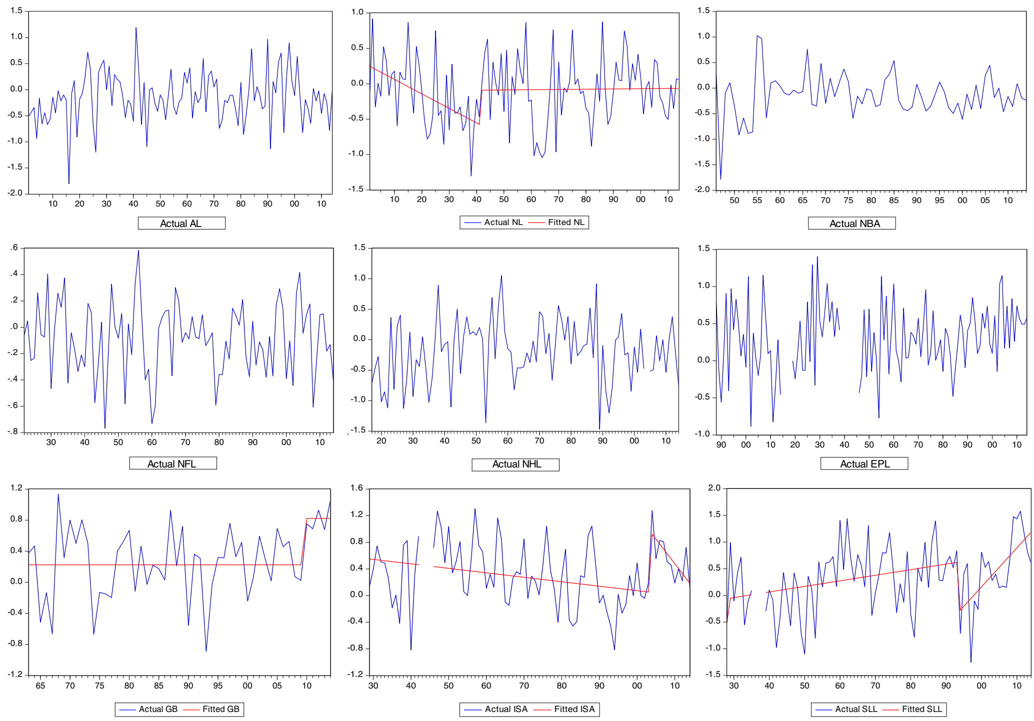
taxes would correctly target the source of imbalance in ESLs, namely, the teams with the top performance. With positive skew and excess frequency in the top tail, mechanisms aimed at the stronger teams would be expected to be effective at reducing competitive imbalance. For example, our results are consistent with the imposition of "financial fair play," at least correctly intended as it strives to put a cap on player spending. However, our results do suggest that caps aimed disproportionately at stronger teams, such as those employed currently in NALs, would prove more effectual.

Because our skew/kurtosis asymmetry result is novel, we also feel obliged to offer our best speculation as to its cause, as an aid to future research. The most obvious to us is the asymmetry of revenue potential in ESLs, compared with what occurs in NALs. This asymmetry is apparent in both TV revenue distribution and for access to lucrative super competitions.

For TV revenue distribution, while equal sharing will aid no particular team, unequal sharing may do so. As an example, SLL adopted its new TV contract system in 1998 that allowed each individual club to sell its media rights independently (Ascari and Gagnepain 2006). This change caused dramatic TV revenue disparity since only two teams, Real Madrid and FC Barcelona, obtained lucrative contracts. They each earned €140 million in the 2014/2015 season whereas Deportivo la Coruna earned only €17.5 million. This change in the SLL matches up quite nicely with the outcome that positive skew would attribute competitive imbalance to the performance of the stronger clubs in ESLs.

Revenues are skewed to the strongest teams in individual leagues because they are the ones that eventually play in lucrative super competitions. If the best teams are continually the best teams, the aggregate impact of recurrent access to the added revenues from super competitions is a likely candidate for the type of imbalance characteristics we have found. Four teams from each ESL go to the UEFA Champions League. Payoff is determined by progress through the rounds of the tournament and shares of qualifying pools and the TV market pool. UEFA Europa League typically includes three teams from each ESL. The

FIGURE 6
Fitted Value of Skew



two champions move on to UEFA Super Cup. In addition, the UEFA Champions League winner earns a bid to FIFA *Club* World Cup (not to be confused with the UEFA World Cup of national teams).

While just being chosen is worth a few hundred thousand Euros, winning can be worth just under €10 million in the Europa Cup and up to tens of millions in the Champions League. And the additional just under €5 million in the Super Cup goes to just two teams. FIFA Club World Cup winners also can make another €5 million. Concentrated entry and success in these super competitions is consistent with our skew/kurtosis findings for ESLs, especially that kurtosis appears greater in ESLs than in NALs, as well as exhibiting increases over the last 15 years in ESLs (except in the EPL).

D. Break Points: Asymmetric Changes in Imbalance, NALs versus EPLs

Our break point results generate both research method and research question suggestions. On research methods, we see four suggestions. First, since we do find break points, the cautions in Davies, Downward, and Jackson (1995) and Dawson and Downward (2005) are in play. They make it quite clear that estimation using level data that span break points can generate spurious correlations.

Second, and related, using the confidence intervals in Table 7, Tables 10 and 9 then show that all four of our skew break points and all six of our kurtosis break points coincide with episodes of normality rejection. Further, these relationships all follow the same skew pattern (negative for NALs and positive for EPLs) and the same kurtosis pattern (platykurtic for NALs and leptokurtic for ESLs). While we are reminded of our own earlier observation in the last section that most of the break point confidence intervals are wide, as future researchers deal with the fact that there are break points, they also should be aware that statistical distribution identification will yield different choices around these break points.

Third, in total (Table 7), there were 10 break points in RSD series and 10 break points in

FIGURE 7
Fitted Value of Kurtosis

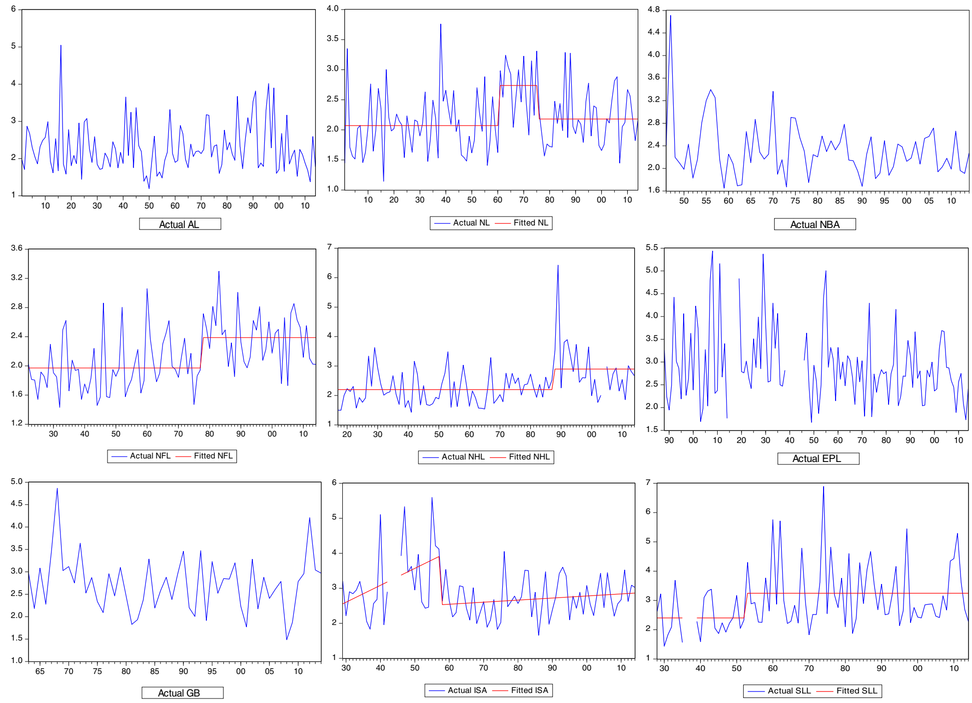


TABLE 9
Normality Rejection Years (Kurtosis Test)

	Playtkurtic	Leptokurtic
AL	1948, 1950, 1977, 1999, 2002, 2007, 2012	1916, 1941
NL	1916, 1956, 1978, 2000, 2007	1938
NBA	1973, 1978, 1990, 1993–1994, 1996, 2007, 2012–2013	1947
NFL	1925, 1944, 1953, 1970, 1973, 1975–1977, 1988, 2003, 2005	—
NHL	2002, 2011	1953, 1989
EPL	1903, 1914, 1949, 1971, 1974, 2013	1892, 1901, 1907–1908, 1911, 1919, 1929, 1933, 1954–1955, 1973, 1984
GB	2008	1968, 2012
ISA	1986	1940, 1947, 1955–1957
SLL	1930, 1935	1953, 1960, 1962, 1968, 1974, 1976, 1981, 1984, 1987, 1994, 1997, 2009–2011

skew and kurtosis series (combined). Since RSD is concerned with the presence (degree) of imbalance, while skew and kurtosis help characterize the way that imbalance occurs, this suggests that both imbalance itself and the *shape* of imbalance from the higher moments have been volatile in the past. But the overriding observation on volatility is that it is, indeed, mostly a thing of the past. Only four break points have occurred during the 2000s, those were all in ESLs, and the most recent was 2010 (GB skew).[16] Thus, just

16. Another point of consistency in break points is for World War II, NL skew, and ISA RSD. Interestingly, as in previous work on break points cited in the text of the paper, *nonoccurrences* are interesting. There are no consistent break points across leagues associated with play suspension during armed conflict (except for ISA) or major economic recession. Except for the break in skew for SLL, 1994, adjacent to the Bosman decision, 1995, no other "free agency" break point is found in any league. There are no break points associated with player unionization or any league policy imposition (drafts, revenue sharing, payroll caps, or payroll taxes).

TABLE 10
Normality Rejection Years (Skewness Test)

	Negative Skewness	Positive Skewness
AL	1916, 1926, 1945, 1991	1941, 1990, 1998
NL	1938, 1961, 1963–1965, 1969	—
NBA	1947	1955
NFL	—	—
NHL	1927, 1935, 1989, 1991–1993, 1999, 2014	1988
EPL	1902, 1911, 1954	1892, 1894, 1901, 1907, 1925, 1927, 1929, 1933, 1935, 1955, 1957, 1960, 1973, 1992, 2003–2004, 2008
GB	1993	1968, 1970, 1972, 1987, 2012, 2014
ISA	1994	1942, 1947–1948, 1950, 1953, 1957, 1963–1964, 1970, 1976, 1987–1988, 2004, 2006–2007
SLL	1942, 1950, 1997	1929, 1960, 1962, 1968, 1974, 1979, 1984, 1986–1987, 1993, 2000, 2009–2013

as careful methods is suggested against spanning break points, and care over distribution identification is suggested, consideration of distributions would do well to include the shape of imbalance dictated by higher moments.

Fourth, it is pretty clear that use of higher moments to capture the complexity of imbalance, and changes in imbalance, should be added to the consideration. The clear difference in kurtosis—leptokurtic for ESLs and platykurtic for NALs—suggests that much of the difference in what fans may care about, translated to actual play on the field, has to do with what happens with the best and worst teams, rather than the behavior around the center of the distribution.

The remainder of our observations about break points concern possible future research topics. No paper can do everything and we do not delve into the questions beyond posing them. However, we do offer some general speculation at the end of this section, again as an aid to investigating these topics.

First, break points in the RSD series, in the three NALs that had any, all improved balance. This is in stark contrast to three of the four ESLs that showed exactly the opposite behavior. In the EPL, GB, and ISA, breaks always coincide with worse balance. In SLL, the breaks were toward better balance but were overcome each time by following, offsetting trends to worse balance. Further, for RSD, all of the break points in NALs that have any occur before 1970 while four of the six break points in European leagues, for RSD, occur after 1970. Explanations for both the timing and the different outcomes vis-à-vis the level of imbalance across leagues awaits future research.

There is a bit more consistency between the leagues on the two continents for break points in the skew time series. In all leagues with breaks, except SLL, skew increased. In SLL, again, while skew increased over time, it did so by overcoming a break in the opposite direction in 1994. Further, the break in skew in the NL was clear back in 1942 while all of the break points in soccer occurred after 1990. Break points in kurtosis, by and large, occur later in NALs. The question for future research is why the increasing impact of the worst teams as the source of imbalance in NALs, compared with the increasing impact of the best teams as a source of imbalance in ESLs?

Finally, some break points coincide with the introduction of the super competitions. Champions League began as the European Cup in 1955–1956, renamed in 1992–1993. Europa League began as the Inter-cities Fairs Cup in 1955–1956, was later renamed UEFA Cup (1971–1972), and finally rebranded with its current name in 2009–2010. Super Cup began in 1972. The first Club World Cup was in 2000, was postponed until 2005, and has been played annually since then. Kurtosis breaks for EPL, ISA, and SLL (all leptokurtic normality violations) occur around the time that Champions League and Europa League were originally founded in 1955–1956. In addition, break points occur coincident with introduction of Super Cup for ISA and SLL (platykurtic violations of normality). This suggested research topic is a bit more fully formed on its own since it involves just why it would be that super competition drove more and increasing inequality in ESLs.

Again, no paper can do everything but we do offer some speculations as an aid to future research. In general, we speculate that the different structures of NALs and ESLs may be responsible for the stark contrasts we find in their winning percentage distributions. NALs operate in closed talent markets, especially compared with ESLs. ESLs also practice promotion and relegation, adding a sense of heightened competition to the bottom teams in the first division and the top teams in the second division. There also are variations in the forms of ownership; true member "clubs" in ESLs but not in NALs. If, as hypothesized throughout the sports economics literature, different ownership forms have different objective functions, then the form of ownership also may influence the distribution of winning percentages.

As noted earlier, another structural difference is that ESLs have true international championships such as the Champions League, Europa League, and Super Cup. NALs proclaim such championships (e.g., the "World" Series in MLB) but they are not truly international. The existence of a true international component to competition by the same teams that play in domestic pro leagues may have influenced the distribution of playing talent and, subsequently, the distribution of winning percentages. The Champions League is relevant to the top-tiered teams in each ESL because only three or four teams in each league advance. Skew and kurtosis might be more sensitive to the creation of the Champions League than RSD would be since increased competition among top-tiered teams is likely to impact the upper tail in winning percentage distribution. Or perhaps it is the particular level of payoff, and the absence of much sharing back to the entire domestic league that drives the talent distribution. Only future work will tell.

Finally, there also is variation in other interesting historical occurrences that may be responsible for the stark contrasts that we find in RSD, skew, and kurtosis comparing NALs and ESLs. In particular, free agency has never coincided with anything interesting in the distribution of winning percentages in NALs. However, the number of breaks occurring shortly after the 1990s, along with all of the other differences in the European leagues, suggests the Bosman ruling in soccer could have altered the distribution of talent in ESLs. Again, our findings suggest a host of interesting future research.

V. CONCLUSION

In this paper, we examine the higher moments of the distribution of winning percentages and discover economic implications of such an examination for nine major sports leagues around the world. The results are useful to both current sports league policy questions and future research.

By-and-large, we fail to reject normality against non-normal stable alternatives but there are rejections. This suggests testing normality in every instance as work on the distribution of winning percentage moves forward. After that, our analysis repeatedly finds stark contrasts in the distribution of winning percentages between NALs and ESLs.

The standard deviation measure appears to have decreased for each of the NALs while the opposite is true for nearly all of the ESLs. Skew results suggest competitive imbalance is due to the weaker teams in NALs but due to stronger teams in ESLs. Finally, while far from the dominant outcome in the leagues we analyzed, violations of normality that do occur are against the platykurtic alternative in NALs but against the leptokurtic alternative in ESLs. In addition to skew, much more of the explanation of winning percentage outcomes will be found in the tails of the distribution for ESLs than for NALs.

Our skewness results have implications for the application of league-wide rules aimed at improving balance. Contrary to current practice in NALs, league policy intervention should be designed to facilitate improvements of weaker teams in NALs. In ESLs, they should be designed to reduce the advantage of the stronger teams.

Additional time series assessment finds structural breaks in the higher moments of the distribution of winning percentages. Interestingly, while few, all of the rejections of normality that we find are associated with break points. Given all of the other stark contrasts in their winning percentage distributions, we speculate that the institutional differences between NALs and ESLs will prove fruitful areas for future research on competitive balance.

REFERENCES

Andrews, D. W. K. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica*, 61(4), 1993, 821–56.

Anscombe, F. J., and W. J. Glynn. "Distribution of the Kurtosis Statistic b 2 for Normal Samples." *Biometrika*, 70, 1983, 227–34.

Ascari, G., and P. Gagnepain. "Spanish Football." *Journal of Sports Economics*, 7(1), 2006, 76–89.

Bai, J., and P. Perron. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica*, 66(1), 1998, 47–78.

———. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics*, 18(1), 2003, 1–22.

———. "Multiple Structural Change Models: A Simulation Analysis," in *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, edited by P. C. B. Phillips, D. Corbae, S. N. Durlauf, and B. E. Hansen. New York: Cambridge University Press, 2006, 212–40.

Cain, L. P., and D. D. Haddock. "Measuring Parity: Tying into the Idealized Standard Deviation." *Journal of Sports Economics*, 7(3), 2006, 330–38.

D'Agostino, R. B. "Transformation to Normality of the Null Distribution of g1." *Biometrika*, 57, 1970, 679–81.

Davies, B., P. Downward, and I. Jackson. "The Demand for Rugby League: Evidence from Causality Tests." *Applied Economics*, 27(10), 1995, 1003–7.

Dawson, A., and P. Downward. "Measuring Short-Run Uncertainty of Outcome in Sporting Leagues." *Journal of Sports Economics*, 6(3), 2005, 303–13.

Depken, C. A. II. "Free-Agency and the Competitiveness of Major League Baseball." *Review of Industrial Organization*, 14(3), 1999, 205–17.

DiNardo, J., and J. Winfree. "The Law of Genius and Home Runs Refuted." *Economic Inquiry*, 48(1), 2010, 51–64.

Doornik, J. A., and H. Hansen. "An Omnibus Test for Univariate and Multivariate Normality." *Oxford Bulletin of Economics and Statistics*, 70(s1), 2008, 927–39.

Eckard, E. W. "Free Agency, Competitive Balance, and Diminishing Returns to Pennant Contention." *Economic Inquiry*, 39(3), 2001, 430–43.

Fama, E. F. "Mandelbrot and the Stable Paretian Hypothesis." *Journal of Business*, 36(4), 1963, 420–29.

———. "The Behavior of Stock-Market Prices." *Journal of Business*, 38(1), 1965, 34–105.

Fama, E. F., and R. Roll. "Some Properties of Symmetric Stable Distributions." *Journal of the American Statistical Association*, 63, 1968a, 817–36.

———. "Parameter Estimates for Symmetric Stable Distributions." *Journal of the American Statistical Association*, 66, 1968b, 331–38.

———. "Parameter Estimates of Symmetric Stable Distributions." *Journal of the American Statistical Association*, 66(334), 1971, 331–8.

Fort, R. "Competitive Balance in North American Professional Sports," in *Handbook of Sports Economics Research*, edited by J. Fizel. Armonk, NY: M.E. Sharpe, Inc, 2006, 190–206.

———. "Comments on 'Measuring Parity'." *Journal of Sports Economics*, 8(6), 2007, 642–51.

———. *Sports Economics*. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2011.

Fort, R., and Y. H. Lee. "Structural Change, Competitive Balance, and the Rest of the Major Leagues." *Economic Inquiry*, 45(3), 2007, 519–32.

———. "Major League Baseball Attendance Time Series: League Policy Lessons," in *The Econometrics of Sports*, edited by P. Rodriguez, S. Kesenne, and J. Garcia. Cheltenham, UK: Edward Elgar, 2013, 35–49.

Fort, R., and J. Maxcy. "Comment: Competitive Balance in Sports Leagues: An Introduction." *Journal of Sports Economics*, 4(2), 2003, 154–60.

Fort, R., and J. Quirk. "Cross-Subsidization, Incentives, and Outcomes in Professional Team Sports." *Journal of Economic Literature*, 33, 1995, 1265–99.

Fort, R., J. Maxcy, and M. Diehl. "Uncertainty by Regulation: Rottenberg's Invariance Principle." *Research in Economics*, 70(3), 2016, 454–67.

Groot, L. *Economics, Uncertainty, and European Football: Trends in Competitive Balance*. Northampton, MA: Edward Elgar, 2008.

Humphreys, B. R. "Alternative Measures of Competitive Balance in Sports Leagues." *Journal of Sports Economics*, 3(2), 2002, 133–48.

Jarque, C. M., and A. K. Bera. "A Test for Normality of Observations and Regression Residuals." *International Statistical Review*, 55(2), 1987, 163–72.

Kolmogorov, A. N. *Foundations of the Theory of Probability*. 2nd ed. New York, NY: Chelsea, 1956. [Original: Kolmogorov, A.N. 1933. Grundbegriffe der Wahrscheinlichkeitsrechnung (in German). Berlin: Julius Springer.]

Krautmann, A. C., and L. Hadley. "Dynasties Versus Pennant Races: Competitive Balance in Major League Baseball." *Managerial and Decision Economics*, 27(4), 2006, 287–92.

Lee, Y. H. "Competitive Balance and Attendance in Japanese, Korean, and U.S. Professional Baseball Leagues," in *International Comparisons in Sports Economics*, edited by R. Fort and J. Fizel. Westport, CT: Praeger, 2004, 281–92.

Lee, Y. H., and R. Fort. "Structural Change in MLB Competitive Balance: The Depression, Team Location, and Integration." *Economic Inquiry*, 43, 2005, 159–69.

———. "Competitive Balance: Time Series Lessons from the English Premiere League." *Scottish Journal of Political Economy*, 59(3), 2012, 266–82.

Lee, Y. H., H. Jang, and S. H. Hwang. "Market Competition and Threshold Efficiency in the Sports Industry." *Journal of Sports Economics*, 16(8), 2015, 853–70.

Lee, Y. H., Y. Kim, and S. Kim. "A Bias-Corrected Estimator of Competitive Balance in Sports Leagues." Working Paper No. 1611, Research Institute for Market Economy, Sogang University, 2016.

Liu, J., S. Wu, and J. V. Zidek. "On Segmented Multivariate Regressions." *Statistica Sinica*, 7, 1997, 497–525.

Mandelbrot, B. "The Variation of Certain Speculative Prices." *Journal of Business*, 36(4), 1963, 394–419.

———. "The Variation of Some Other Speculative Prices." *Journal of Business*, 40(4), 1967, 393–413.

Martins, A. M., and S. Cro. "The Demand for Football in Portugal: New Insights on Outcome Uncertainty." *Journal of Sports Economics*, 19(4), 2018. https://doi.org/10.1177/1527002516661602.

May, J. M. "Extended and Corrected Tables of the Upper Percentage Points of the Studentized Range." *Biometrika*, 39, 1952, 192–93.

Mills, B., and R. Fort. "League-Level Attendance and Outcome Uncertainty in U.S. Pro Sports Leagues." *Economic Inquiry*, 52(1), 2014, 205–18.

Noll, R. G. "Professional Basketball." Studies in Industrial Economics Paper No. 144, Stanford University, 1988.

Owen, P. D. "Limitations of the Relative Standard Deviation of Win Percentages for Measuring Competitive Balance in Sports Leagues." *Economic Letters*, 109(1), 2010, 38–41.

———. "Measuring Parity in Sports Leagues with Draws: Further Comment." *Journal of Sports Economics*, 13(1), 2012, 85–95.

Owen, P. D., and N. King. "Competitive Balance Measures in Sports Leagues: The Effects of Variation in Season Length." *Economic Inquiry*, 53(1), 2015, 731–44.

Owen, P. D., M. Ryan, and C. R. Weatherston. "Measuring Competitive Balance in Professional Team Sports Using the Herfindahl-Hirschman Index." *Review of Industrial Organization*, 31(4), 2007, 289–302.

Quirk, J., and R. D. Fort. *Pay Dirt: The Business of Professional Team Sports*. Princeton, NJ: Princeton University Press, 1992.

Schmidt, M. B., and D. J. Berri. "Competitive Balance and Attendance: The Case of Major League Baseball." *Journal of Sports Economics*, 2(2), 2001, 145–67.

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

Shapiro, S. S., and M. B. Wilk. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika*, 52(3/4), 1965, 591–611.

Szymanski, S. "The Economic Design of Sporting Contests." *Journal of Economic Literature*, 41(4), 2003, 1137–87.

Thadewald, T., and H. Büning. "Jarque–Bera Test and Its Competitors for Testing Normality—A Power Comparison." *Journal of Applied Statistics*, 34(1), 2007, 87–105.

Utt, J., and R. Fort. "Pitfalls to Measuring Competitive Balance with Gini Coefficients." *Journal of Sports Economics*, 3(4), 2002, 367–73.

Winfree, J., and R. Fort. "Nash Conjectures and Talent Supply in Sports League Modeling: A Comment on Current Modeling Disagreements." *Journal of Sports Economics*, 13(3), 2012, 306–13.

Yao, C. C. "Estimating the Number of Change-Points via Schwarz' Criterion." *Statistics & Probability Letters*, 6(3), 1988, 181–89.